May 2023

1st EFORT European Consensus

Medical & Scientific Research Requirements for the Clinical Introduction of Artificial Joint Arthroplasty Devices

Chairs:
Thomas M. Grupp
Søren Overgaard

EFORT Implant & Patient Safety Initiative:
Klaus-Peter Günther

Scientific Committee:
Ashley Blom
Luca Cristofolini
Eduardo Garcia-Rey
Alexander Giurea
Bernd Grimm
Marcus Jäger
Dennis Janssen
Philippe Massin
Francesco Siccardi

Research Assistants:
Sabine Rusch
Ronja A. Schierjott

#EOTEP

Table of Contents

Editorial

Dear Colleagues,

Innovations in Orthopaedics and Traumatology have contributed to the achievement of a high-quality level of care in musculoskeletal disorders and injuries over the past decades. The applications of new implants as well as diagnostic and therapeutic techniques in addition to implementation of clinical research, have significantly improved patient outcomes, reduced complication rates and length of hospital stay in many areas. Further innovations in the fields of artificial intelligence, computer-assisted surgical procedures, augmented-reality-based applications, 3D-printed implants and instruments, new biomaterials, and smart implants are currently being developed which may in addition reduce burden of disease and morbidity associated with surgical treatment in the future.

However, the introduction of new treatment techniques, modified strategies or innovative implants must strictly follow rules of proper evaluation, regulation and training in order to know the results of the initiative, and not to put patients at unnecessary risk.

With the implementation of the new "Medical Device Regulation (MDR)" by the European Parliament, several already existing rules for the introduction of innovative products and instruments have been tightened, and even new rules have been established. As the approval process for medical devices in Europe will become more stringent, compared to the Medical Device Directive (MDD) more robust clinical and pre-clinical data will be required.

The regulatory framework is extensive, but there is a lack of understanding and clarity in daily practice what the meaning of clinical & pre-clinical evidence as required by the MDR is. Thus, understanding and clarity are of utmost importance for introduction of new implants and implant-related instrumentation in combination with surgical technique to ensure a safe use of implants and treatment of patients.

Therefore the "EFORT Patient & Implant Safety Initiative" (IPSI), was launched by the European Federation of National Associations of Orthopaedics and Traumatology (EFORT) with an inaugural workshop on January 21, 2020, in Brussels with participation of a steering group invited by the EFORT Board. The recommendations IPSI WG1 "Introduction of Innovations" are based on the results of the inaugurational workshop and a consecutive Delphi consensus process [Overgaard et al. EFORT Open Reviews 2023]. They provide surgeons, researchers, implant manufacturers as well as patients and health authorities with a consensus of the development, implementation, and dissemination of innovation in the field of arthroplasty. In dialogue with the EFORT Board the chairs of IPSI WG1 have designated a Scientific Committee who have put in place a process for a more intensive and detailed 1st EFORT European Consensus Initiative involving EFORT National Member Societies, European Specialty Societies as well as International Expert Delegates. Due to the ubiquitous and increasing need for arthroplasty surgery first activities concentrated on the topics of new implants and implant-related instrumentation in this area, where innovations have fundamental consequences for patients as well as surgeons and manufacturers. Aim of the consensus is also to build a fundament for quality assured development of innovative therapies and medical devices, which is important for the translation process.

The objectives of the 1st EFORT European Consensus on "Medical & Scientific Research Requirements for the Clinical Introduction of Artificial Joint Arthroplasty Devices" were foremost to focus on patient & implant safety through performance requirements for medical devices in this specific field. The intended key outcomes are consented, practical pathways to maintain innovation and optimisation of orthopaedic products and workflows within the boundaries of MDR 2017/745. Open Access practical guidelines based on adequate, state of the art pre-clinical and clinical evaluation methodologies for the introduction of joint replacements and implant-related instrumentation shall provide hands-on orientation for orthopaedic surgeons, research institutes and laboratories, orthopaedic device manufacturers, Notified Bodies but also for National Institutes and authorities, patient representatives and further stakeholders.

To achieve this, a number of research questions of importance were defined resulting in practical guidelines. The Conference Chairs & Scientific Committee members conducted a Delphi methodology asking groups of international expert delegates to answer the assigned research question(s) based on scientific research (systematic literature reviews), their published extensive expertise and utmost interest [Grupp et al. EFORT Open Reviews 2023]. The resulting Draft Consensus Statements have been circulated for review among all participants in preparation of the final discussion and voting at the 1st EFORT European Consensus Conference taking place June 22-23, 2021, in Dresden.

We would like to acknowledge and thank the Scientific Committee members, all International Expert Delegates, the Delegates from European National & Specialty Societies and the Editorial Team for their outstanding contributions and support during this EFORT European Consensus and hope that during the development, testing, certification and clinical introduction of innovation in the field of arthroplasty the following guidelines will contribute to a better understanding and clarity in daily practice. Enjoy reading!

Yours sincerely,

Klaus-Peter Günther
EFORT President 2020-21
EFORT Past-President
Initiator EFORT Implant & Patient Safety Initiative (IPSI)

Søren Overgaard
Chair EFORT Science Committee &
Chair EFORT Implant & Patient Safety Initiative – Working group I
Chair 1. EFORT European Consensus

Thomas M. Grupp
Chair EFORT Implant & Patient Safety Initiative – Working group I "Introduction of Innovation"
Chair 1. EFORT European Consensus

Scientific Committee

Ashley Blom | United Kingdom
Luca Cristofolini | Italy
Eduardo García-Rey | Spain
Alexander Giurea | Austria
Bernd Grimm | Luxembourg
Marcus Jäger | Germany
Dennis Janssen | The Netherlands
Philippe Massin | France
Francesco Siccardi | Switzerland

International Experts

Hassan Achakri | Switzerland
George Babis | Greece
Francesco Benazzo | Italy
Dario Bergadano | Switzerland
Massimiliano Bernardoni | Italy
Marc Bohner | Switzerland
Barbara Bordini | Italy
Joshua Bridgens | United Kingdom
Vincenzo Carbone | Italy
Justin Cobb | United Kingdom
José Cordero-Ampuero | Spain
Marcel Dreischarf | Germany
Georg N. Duda | Germany
Ingrid Dupraz | Germany
Matthias Fink | Germany
John Fisher | United Kingdom
Faris Haddad | United Kingdom
Alister Hart | United Kingdom
Petra Heesterbeek | The Netherlands
Bernardo Innocenti | Belgium
Volkmar Jansson | Germany
Christian Kaddick | Germany
Franz Kainberger | Austria
Johan Kärrholm | Sweden
Jan Philippe Kretzer | Germany
Ricardo Larrainzar-Garijo | Spain
Christoph Lohmann | Germany
Anne Lübbeke | Switzerland & United Kingdom
Fabio Mancino | Italy
Bernard Masson | France
Richard T. Mayer | Switzerland
Claudia Mazzà | United Kingdom
Michael Morlock | Germany
Rob G.H.H. Nelissen | The Netherlands
Maria Angeles Pérez Ansón | Spain
Gwendolen Reilly | United Kingdom
Claude Rieker | Switzerland
Sabine Rusch | Germany
Maria Luisa Ruspi | Italy
Ronja A. Schierjott | Germany
Christoph Schilling | Germany
Arndt-Peter Schulz | Germany
William R. Taylor | Switzerland
Francesco Traina | Italy
Walter Van der Weegen | The Netherlands
Marco Viceconti | Italy
Henning Windhagen | Germany
Matthias Woiczinski | Germany
Ina Wüstefeld | Germany

Editorial Team

Sabrina Marchal | Switzerland
Ronja Schierjott | Germany
Sabine Rusch | Germany
Susan Davenport | Switzerland
Adrian Ott | Switzerland

Results Overview

Topics and results overview of the 1st EFORT European Consensus on "Medical & Scientific Research Requirements for the Clinical Introduction of Artificial Joint Arthroplasty Devices"

1 | BIOLOGICAL SAFETY / BIOCOMPATIBILITY & STERILITY

Biological Safety Determination of Medical Devices Considering Limited, Prolonged, and Long-term Implantation

R.T. Mayer, D. Bergadano, I. Wüstefeld, M. Bohner

Voting result: 97% | 97% Agree / 3% Disagree / 0% Abstain

2 | PRE-CLINICAL METHODS

Pre-Clinical Testing in the Field of Arthroplasty: Potentials, Limitations and Demands Regarding Test Methodology

L. Cristofolini, T.M. Grupp, C. Kaddick, M. Morlock, M.L. Ruspi, D. Janssen

Voting result: 100% | 100% Agree / 0% Disagree / 0% Abstain

3 | INTERFACE COMPATIBILITY / INTERFACE GEOMETRY

Interface Compatibility of Joint Arthroplasty Devices – Geometrical, Dimensional and Functional Assessment of Implant/Instrument Interfaces

C. Rieker, M. Bernardoni, C. Schilling, M. Woiczinski, J. Bridgens

Voting result: 98% | 98% Agree / 0% Disagree / 2% Abstain

4 | MECHANICAL COMPONENT TESTING THA (static/dynamic)

Methods in Total Hip Arthroplasty Mechanical Component Testing – Discussing Standard and Additional Test Protocols for Contemporary Product Development

M. Bernardoni, L. Cristofolini, J.P. Kretzer

Voting result: 93% | 93% Agree / 2% Disagree / 5% Abstain

5 | MECHANICAL COMPONENT TESTING TKA (static/dynamic)

Part 1

Standard Test Methods in Total Knee Arthroplasty to Establish that the Implant will Withstand the Endurance Habitual and Peak Loads that Must Reasonably be Expected

C. Kaddick, C. Schilling, D. Janssen, J.P. Kretzer

Voting result: 98% | 98% Agree / 2% Disagree / 0% Abstain

Part 2

Additional Test Methods in Total Knee Arthroplasty to Establish that the Implant will Withstand the Endurance Habitual and Peak Loads that Must Reasonably be Expected

C. Kaddick, C. Schilling, D. Janssen

Voting result: 91% | 91% Agree / 3% Disagree / 6% Abstain

Part 3

Additional Numerical Test Methods in Total Knee Arthroplasty to Establish that the Implant will Withstand the Endurance Habitual and Peak Loads that Must Reasonably be Expected

C. Kaddick, C. Schilling, D. Janssen

Voting result: 95% | 95% Agree / 2.5% Disagree / 2.5% Abstain

6 | MECHANICAL COMPONENT TESTING (clinical perspective)

Endurance Habitual and Peak Loads in Hip and Knee Arthroplasty –Review and Summary of Aspects to be Considered

M. Morlock, J.P. Kretzer, R.A. Schierjott, F. Traina, R. Larrainzar-Garijo, G.N. Duda, C. Kaddick

Voting result: 90% | 90% Agree / 3% Disagree / 7% Abstain

7 | BIOTRIBOLOGY (wear simulation, wear debris and biological response)

Biotribological Methodologies in Wear Simulation and Wear Debris Characterisation in Hip and Knee Arthroplasty – Review and Introduction of Thresholds for Standard Testing

T. M. Grupp, C. Kaddick, C. Rieker, J.P. Kretzer, J. Fisher

Voting result: 98% | 98% Agree / 0% Disagree / 2% Abstain

8 | BIOTRIBOLOGY (beyond standard testing)

Biotribological Methodologies in Wear Simulation, Evaluation of Wear Debris Release and Biological Response in Hip and Knee Arthroplasty – Review and Guidance on Additional Methodologies to Complement Standard Testing

T.M. Grupp, C. Kaddick, C. Rieker, J. P. Kretzer, J. Fisher

Voting result: 95% | 95% Agree / 0% Disagree / 5% Abstain

9 | BIOTRIBOLOGY (clinical perspective)

Methodologies to Evaluate Linear Penetration and Volumetric Wear in Clinical Studies - Imaging Techniques for the Evaluation Of Articulations in THA, TKA to Assess Wear and Function of the Joint Throughout the Expected Implant Lifetime

M. Jäger, M. Dreischarf, T.M. Grupp, C. Rieker

Voting result: 93% | 93% Agree / 2% Disagree / 5% Abstain

10 | BIOTRIBOLOGY (clinical follow-up)

Detection of Wear/Debris Complications in Total Hip and Knee Arthroplasty at an Early Follow-up with Regard to Complementary Tests for "Silent Bone Loss" Detection and Risky Modular Implants

E. Garcia-Rey, J. Cordero-Ampuero, G. Babis, F. Benazzo, M. Morlock

Voting result: 88% | 88% Agree / 10% Disagree / 2% Abstain

11 | SIZE RANGE AND ANOTOMICAL DESIGN OF IMPLANTS

Appropriateness of Implant Geometry, Sizing Range and Increments for Reconstruction of Anatomical Structures in Hip and Knee Arthroplasty – Review and Suggestion of Methods for Pre-clinical Evaluation

D. Janssen, M. Bernardoni, I. Dupraz, R.A. Schierjott

Voting result: 98% | 98% Agree / 2% Disagree / 0% Abstain

12 | SIZE RANGE AND ANOTOMICAL DESIGN OF IMPLANTS (clinical perspective)

Appropriateness of Implant Geometry, Sizing Range and Increments for Reconstruction of Anatomical Structures in Hip and Knee Arthroplasty – Review and Suggestion of Methods for Clinical Evaluation

F. Benazzo, B. Grimm, C. Mazzà, F. Mancino, R.A. Schierjott

Voting result: 97% | 97% Agree / 0% Disagree / 3% Abstain

13 | MODULARITIES/INTERFACES

Pre-clinical Assessment of the in vivo Behaviour of Non-articulating Interfaces Between Implant Components Concerning the Consequences of Micro Motion or Corrosion Processes – Test Methodology, and Requirements for Testing and Characterization of Implant Modularities

J.P. Kretzer, T.M. Grupp, C. Kaddick, R.T. Mayer, M. Morlock

Voting result: 92% | 92% Agree / 3% Disagree / 5% Abstain

14 | MODULARITIES/INTERFACES (clinical perspective)

Clinical Assessment of the in vivo Behaviour of Interfaces Between Non-Articulating Implant Components Concerning the Consequences of Micro Motion or Corrosion Processes

F. Traina, M. Morlock, R.T. Mayer, A. Hart

Voting result: 84% | 84% Agree / 5% Disagree / 11% Abstain

15 | IMPLANT FIXATION (cemented implants)

Standard Test Methods for the Pre-Clinical Assessment of Primary and Secondary Stability in Cemented Total Joint Arthroplasty Taking into Consideration Physiological Force Transmission and Long-Term Fixation with Regard to the Clinical Application

L. Cristofolini, T.M. Grupp, V. Jansson, R.T. Mayer

Voting result: 93% | 93% Agree / 7% Disagree / 0% Abstain

16 | IMPLANT FIXATION (cementless implants)

Test Methods for the Pre-Clinical Assessment of Primary and Secondary Stability in Cementless Total Joint Arthroplasty Taking into Consideration Physiological Force Transmission, Stress Shielding and Long-Term Fixation with Regard to the Clinical Application

D. Janssen, C. Schilling, J.P. Kretzer, R.T. Mayer

Voting result: 97% | 97% Agree / 0% Disagree / 3% Abstain

17 | IMPLANT FIXATION (clinical perspective)

How to Assess Primary and Secondary Stability of Orthopaedic Joint Replacement Devices in a Clinical Setting Considering also How to Obtain/ Ensure Optimal Force Transmission into the Underlying Bone

J. Kärrholm, M. Dreischarf, R.T. Mayer, R.G.H.H. Nelissen

Voting result: 90% | 90% Agree / 2.5% Disagree / 7.5% Abstain

18 | IMPLANT FIXATION (clinical methods)

Radiologic Methods and Parameters to Estimate Primary Stability of Implant Fixation to the Bone - Discussing Recommended Methods and Time Points for Evaluating Subsidence/Loosening of Implant Components and Evaluating Implant Fixation Depending on the Implant And Fixation Material

J. Kärrholm, M. Dreischarf, J. Cordero-Ampuero, P. Heesterbeek, R.T. Mayer, R.G.H.H. Nelissen

Voting result: 97% | 97% Agree / 0% Disagree / 3% Abstain

19 | JOINT STABILITY AND KINEMATICS

Assessment of Functional Joint Stability and Movement Performance after Total Joint Arthroplasty Regarding the Ability of an Implant to Enable the Reconstruction of a Functionally Satisfying and Stable Joint, Including an Appropriate Range of Motion and Best Possible Preservation / Restoration of Kinematics

W.R. Taylor, B. Innocenti, G.N. Duda, T.M. Grupp, M. Woiczinski

Voting result: 98% | 98% Agree / 2% Disagree / 0% Abstain

20 | TRANSFERABILITY OF RESULTS (in between devices)

Transferability of Pre-Clinical/Clinical Results of a Product to Another Device Taking into Consideration General Criteria of Equivalence Devices/ Reference Products and Limitations for the Transferability of Results Across Different Variants of One Implant

D. Bergadano, J. Bridgens, T.M. Grupp, A-P. Schulz

Voting result: 92% | 92% Agree / 8% Disagree / 0% Abstain

21 | TRANSFERABILITY OF RESULTS (pre-clinical/clinical)

Transferring Pre-Clinical Results of Joint Replacement Devices into the Clinical Setting

M. Jäger, F. Traina, A. Giurea, T.M. Grupp, S. Rusch

Voting result: 90% | 90% Agree / 2% Disagree / 8% Abstain

22 | EVALUATION OF INSTRUMENTS AND USABILITY

Part 1

Pre-Clinical Evaluation of Instruments and Usability with Regard to Handling, Workflow and Functionality

A. Giurea, F. Benazzo, A. Blom, M. Bernardoni, C. Schilling, F. Traina, R.T. Mayer, S. Overgaard

Voting result: 95% | 95% Agree / 0% Disagree / 5% Abstain

Part 2

Clinical Evaluation of Instruments and Usability with Regard to Handling, Workflow and Functionality

A. Giurea, F. Benazzo, A. Blom, M. Bernardoni, C. Schilling, F. Traina, R.T. Mayer, S. Overgaard

Voting result: 97% | 97% Agree / 0% Disagree / 3% Abstain

23 | MODIFICATIONS/ADJUSTMENTS

Requirements and Considerations for the Implementation of Modifications During the PMCF Phase of a Joint Replacement Device Based on the Functional Relevance of the Adjustment

J. Bridgens, M. Bernadoni, C. Schilling, S. Rusch, P. Massin, R. Larrainzar-Garijo, F. Traina, V. Jansson

Voting result: 95% | 95% Agree / 2.5% Disagree / 2.5% Abstain

24 | PRE-CE STUDIES/SAFETY STUDIES (potentials/limitations)

Potentials and Limitations of a Pre-CE Study (or Safety Study) in the Field of Arthroplasty

A. Blom, D. Bergadano, I. Wüstefeld, J. Cobb, F. Haddad, M. Jäger, H. Achakri, M. Fink, A-P. Schulz

Voting result: 95% | 95% Agree / 2.5% Disagree / 2.5% Abstain

25 | PRE-CE STUDIES/SAFETY STUDIES (study design)

Requirements to Study Design of Pre-CE Studies / Safety Studies

A. Blom, D. Bergadano, I. Wüstefeld, J. Cobb, F. Haddad, M. Jäger, H. Achakri, M. Fink, A-P. Schulz

Voting result: 92% | 92% Agree / 3% Disagree / 5% Abstain

26 | PERIOPERATIVE AND SHORT-TERM POSTOPERATIVE (SERIOUS) ADVERSE EVENTS

Approach for the Pre-Clinical Investigation of Adverse Events or Complications Related to the Clinical Application of Total Joint Arthroplasty Devices and their Implantation Procedure

F. Siccardi, S. Rusch, S. Overgaard, A. Giurea, T.M. Grupp, A-P. Schulz

Voting result: 95% | 95% Agree / 0% Disagree / 5% Abstain

27 | PERIOPERATIVE AND SHORT-TERM POSTOPERATIVE (SERIOUS) ADVERSE EVENTS (clinical perspective)

Approach for the Clinical Investigation of Adverse Events or Complications Related to the Clinical Application of Total Joint Arthroplasty Devices and their Implantation Procedure

F. Siccardi, S. Rusch, S. Overgaard, A. Giurea, T.M. Grupp, A-P. Schulz

Voting result: 92% | 92% Agree / 3% Disagree / 5% Abstain

28 | REVISION RATE/SURVIVAL TIME

Assessing Revision Rates, Lifetime and Survival Time of Total Joint Arthroplasty Implants with Regard to Benchmark Values, Pre-CE Studies, Influencing Factors and Relevant Parameters

V. Jansson, A. Blom, B. Bordini, A. Lübbeke

Voting result: 97% | 97% Agree / 0% Disagree / 3% Abstain

29 | PMCF (post-market clinical follow-up) STUDIES

Potentials and Limitations of PMCF Studies in the Field of Total Joint Arthroplasty – Discussing Study Design, Parameters, and Alternatives

A. Lübbeke, H. Achakri, D. Bergadano, J. Bridgens, I. Wüstefeld, M. Jäger, R. Larrainzar-Garijo, H. Windhagen, P. Massin, E. Garcia-Rey

Voting result: 95% | 95% Agree / 0% Disagree / 5% Abstain

30 | REGISTRY STUDIES

Potentials of Registry Studies in Total Joint Arthroplasty Considering Quality and Quantity of Data, Relevant Parameters and their Potential to Increase the Exploratory Power of Pre-CE Studies

V. Jansson, A. Blom, B. Bordini, S. Overgaard, R.G.H.H. Nelissen

Voting result: 95% | 95% Agree / 2.5% Disagree / 2.5% Abstain

31 | FUNCTIONALIZED IMPLANTS/BIOMATERIALS/SURFACES/INNOVATIONS

Functionalized Surfaces or Novel Aspects in Hip and Knee Arthroplasty – Review and Proposal of a Stepwise Analysis Approach

G.N. Duda, M. Jäger, B. Masson, E. Garcia-Rey, M.A. Pérez Ansón, G. Reilly, R.A. Schierjott

Voting result: 89% | 89% Agree / 0% Disagree / 11% Abstain

32 | IN SILICO TRIALS (Big Data Analytics, Machine Learning, System Biology models, system physiology models)

Part 1

In Silico Trials Methodologies within the Development, Pre-Clinical Assessment and Clinical Evaluation Process of Total Joint Arthroplasty Implants

M. Viceconti, B. Grimm, W. Van der Weegen, F. Traina, I. Wüstefeld, C. Mazzà, M. Dreischarf, C. Lohmann

Voting result: 97% | 97% Agree / 0% Disagree / 3% Abstain

Part 2

In Silico Trials Methodologies within the Development, Pre-Clinical Assessment and Clinical Evaluation Process of Total Joint Arthroplasty Implants, Instruments or Procedures

M. Viceconti, B. Grimm, W. Van der Weegen, F. Traina, I. Wüstefeld, C. Mazzà, M. Dreischarf, C. Lohmann

Voting result: 95% | 95% Agree / 0% Disagree / 5% Abstain

NOTE:
Please note that the following research questions have not been addressed:
Biological safety/biocompatibility & sterility (clinical perspective)
Joint stability & kinematics (clinical perspective)
Functional result/clinical outcome
One research topic was withdrawn during the Conference due to quality reasons.

RESEARCH TOPIC

BIOLOGICAL SAFETY / BIOCOMPATIBILITY & STERILITY

Research Topic 1

Biological Safety Determination of Medical Devices Considering Limited, Prolonged, and Long-term Implantation

RESEARCH TOPIC 1

Biological Safety Determination of Medical Devices Considering Limited, Prolonged, and Long-term Implantation

Authors

Richard T. Mayer[1], Dario Bergadano[2], Ina Wüstefeld[3], Marc Bohner[4]

[1]Johnson & Johnson, Office of the Chief Medical Officer New Brunswick NJ, Zuchwil, Switzerland

[2]Medacta International SA, Castel San Pietro, Switzerland

[3]Aesculap AG Medical Scientific Affairs, Tuttlingen, Germany

[4]Robert Mathys Foundation RMS, Bettlach, Switzerland

1. Abstract

Question: Biological safety / Biocompatibility & Sterility

How can the biological safety of a final finished medical device with limited, prolonged, and long-term implantation be established (including potential degradation products and novel materials/indications)?

Summary/Recommendation:

All biological evaluations shall be carried out within a risk management framework according to ISO 10993-1:2018. Clinical history of safe use data/post-market surveillance data should be utilised as input to mitigate the need to perform biological/chemical testing.

Biological evaluation within a risk management procedure does not mean testing of the device is required. Further, established marketed products are not mandated to be retested due to changes within ISO 10993-1:2018 (e.g. a marketed device was cleared in a past revision of ISO 10993-1:2018, there is no need to perform testing when a new ISO 10993-1:2018 revision is released).

Leveraging acceptable biological safety data from predicate devices for new device development is an acceptable strategy to perform a biological evaluation within a risk management framework on a new device, granted the predicate is an acceptable comparator. For leveraging to be acceptable, appropriate equivalence in material(s) of construction and manufacturing processes/aids must be established. It must be confirmed that the predicate device is a worst case/representative device/coupon in comparison to the new device to be evaluated.

If testing is carried out as part of a biological evaluation, utilise worst case representative final finished device/coupon.

2. Level of Evidence

N/A

3. Consensus Delegate Vote

97% – unanimous, strongest consensus (97% agree / 3% disagree / 0% abstain)

4. Graphical Abstract

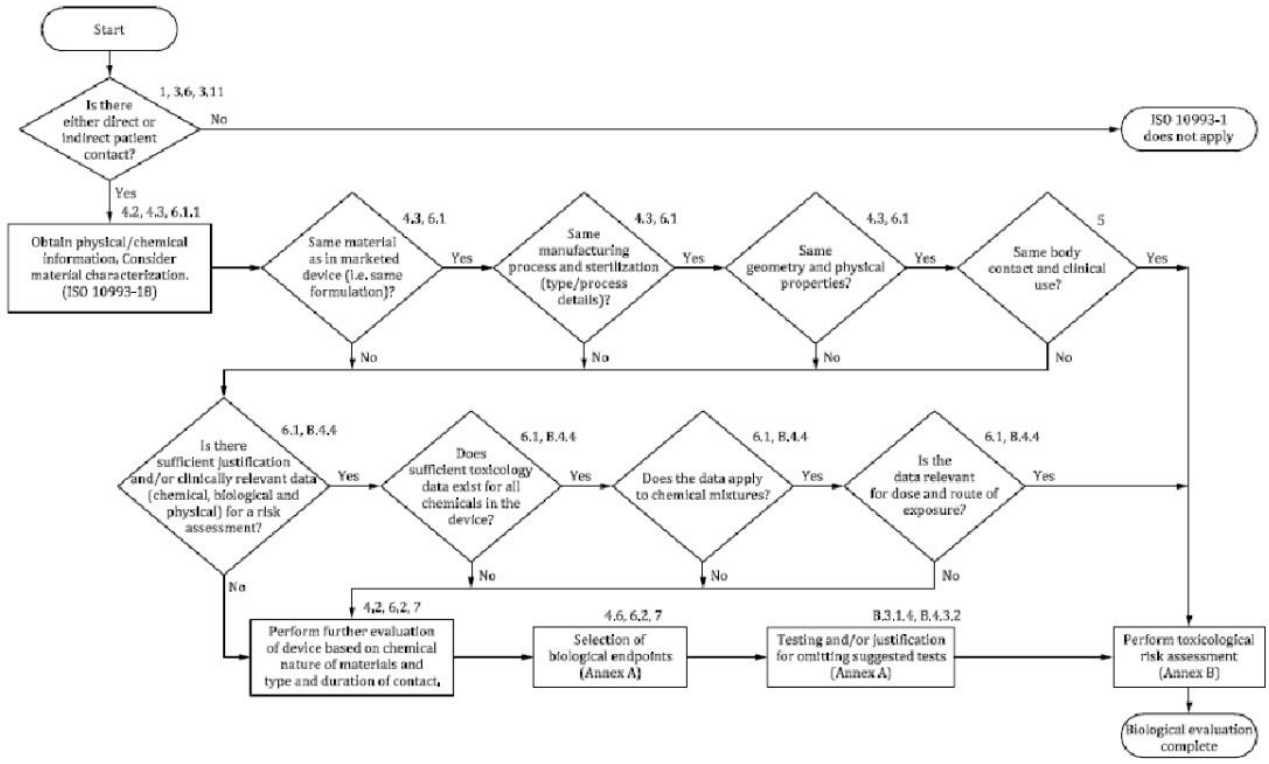


Figure 1: Biological evaluation flowchart (Figure 1 from ISO 10993-1:2018) to be utilised when determining the activities/evaluation needed to be performed to qualify a device as biocompatible.

5. Search Strategy

Not applicable. Question relates solely to current ISO standards.

6. Rationale

1. Biological Evaluation according to ISO 10993-1:2018

Biological Evaluation of medical devices should be carried out according to ISO 10993-1:2018. ISO 10993-1:2018 is the overarching state of the art standard utilised and recognised worldwide to assess medical devices for their biological safety. The standard delineates endpoints to be considered for the biological safety evaluation of medical devices as shown in Annex A. However, selecting an appropriate testing battery requires professional judgement and is not a simple check-box exercise.

Further, ISO 10993-1:2018 shall not be used to mandate re-testing of historical products assessed previously using the appropriate edition of this document at the time of the assessment.

ISO 10993-1:2018 has language throughout guiding medical device manufacturers on how they should be assessing their products for biological safety.

2. Biological Evaluation for Medical Device manufacturers according to ISO 10993-1:2018

At a Medical Devices manufacturer, biological evaluation is an integral part of R&D product development and is performed in accordance with the risk management approach as outlined in ISO 10993-1:2018. Risk assessments are based on the categorisation of the devices, as well as the level of historical information with regard to the device (e.g. existing biocompatibility testing, clinical history of safe use, characterisation of materials of composition, certain manufacturing controls, etc.) in order to determine if additional chemical and biological testing is required within the overall biological evaluation. Manufacturing steps such as, but not limited to, passivation, anodisation, and electropolishing, performed in accordance with established industry standards (e.g. passivation per ASTM F86-13) are acceptable manufacturing controls to waive endpoint testing, as long as no additional manufacturing materials are in contact with the device subsequent to the manufacturing step performed.

3. Example scenarios

Depending on the considerations above, biological evaluations will take a variety of forms. The examples outlined below are a few ways that a Medical Devices manufacturer can evaluate metal alloy implants and instruments to establish an acceptable level of biological risk. The first step of a biological evaluation within a risk management process is to obtain physical and chemical information according to the recommendations of ISO 10993-18:2020. When the medical device to be evaluated is comprised of well-established materials (e.g. through literature/clinical data gathering), such as metal alloys with a long history of safe clinical use, and well-established manufacturing processes, chemical, and biological testing of a worst case/representative part/coupon may not be required.

Per ISO 10993-1:2018, the endpoints to be considered for the biological safety evaluation of external communicating devices having limited (≤ 24 hours) contact with tissue/bone/dentin are cytotoxicity, sensitisation, material mediated pyrogenicity, acute systemic toxicity, and irritation or intracutaneous reactivity. The endpoints to be considered for the biological safety of implant devices with long term (> 30 days) contact with patient's tissue/bone are cytotoxicity, sensitisation, irritation or intracutaneous reactivity, material mediated pyrogenicity, acute systemic toxicity, subacute and sub-chronic toxicity, chronic toxicity, implantation effects, genotoxicity, and carcinogenicity. Certain manufacturing controls can be used as support to justify and waive required biological endpoint testing. Examples include (but are not limited to):

- Validated/qualified manufacturing procedures via manufacturing process flows/manufacturing contact material lists/aids.
- Cleaning validation of the devices that contain testing with acceptance criteria.
- The inclusion of surface finishing manufacturing processes (such as passivation, electropolishing, anodisation) during manufacturing limits the device associated residue for chemicals used prior to the process and therefore mitigates the risk from such manufacturing chemicals. The above manufacturing processes conducted in accordance with a recognised consensus, is an acceptable justification to support biocompatibility of an implant or instrument if no additional chemicals are used after surface finishing.

Example 1: Implant device, long-term tissue/bone contacting. New product development, with established materials and geometrical configurations:

- Following the ISO 10993-1:2018 flowchart and considering the device as an equivalent device to an already marketed product, a biological evaluation can be conducted with no additional biological testing required. The biological evaluation should include (but is not limited to):
    - Chemical characterization according to ISO 10093-18:2020
    - Demonstration of equivalence of ISO 10093-18:2020 results with those of the comparator device
    - Risk assessment of equivalence of materials of construction and manufacturing/processing aids used for comparator device compared to proposed device
    - History of safe clinical use data (of the comparator device)
    - Any previous biocompatibility data (of the comparator device)

Conclusion: No biological testing required, based on compliance with ISO 10993-1:2018 figure 1 flowchart, and including proper justifications for waiving biological endpoint testing (as shown below in Section 4).

Example 2: Implant device, long-term tissue/bone contacting. Marketed device. Resubmission due to a change in regulations (e.g. MDR)

- Following the ISO 10993-1:2018 flowchart (Fig. 1) and considering the device is a marketed product, a biological evaluation could be conducted with no additional biological testing required. However, if there are insufficient controls in place and missing/insufficient details of manufacturing such as processing aids used, or missing/incomplete assessment of biocompatibility endpoints, the following approach could be suggested:
    - Select an appropriate worst case/representative part/coupon for testing to accommodate a large portfolio of parts in scope
    - Chemical characterisation according to ISO 10993-18:2020
    - Toxicological risk assessment in accordance with ISO 10993-17:2002
    - Cytotoxicity testing according to ISO 10993-5:2009, if necessary (as a sensitivity indicator of potential toxicity)
    - History of safe clinical use data (product specific released post market surveillance reports, and/or risk management reports)

If the Toxicological Risk Assessment (TRA) determines acceptable amounts of analysed compounds found from a device, and cytotoxicity testing in accordance with ISO 10993-5:2009 meets acceptance criteria, no further endpoint testing should be necessary. If TRA shows unacceptable amounts, further biological endpoint testing may be required and/or refined chemical characterisation may be required to confirm biological safety.

Conclusion: Biological/chemical characterisation testing required, based on compliance with ISO 10993-1:2018. Certain biological endpoint testing can be waived utilising the justifications listed in section 4.

Example 3: Externally communicating device (Instrument) with limited tissue/bone contact. New product development, with established materials and geometrical configurations:

- Following the ISO 10993-1:2018 flowchart and considering the device as an equivalent device to an already marketed product, a biological evaluation can be conducted with no additional biological testing required. The biological evaluation should include (but is not limited to):
    - Chemical characterization according to ISO 10093-18:2020
    - Demonstration of equivalence of ISO 10093-18:2020 results with those of the comparator device
    - History of safe clinical use data (of the comparator device),
    - Any previous biocompatibility data (of the comparator device)
    - Leveraging of manufacturing controls (e.g. cleaning validation, process validation evaluations)

Conclusion: No testing required, based on compliance with ISO 10993-1:2018 figure 1 flowchart, and including proper justifications for waiving biological endpoint testing (as shown below in Section 4).

Example 4: Externally communicating device (Instrument) with limited tissue/bone contact. Marketed device. Resubmission due to a change in regulations (e.g. MDR)

- Following the above flowchart and considering the device is a marketed product, a biological evaluation could be conducted with no additional biological testing required. However, if there are insufficient controls in place and missing/insufficient details of manufacturing such as processing aids used, or missing/incomplete assessment of biocompatibility endpoints, the following approach could be suggested:
    - Select a proper worst case/representative part/coupon for testing to accommodate a large portfolio of parts in scope
    - Chemical characterisation in accordance with ISO 10993-18:2020
    - Toxicological assessment according to ISO 10993-17:2002
    - Cytotoxicity testing according to ISO 10993-5:2009, if necessary (as a sensitivity indicator of potential toxicity)
    - History of safe clinical use data (product specific post market surveillance reports, and/or risk management reports)

Conclusion: Biological/chemical characterisation testing required, based on compliance with ISO 10993-1:2018. Certain biological endpoint testing can be waived utilising the justifications listed in section 4.

4. Additional information for biological evaluation according to ISO 10993-1:2018

According to Section 3 above, the following justifications for waiving endpoint testing can be utilised:

- The ISO guidance recommends minimising animal testing whenever biological safety can be established from other lines of evidence, such as in vitro or chemical tests. For example, Section 4.4 of ISO 10993-2:2006 states, "*For the purposes of the ISO 10993 series, animal tests shall only be deemed to be justified … when no suitable scientifically validated test method not involving the use of living animals is reasonably and practically available; and when relevant reduction and refinement strategies have been identified and implemented including, if appropriate, obtaining test data from manufacturers and suppliers, and literature searches for toxicity and biocompatibility data.*" Section 6.2 of ISO 10993-1:2018 states, "*Additional in vivo testing shall not be carried out where the existing non-clinical and clinical data, including history of safe use, meet the requirements of biological evaluation and therefore further animal testing would be unethical.*"

- Section 4.4 of ISO 10993-1:2018 states that, "*Testing is usually not necessary when sufficient information is already available to perform a risk assessment of the material and/or the medical device … biological testing is usually not necessary, if material characterisation (e.g. physical and chemical) demonstrates equivalence to a previously assessed medical device or material with established safety.*"
- Section 4.10 of ISO 10993-1:2018 states that, "*The biological evaluation shall take into account preclinical tests, clinical investigations, post-market experience from similar medical devices or materials, and other relevant information.*" This information can be leveraged as described in Section 4.11: "*Where recommendations for endpoint assessment per Annex A are different from prior published versions of this document, a history of safe clinical use can be used to document why additional testing on a commercially-marketed medical device is not needed.*"
- Regarding the new addition of tests to ISO 10993-1:2018, a rationale for considering acute systemic toxicity and material-mediated pyrogenicity testing is provided in Annex A.2, which states, "*extractables/leachables can be introduced … to the systemic circulation, lymphatic system, …*". Note that, for acute systemic toxicity, adverse effects are defined in ISO 10993-11:2017 as those which are immediate (within 72 hours) and that produce a systemic response from only a single dose or an exposure lasting ≤ 24 hours.
- Regarding the relevance of testing for material-mediated pyrogenicity (i.e., non-endotoxin related), Annex G of ISO 10993-11:2017 states, "*It is not necessary to test all new medical devices for in vivo pyrogenicity. However, materials containing substances that have previously elicited a pyrogenic response, and/or new chemical entities where the pyrogenic potential is unknown should be evaluated for material-mediated pyrogenicity.*" Annex G enumerates a list of currently known substances that have been shown to elicit a pyrogenic response, including endogenous pyrogens, prostaglandin, inducers, drug substances that disrupt thermoregulatory centers, organic chemicals that act as uncoupling agents of oxidative phosphorylation, certain naphthylamines, bacterial endotoxins, neurotransmitters, and certain nickel salts.
- The characterisation of extractables and leachables (ISO 10993-18:2020), together with a toxicological risk assessment (ISO 10993-17:2002) may be utilised in lieu of conducting biological tests for sub-chronic and chronic systemic toxicity, genotoxicity, and carcinogenicity. To this effect, ISO 10993-1:2018, Section 4.3 states, "*Chemical characterisation with an appropriate toxicological threshold can be used to determine if further testing is needed (see Annex B, ISO 10993-17:2002 and ISO 10993-18:2020).*" Section 6.3.1, Subpart b) also states, "*The choice of test procedures shall take into account: … 4) that certain biological tests (i.e. those designed to assess systemic effects) are not justifiable where the presence of leachable chemicals has been excluded (in accordance with ISO 10993-18:2020), or where chemicals have a known and acceptable toxicity profile, allowing the safe use by evaluation in accordance with ISO 10993-17:2002 and risk assessment in accordance with ISO 14971:2019.*" Annex B, Section B.4.3.2 states, "*In the following circumstances, a correctly conducted risk assessment can provide justification for not carrying out long-term testing, where the nature and extent of exposure confirms that a patient is being exposed to very low levels of substances, below relevant toxicological thresholds.*"

How are implant and instrument materials affected by the different sterilisation methods and what must therefore be considered, also regarding the testing of the products?

- The bulk properties of metal alloys are not fundamentally impacted/changed by sterilisation methods such as steam, radiation, and ETO, so the mechanical properties would not be fundamentally impacted/changed. However, the surface properties may be affected (e.g. oxide layer thickness of Ti alloys), which has a direct influence on the biological properties.
- Sterilisation of polymeric materials may cause degradation requiring additional endpoint evaluation as per ISO 10993-1:2018 (See answers to questions 1 and 2 above for typical medical device biological evaluation).

7. References

DIN EN ISO 10993-1: Biological evaluation of medical devices – Part 1: Evaluation and testing within a risk management process (ISO 10993-1:2018); International Organisation for Standardisation: Geneva, Switzerland, 2018.

DIN EN ISO 10993-2: Biological evaluation of medical devices – Part 2: Animal welfare requirements (ISO 10993-2:2006); International Organisation for Standardisation: Geneva, Switzerland, 2006.

DIN EN ISO 10993-5: Biological evaluation of medical devices – Part 5: Tests for in vitro cytotoxicity (ISO 10993-5:2009); International Organisation for Standardisation: Geneva, Switzerland, 2009.

DIN EN ISO 10993-11: Biological evaluation of medical devices – Part 11: Tests for systemic toxicity (ISO 10993-11:2017); International Organisation for Standardisation: Geneva, Switzerland, 2017.

DIN EN ISO 10993-17: Biological evaluation of medical devices – Part 17: Establishment of allowable limits for leachable substances (ISO 10993-17:2002); International Organisation for Standardisation: Geneva, Switzerland, 2002.

DIN EN ISO 10993-18: Biological evaluation of medical devices – Part 18: Chemical characterization of medical device materials within a risk management process (ISO 10993-18:2020); International Organisation for Standardisation: Geneva, Switzerland, 2020.

DIN EN ISO 10993-19: Biological evaluation of medical devices – Part 19: Physico-chemical, morphological and topographical characterization of materials (ISO 10993-19:2020); International Organisation for Standardisation: Geneva, Switzerland, 2020.

DIN EN ISO 14971: Medical devices – Application of risk management to medical devices (ISO 14971:2019); International Organisation for Standardisation: Geneva, Switzerland, 2019.

RESEARCH TOPIC

PRE-CLINICAL METHODS

Research Topic 2

Pre-Clinical Testing in the Field of Arthroplasty: Potentials, Limitations and Demands Regarding Test Methodology

RESEARCH TOPIC 2

Pre-Clinical Testing in the Field of Arthroplasty: Potentials, Limitations and Demands Regarding Test Methodology

Authors

Luca Cristofolini[1], Thomas M. Grupp[2,3], Christian Kaddick[4], Michael Morlock[5], Maria Luisa Ruspi[1], Dennis Janssen[6]

[1]Department of Industrial Engineering, School of Engineering and Architecture, Alma Mater Studiorum – Università di Bologna, Bologna, Italy

[2]Aesculap AG Research & Development, Tuttlingen, Germany

[3]Ludwig Maximilians University Munich, Department of Orthopaedic and Trauma Surgery, Musculoskeletal University Center Munich (MUM), Campus Grosshadern, Munich, Germany

[4]EndoLab Mechanical Engineering GmbH, Riedering, Germany

[5]Hamburg University of Technology, Hamburg, Germany

[6]Radboud University Medical Center, Orthopaedic Research Lab, Research Institute for Medical Innovation, Nijmegen, the Netherlands

1. Abstract

Questions: Pre-clinical methods

What are potentials and what are limitations of pre-clinical testing in the field of arthroplasty?

1. Which parameters can actually be assessed/evaluated with regard to e.g. device testing (mostly standardized) and functional testing of the implant (e.g. load transfer, implant stability etc)?
2. Which are the known aspects and failure modes to be tested for providing sufficient pre-clinical evidence?
3. Are there additional alternatives or are there additional ways to generate supplementary data/information?

Which demands must the test methodology of pre-clinical testing in the field of arthroplasty meet?

1. How is the evaluation of safety and performance of an implant possible at an early stage in the development process?
2. With the introduction of a new design principle or material, should there be a risk assessment of potentially new failure modes and consequently additional testing required for the highest risk mode(s)?
3. Which number of specimens is reasonable in pre-clinical testing and how should the number be determined?

Summary/Recommendation:

The main reasons of non-septic implant failure according to implant registries, can be assessed pre-clinically to different extents:

- Aseptic loosening: reliable non-standardized in vitro protocols and numerical methods exist. They should be part of the pre-clinical testing.
- Periprosthetic post-operative bone fracture are related to notching and stress concentration. In vitro testing is possible and should be performed wherever this concern exist.
- Joint instability, component mispositioning and malalignment depend significantly also on the patient and the surgeon. While some numerical simulations are applicable, they cannot be expected to completely overcome this problem.

Failure of components accounts for a minority of revisions:

- Fatigue fracture has been drastically reduced by application of the standard tests, which seem to be effective.
- Wear still affects many implant types. The standard tests are poorly reproducible

Two aspects should be considered in the early stages of development:

- Prevention of foreseeable failure scenarios: these should be identified with a formal analysis of the failure tree (e.g. FMEA) and each failure mode should be assessed either with a numerical simulation, or with an experiment, or both.
- Verification of the claimed strengths: again, if a device is intended to overcome limitations of existing ones, this claim should be tested.

While indications are given for the required sample size in standardized testing, non-standard tests require different sample sizes, that should be estimated based on test uncertainty, least significant difference, and with the general aim.

As failure of a THR or a TKR is a relatively rare event, the possible risks associated with extreme cases, a sensitivity analysis should be performed whenever concerns exist about a specific design. A combination of in vitro testing and in silico simulations is often the best way to obtain reliable and extensive information. It is important that such in silico models are validated on relevant datasets.

2. Level of Evidence

Moderate

3. Consensus Delegate Vote

100% – unanimous, strongest consensus (100% agree / 0% disagree / 0% abstain)

4. Graphical Abstract

5. Search Strategy

The search was carried out in three phases:
1) In the preliminary phase, the incidence of the different loosening failure scenarios was investigated. The Hip and Knee registries of arthroplasty were interrogated. Only registries meeting the following criteria were considered:

- Covering a population of at least 4 Million citizens
- Including at least 1 000 revision cases
- Follow up of at least 4 years

2) The ISO and ASTM database were interrogated for relevant test methods.
3) To investigate the published test methods, journal papers were searched, using PubMed (National Library of Medicine)

Search string:

(pre-clinical OR preclinical OR in vitro OR biomechanical) (testing OR assessment OR validation) AND ("hip" OR "knee") AND ("implant" OR "prosthesis" OR "stem" OR "acetabulum" OR "tibial" OR "femoral" OR "patellar" OR "device")

Timeframe: ALL (1920-2020)

While the entire database, was searched, here only a selection of the relevant ones is presented.

6. Rationale

Which are the known aspects and failure modes to be tested for providing sufficient pre-clinical evidence?

An analysis of the implant registries (Table 1 for the hip, and Table 2 for the knee) helps identify the most relevant failure modes and scenarios.

Three main categories for "failure" can be defined:

(1a) component failure due to bulk material failure: strength can be assessed by mechanical testing but since we do not know the loading in the body, absolute guide values are difficult to establish.

(1b) component failure due to wear and aging – sophisticated simulators are available. However, many possible loading configurations exist. There should be a consensus on a few loading regimes (walking, stumbling?, stair climbing? resting?). Problem of joint fluid composition and simulation of aging)

(2) interface failure: the interaction with the host bone (load transfer, interface stress, micromotions) can be assessed by animal studies (uncemented) or laboratory tests (cemented and uncemented)

(3) "process failure" (e.g. dislocation, infection, intra-operative peri-prosthetic fracture), which are not really directly device specific.

Apart from sepsis, the registries report the highest incidence for:

- aseptic loosening of the different components (both THR and TKR)
- joint instability (both THR and TKR)
- periprosthetic post-operative bone fracture (mainly THR)
- component mispositioning and malalignment (mainly TKR).

Failure of one of the components accounts for a minority of revisions, mainly due to:

- Wear (both THR and TKR)
- Fatigue fracture (mainly THR)

All these failure modes potentially affect contemporary joint arthroplasty and should be tested pre-clinically.

While component failure represents a minority of revision cases, standardized testing of the individual components seems to be adequate for the current concepts of THR and TKR. Conversely, the high incidence of failures related to bone-implant interaction and to surgery suggest that more pre-clinical testing should be extended to cover such functional aspects.

Which parameters can actually be assessed/evaluated with regard to e.g. device testing (mostly standardized) and functional testing of the implant (e.g. load transfer, implant stability etc)?

The different modes of failure of the device in itself for THR are covered by current standards and mandatory tests (Table 3):

- Fatigue failure of the hip stem, neck and head are addressed by the different parts of the ISO 7206, both for monoblock and modular versions;
- Connection of modular hip stems is addressed by ASTM F2580-18
- The method to test the deformation of the acetabular component is described in ISO 7206-12
- Methods for testing static and fatigue strength of the femoral head are detailed in ISO 11491 and in ASTM F2345-03.
- The method to test adhesion strength of hydroxyapatite are described in ISO 13779-4 and ASTM F1147-05
- Methods to test wear of the head and socket are described in ISO 14242 and ASTM F2025. It must be noted that in this field diverging results are often reported, suggesting that the standard is not adequately reproducible.

The modes of failure of the device for TKR are covered by current standards and mandatory tests (Table 4):

- Fatigue failure of the tibial tray should be tested according to ISO 14789 and ASTM F1800 and F3140 (both bi-condylar and unicondylar)
- Wear of the meniscal bearing is addressed by ISO 14243 (wear of the patellar component is not addressed). It must be noted that in this field diverging results are often reported, suggesting that the standard is not adequately reproducible.

Functional testing is not covered by any standards, due to the complexity of standardizing a test that includes an anatomical specimen, and the surgical preparation. However, the scientific literature offers a series of in vitro and in silico methods to assess the interaction between the THR implant and the host bone (Table 3):

- Methods to test the primary stability of uncemented hip stems have been published since the 1990ies ((1)(2)(3)). Due to the complexity of modeling the viscoelastic nature of bone, interface fit, and friction, the most reliable method is definitely in vitro testing (as opposed to numerical modeling). There is an agreement about the importance of torsional loading (4).
- Both in vitro (5–9) (10) and in silico ((11)(12–14)) methods have been published to assess the long term risk of loosening of cemented hip stems. In this case, more severe loading and large numbers of cycles must be taken into account. The importance of such in vitro tests relates to the fact that actual interface properties, cement fatigue behaviour etc. are incorporated. In silico models (if quantitatively validated (15–17) can be used to assess magnitudes that are difficult to measure experimentally (e.g. interface stress) and explore other scenarios such as simulated tissue adaptation.
- Methods to test the primary stability of uncemented acetabula have been developed. Simplified tests (e.g. based on polyurethane foam blocks or other bone surrogates (18,19) (20,21) are suitable only in a first instance, for preliminary exploration. Because of the complex viscoelastic nature of bone, interface fit, and friction, the most reliable method is definitely in vitro testing of cadaveric specimens ((22–27)). Validated numerical model con provide additional insights (28).
- Peri-prosthetic bone resorption due to stress shielding was deemed a possible failure mode until the late 1990ies and was initially investigated with the aid of numerical models (29,30). Some concerns can still exist for the stiffest stem designs (e.g. revision uncemented hip stems). Reliable methods to measure the alteration of bone strains due to implant have been published (31–34).
- Post-operative peri-prosthetic bone fractures in some case occur. Due to the complex nature of this phenomenon (which depends on bone preparation, implant design, implant positioning), actual implants should be tested (33,35)).

Similarly, in vitro methods have been published to the assessment of the functional performance of TKR implants (Table 4):

- The long-term stability of the femoral component should be tested including the complex kinematics of the knee, with a loading protocol that can be adapted from the ISO14243 protocol for wear testing ((36,37)).
- The load transfer of the tibial component can be measured both in vitro (38–41) and simulated numerically (42–46)17 strain rosettes were attached to a composite tibia (model 3101, Pacific Research Laboratories, Vashon, Washington, USA.

Are there additional alternatives or are there additional ways to generate supplementary data/ information?

As failure of a THR or a TKR is a relatively rare event, it is clear that such failures are associated with circumstances that deviate from the "average implant in the average patient". To assess the possible risks associated with extreme patients (e.g. high BMI, highly osteoporotic) a sensitivity analysis should be performed whenever concerns exist about a specific design. Similarly, failures caused by sub-optimal implantation (e.g. implant mispositioning, improper use of acrylic cement, over/under-reaming) are unlikely to be detected in a Phase I or Phase II clinical trial, as highly expert and trained surgeons would be recruited: such failures occur when an implant undergoes widespread use, in a large number of centers. To detect such risks pre-clinically, the effects of all foreseeable surgical errors should be evaluated. The most flexible tool in this case are in silico simulations, which allow cost-effective exploration of multiple factors (e.g. Monte Carlo simulation (47) (48,49) (50)). It is important that such in silico models are extensive validated on relevant datasets.

How is the evaluation of safety and performance of an implant possible at an early stage in the development process?

While testing of physical specimens is advantageous (and in some aspects mandatory) in the last phase of the pre-clinical validation, no strict rule applies to the early stages of development. However, two aspects should be considered:

- Prevention of foreseeable failure scenarios: these should be identified with a formal analysis of the failure tree (e.g. FMEA) and each failure mode should be assessed either with a numerical simulation, or with an experiment, or both.
- Verification of the claimed strengths: again, if a device is intended to overcome limitations of existing ones, this claim should be with a combination of in vitro testing and in silico simulations.

Examples of design optimization based on a combination of numerical modeling and in vitro testing are present in the literature (28,48,51,52)(16,53)).

With the introduction of a new design principle or material, should there be a risk assessment of potentially new failure modes and consequently additional testing required for the highest risk mode(s)?

New concepts for existing devices or use of innovative materials represent an additional complication with respect to the state-of-the-art as new risks and new modes of failures could apply. Dramatic examples include the 3M Capital stem (54) and the ASR metal-on-metal resurfacing (55),(56).

The only viable option to try and detect such new failure mechanisms is to perform an extensive preclinical test campaign mainly using in vitro tests (this is the only way to incorporate the anatomy, natural materials, interface properties etc). In all cases, unforeseeable failure modes might be missed as experimental and numerical models can only focus on expected scenarios. For this reason, very cautions stepwise clinical testing should be included.

Which number of specimens is reasonable in pre-clinical testing and how should the number be determined?

For pre-clinical testing of the component itself (e.g. fatigue, or wear) the standardized tests in most cases prescribe the sample size required. Indeed, there are several ISO and ASTM standards as well as FDA guidance documents which define for fatigue testing of orthopaedic implants a sample size of n = 5 or n = 6.

When it comes to functional testing, no standard exists. It is difficult to determine a specific number of specimens to be used in pre-clinical testing as this depends on the test reproducibility, on the possible inter-specimen variability, and on the least significant difference to be detected. For these reasons, a sample size analysis should be always performed based on previous similar studies or on pilot tests (57). An actual ASTM F 04.22 draft describes a standard practice for the determination of sample sizes and comparative statistical analysis of test data for test methods on medical devices in arthroplasty which generate measurable or quantitative values (58). This practice explicitly applies to implant test methods in arthroplasty, which generate quantitative values in quasi-static testing. It does not apply to fatigue testing or wear testing (see Consensus Statement 7 "Biotribology" for suitable specimen numbers in wear testing). For performance verification the ASTM draft at the current state requires testing no fewer than 5 specimens of each test implant device. A minimum of five specimens is foreseen to compute the sample size necessary to perform a comparative test. These five data points may come from prototype testing, pilot testing, performance data for a similar reference in the literature, or data from a previous test. An actual comparison of a new device to a similar reference device with clinical history shall be performed to determine superiority, equivalency or non-inferiority. As a thumb rule, the minimum number of specimens to be considered in several types of in vitro tests is between 6 and 10 (e.g. (26,32,59,60): this typically allows detecting differences between designs, if test reproducibility is granted. It must be noted that this small sample size allows exploring what happens in a typical scenario, e.g. average patient or the worst-case. Conversely, much larger samples are needed to account for the wide inter-patient variability.

Table 1: Incidence of the different failure modes of THR according to the different implant registries.

	Range	RIPO Emilia Romagna	RIAP Italy*	Germany	Netherlands *	Norway **	Sweden		Uk	New Zealand	Australia
	Reference population	Emilia Romagna	(Lombardy, Tuscany, Marche, Apulia, Basilicata, Calabria, Sicily, and Campania, Bolzano, Trento and two hospital structures "Policlinico Città di Alessandria" and "Santa Maria della Misericordia" of Udine	Germany	Netherlands	Norway	Sweden		Uk	New Zealand	Australia
Diagnosis in revision	**%**	*(% of number of revisions)*	*(% of number of revisions)*	*(% of number of revisions)*	*(% of number of revisions)*	*(% of number of revisions)*	*dual mobility cup cemented (% of number of revisions)*	*dual mobility cup uncemented (% of number of revisions)*	*(% of number of revisions)*	*(% of number of revisions)*	*(% of number of revisions)*
Loosening of acetabulum component	15.6-30.9	30.9	18.1	15.6	22.8	30.8				21.7	
Loosening of femur component	8.1-31.1	13.3	10.8	10.9	18.9	31.1	8.1	29.2		17.2	
Loosening total aseptic	3.3-18.3	18.3	7.0	3.3							24.0
Loosening septic	1.1-60.4	1.1	8.8	15.2	19.1	6.5	60.4	41.7	12.3	12.2	
Loosening aseptic	26.2								26.2		
Prosthesis breakage*	1.8-5.2	5.2	3.5	1.8					2.5		
modular necks (% of Prosthesis breakage)		*34.4*									
liners (% of Prosthesis breakage)		*22.2*									
heads (% of Prosthesis breakage)		*14.5*									
stems (% of Prosthesis breakage)		*12.3*									
cups (% of Prosthesis breakage)		*12.5*									
liner and head (% of Prosthesis breakage)		*1.7*									
not specified (% of Prosthesis breakage)		*2.3*									
Primary instability	0.7	0.7									21.6
Prosthesis dislocation	9.2-19.0	9.2	13.9	11.7	19.0	8.3	16.2	16.7	12.0	18.6	
Inlay wear	4.2-18.0	4.2			18.0						
Wear	3.1-9.2		8.2	8.1		3.1			9.2		
Metallosis	0.8	0.8									
Femur fracture	3.7-13.2		12.3	10.9	13.2	3.7	10.7	12.5	10.4	12.5	21.2
Acetabulum fracture	0.2	0.2									
Bone fracture	7.1	7.1									
Pain	1.7-15.2	1.7	5.6			6.4			3.6	15.2	
Trauma	0.2	0.2									
Osteolysis acetabular (no loosening)	1.6					1.6					
Osteolysis femur	1.9					1.9					
Lysis	0.9-8.9		2.7	0.9					8.9		2.1
Condition after removal	9.2			9.2							
Girdlestone situation	2.1-5.5				5.5	2.1					
Two steps prosthesis removal	5.0	5.0									
Symptomatic mom bearing	3.7				3.7						
Previous prosthesis removal	1.6		1.6								
Malalignment	1.6-3.2			1.6					3.2		
Adverse reaction to particulate debris	7.8								7.8		
Peri-articular ossification	0.5-1.8	0.5			1.8						
Progression of arthrosis	0.3			0.3							
Disease progression	0.1		0.1								
Head-socket size mismatch	0.3								0.3		
Missing information	0.4					0.4					
Other	0.0-11.7	1.5	7.4	10.5	11.3	4.2	4.6	0.0	3.8		11.7

*= one patient may have more than one reason for revisionor re-surgery. As such the total proportion is over 100%
**= Diseases are not mutually exclusive

Table 2: Incidence of the different failure modes of TKR according to the different implant registries.

	Range	RIPO Emilia Romagna	RIAP Italy*	Germany	Netherlands *	Norway **			Sweden ***			Uk	New Zealand	Australia
	Reference population	Emilia Romagna	(Lombardy, Tuscany, Marche, Apulia, Basilicata, Calabria, Sicily, and Campania, Bolzano, Trento and two hospital structures "Policlinico Città di Alessandria" and "Santa Maria della Misericordia" of Udine	Germany	Netherlands	Norway			Sweden			Uk	New Zealand	Australia
Diagnosis in revision	%	-	-	-	-	total knee prosthesis with patella	total knee prosthesis without patella	Unicondylar knee prosthesis	TKA-O	TKA-RA	UKA-OA	-	-	-
		(% of number of revisions)	(% of number of revisions)	(% of number of revisions)	(% of number of revisions)	(% of number of operations)			(% of number of revisions)			(% of number of revisions)	(% of number of revisions)	(% of number of revisions)
Aseptic loosening	17.0-38.4	38.4							17.0	29.0	28.0	29.6		
Loosening of femoral component	2.6-9.0	2.6	5.5	4.4	9.0	0.9	0.4	2.3					8.0	
Loosening of patellar component	0.2-20		0.2	0.6	1.7	0.5	-	-					20.0	32.2
Loosening of several components	10.7-26.0		26.0	10.7										
Loosening of tibial component	9.3-24.0	9.6	9.9	9.3	21.0	2.1	0.9	2.4					24.0	
Septic loosening	2.6-24.0	2.6	15.5	14.7	18.0	1.5	0.8	0.3	31.0	27.0	3.0	24.0	23.0	
Patellar reasons	2.0-20.9								19.0	4.0	2.0			20.9
Arthrofibrosis	4.2-4.8			4.2	4.8									4.7
Progression of arthities	8.6-16.5				8.6							16.5		
Breakage of prosthesis	0.5-2.0	0.5	1.4	2.0										
Condition after removal	10.9			10.9										
Dislocation prosthesis	2.2-3.2	2.2	2.4									3.2		
Dislocation of patella	2.4				2.4	0.2	0.1	-						
Dislocation no patella	-					0.1	0.1	-						
Fractured spacer	0.1		0.1											
Periprosthetic fracture	1.2-4.5	1.7	1.2	3.0	1.9	0.3	0.2	0.5				4.5		
Implant wear	2.0-10.8	3.7	2.4	5.7	7.3				2.0	4.0	10.0	10.8		
Instability	1.8-26.3	1.8	3.5		26.3	1.3	0.6	0.8	14.0	9.0	10.0	14.7		11.2
Ligament instability	8.9			8.9										
Malalignment / rotation revision	1.8-12.6			1.8	12.6	0.7	0.3	0.9				5.5		
Osteolysis with fixed component	1.1			1.1										
Femoral components				*0.3*										
Tibial tray				*0.3*										
Patellar components				*0.1*										
Several components				*0.4*										
Pain	9.9-25.0	9.9	16.7		21.0	1.7	1.8	5.7				10.1	25.0	10.7
Trauma	0.5	0.5												
Progression of desease	0.9-32	0.9	2.3						13.0	20.0	32.0			
Restricted mobility	4.0			4.0										
Stiffness	1.0-4.9	1.0	1.6									4.9		
Defect polyethylene	-					1.2	0.2	0.7						
Two steps prosthesis removal	18.4	18.4												
Revision of knee removal	5.6				5.6									
Missing	-					0.1	0.1	0.2						
Other	3.0-20.4	6.3	11.4	18.7	8.0	0.7	0.4	1.1	3.0	4.0	14.0	8.8		20.4

*= one patient may have more than one reason for revisionor re-surgery. As such the total proportion is over 100%
**= Revision causes are not mutually exclusive
***= numbers taken from a histagram chart

TKA O = Total Knee Arthroplasty in Osteoarthrities
TKA RA = Total Knee Arthroplasty in Rheumatoid Arthritis
UKA OA = Unicompartmental Knee Arthroplasty

Table 3: Modes of failure and related test methods for hip prostheses.

Type	**Component(s) involved**	**Description**	**Available test method?**	**Are test methods adequate?**	**Future directions**	**Dealt with in other sections?**	**Reference**
(device related, functional, surgeon-dependent, etc)	*Stem/ neck/ head/ resurfacing/ insert/ cup+ primary/ revision*		*(standardized, non-standardized, none)*	*(based on incidence of mode of failure)*	*(if previous column different from "yes")*	*(if yes, indicate # of relevant section)*	*If available, fill here (standard #, DOI, PMID or full ref)*
Device	Stem	Stem breakage	Standardized	Yes		**5**	
Device	Neck	Neck breakage	Standardized	Yes		**5**	
Device	Head	Head burst	Standardized	Yes		**5**	
Device	Head + insert	Wear of bearing surfaces	Standardized	Partly (not, on some cases such as MoM)		**6**	
Functional	Stem + cup	Long-term loosening (as a multifactorial phenomenon, including effect of wear debris and bone resorption)	Only for specific aspects,	No	Guidelines for in addressing multifactorial problem?		
Functional + surgeon	Stem (primary and revision)	Short-term lack of stability of uncemented stem	Non-standardized (both experimental and numerical)	Yes, when performed	Guidelines for in vitro testing? In silico simulation of worst cases?	**9**	
Functional + surgeon	Cup (primary and revision)	Short-term lack of stability of uncemented cup	Non-standardized (both experimental and numerical)	Yes, when performed	Guidelines for in vitro testing? In silico simulation of worst cases?	**9**	
Functional + surgeon	Stem and cement (primary and revision)	Long-term loosening of cemented stem	Non-standardized (both experimental and numerical)	Yes, when performed	Guidelines for in vitro testing? In silico simulation of worst cases?	**9**	
Functional + surgeon	Cup (primary and revision)	Short-term lack of stability of uncemented cup	Non-standardized (both experimental and numerical)	Yes, when performed	Guidelines for in vitro testing? In silico simulation of worst cases?	**9**	
Functional + surgeon	Stem	Periprosthetic intra-op fractures (uncemented)	Non-standardized (both experimental and numerical)	?	?	**12, 15**	
Functional + surgeon	Resurfacing	Periprosthetic post-op fractures (notching)	Non-standardized (both experimental and numerical)	?	?		
Functional + surgeon	Cup	Periprosthetic intra-op fractures (uncemented)	Non-standardized (both experimental and numerical)	?	?		
Functional + surgeon	Cup / stem	Implant mis-sizing	none			**7**	
Functional + surgeon + instrumentation	Cup / stem /resurfacing	Implant mis-positioning / misaligning	Some non-standard	partially	Assess numerically tolerance to surgical error?	**4A/B**	
Functional + surgeon + indications		Luxation	Some non-standard	partially	Build recommendations numerically?	**10**	

Table 4: Modes of failure and related test methods for knee prostheses.

Type	Component(s) involved	Description	Available test method?	Are test methods adequate	Future directions	Dealt with in other sections?	Reference
(device related, functional, surgeon-dependent, etc)	*Tibial tray / meniscus/ femoral component /patellar*		*(standardized, non-standardized, none)*	*(based on incidence of mode of failure)*	*(if previous column different from "yes")*	*(if yes, indicate # of relevant section)*	*If available, fill here (standard #, DOI, PMID or full ref)*
Device	Tibial tray	Breakage	Standardized	Partly (the load levels defined are not suitable for all design families)		**5**	
Device	Moving elements	Wear of meniscal bearing	Standardized	Partly (not, on some cases such as MoM)		**6**	
Functional	Tibial / femoral	Long-term loosening (as a multifactorial phenomenon, including effect of wear debris and bone resorption)	Only for specific aspects,	No	Guidelines for in addressing multifactorial problem?		
Functional + surgeon	Femoral and cement (primary and revision)	Long-term loosening of cemented femoral component	Non-standardized (both experimental and numerical)	Yes, when performed	Guidelines for in vitro testing? In silico simulation of worst cases?	**9**	
Functional + surgeon	Tibial and cement (primary and revision)	Long-term loosening of cemented tibial component	Non-standardized (both experimental and numerical)	Yes, when performed	Guidelines for in vitro testing? In silico simulation of worst cases?	**9**	
Functional + surgeon	Tibial / femoral /meniscus	Implant mis-sizing	none			**7**	
Functional + surgeon + instrumentation	Tibial / femoral	Implant mis-positioning / misaligning	Some non-standard	partially	Assess numerically tolerance to surgical error?	**4A/B**	
Functional + surgeon + indications		Knee pain	Non-standard, very preliminary	no	Build recommendations numerically?	**10**	

7. References

1. Harman MK, Toni A, Cristofolini L, Viceconti M. **Initial stability of uncemented hip stems: an in-vitro protocol to measure torsional interface motion. Med Eng Phys. 1995 Apr;17(3):163–71.**

2. Monti L, Cristofolini L, Viceconti M. **Methods for quantitative analysis of the primary stability in uncemented hip prostheses. Artif Organs. 1999 Sep;23(9):851–9.**

3. McKellop H, Ebramzadeh E, Niederer PG, Sarmiento A. **Comparison of the stability of press-fit hip prosthesis femoral stems using a synthetic model femur. J Orthop Res. 1991;9:297–305.**

4. Cristofolini L, Duda G, Prendergast PJ. **Inter-laboratory consensus about stability testing of cemented and cementless hip stems. In: European Orthopaedic Research Society Transactions Vol 16, Baldini N, Savarino L eds, CLUEB Publ, Bologna: O66. 2006.**

5. Cristofolini L, Saponara Teutonico A, Monti L, Cappello A, Toni A. **Comparative in vitro study on the long term performance of cemented hip stems: validation of a protocol to discriminate between "good" and "bad" designs. J Biomech. 2003 Nov;36(11):1603–15.**

6. Cristofolini L, Saponara Teutonico A, Savigni P, Erani P, Viceconti M. **Preclinical assessment of the long-term endurance of cemented hip stems. Part 1: effect of daily activities--a comparison of two load histories. Proc Inst Mech Eng H. 2007 Aug;221(6):569–84.**

7. Cristofolini L, Erani P, Savigni P, Bordini B, Viceconti M. **Preclinical assessment of the long-term endurance of cemented hip stems. Part 2: in-vitro and ex-vivo fatigue damage of the cement mantle. Proc Inst Mech Eng H. 2007 Aug;221(6):585–99.**

8. Cristofolini L, Erani P, Grupp TM, Jansson V, Viceconti M. **In vitro long-term fatigue endurance of the secondary "Cement Injection Stem" hip prosthesis. Artif Organs. 2007 Jun;31(6):441–51.**

9. Morellato K, Grupp TM, Bader U, Sungu M, Fink B, Cristofolini L. **Standard and line-to-line cementation of a polished short hip stem: Long-term in vitro implant stability. J Orthop Res. 2018;36(10):2736–44.**

10. McCormack BAO, Prendergast PJ, Gallagher DG. **An experimental study of damage accumulation in cemented hip prostheses. Clin Biomech. 1996;11(4):214–9.**

11. Lennon AB, Prendergast P. **Evaluation of cement stresses in Finite Element Analyses of cemented orthopaedic implants. J Biomech Eng. 2001;123(6):623–8.**

12. Stolk J, Maher S, Verdonschot N, Prendergast PJ, Huiskes R. **Can finite element models detect clinically inferior cemented hip implants? Clin Orthop Relat Res. 2003;409:138–50.**

13. Verdonschot N. **Philosophies of stem designs in cemented total hip replacement. Orthopedics. 2005;28(8 Suppl):833–40.**

14. Stolk J, Verdonschot N, Murphy BP, Prendergast PJ, Huiskes R. **Finite element simulation of anisotropic damage accumulation and creep in acrylic bone cement. Engng Fract Mater. 2004;71:513–28.**

15. Babuska I, Oden JT. **Verification and validation in computational engineering and science: basic concepts. Comput Methods Appl Mech Eng. 2004;193(36–38):4057–66.**

16. Henninger HB, Reese SP, Anderson AE, Weiss JA. **Validation of computational models in biomechanics. Proc Inst Mech Eng H. 2010;224(7):801–12.**

17. Cristofolini L, Schileo E, Juszczyk M, Taddei F, Martelli S, Viceconti M. **Mechanical testing of bones: the positive synergy of FE models and in vitro experiments. Philos Trans Math Phys Eng Sci. 2010;368:2725–63.**

18. Szivek JA, Thompson JD, Benjamin JB. **Chracterization of three formulations of a synthetic foam as models for a range of human cancellous bone types. J Appl Biomater. 1995;6:125–8.**

19. Schierjott RA, Hettich G, Baxmann M, Morosato F, Cristofolini L, Grupp TM. **Primary stability of a press-fit cup in combination with impaction grafting in an acetabular defect model. J Orthop Res Off Publ Orthop Res Soc. 2021;39(5):929–40.**

20. C. Markel, Nivedita Hora, Michele G D. **Press-fit stability of uncemented hemispheric acetabular components: a comparison of three porous coating systems. Int Orthop. 2002 Apr 1;26(2):72–5.**

21. Crosnier EA, Keogh PS, Miles AW. **The effect of dynamic hip motion on the micromotion of press-fit acetabular cups in six degrees of freedom. Med Eng Phys. 2016 Aug;38(8):717–24.**

22. Beckmann NA, Bitsch RG, Gondan M, Schonhoff M, Jaeger S. **Comparison of the stability of three fixation techniques between porous metal acetabular components and augments. Bone Jt Res. 2018;7(4):282–8.**

23. Dickinson AS, Taylor AC, Browne M. **The influence of acetabular cup material on pelvis cortex surface strains, measured using digital image correlation. J Biomech. 2012;45(4):719–23.**

24. Morosato F, Traina F, Cristofolini L. **Effect of different motor tasks on hip cup primary stability and on the strains in the periacetabular bone: An in vitro study. Clin Biomech. 2019;70:137–45.**

25. Morosato F, Traina F, Cristofolini L. **A reliable in vitro approach to assess the stability of acetabular implants using digital image correlation. Strain. 2019;55(4 (e12318)):1–12.**

26. Morosato F, Cristofolini L, Castagnini F, Traina F. **Effect of cup medialization on primary stability of press-fit acetabular cups. Clin Biomech. 2020 Dec 1;80(paper 105172):1–6.**

27. Morosato F, Traina F, Schierjott RA, Hettich G, Grupp TM, Cristofolini L. **Primary Stability of Revision Acetabular Reconstructions Using an Innovative Bone Graft Substitute: A Comparative Biomechanical Study on Cadaveric Pelvises. Materials. 2020;13(19):1–11.**

28. Clarke SG, Phillips AT, Bull AM. **Validation of FE micromotions and strains around a press-fit cup: introducing a new micromotion measuring technique. Ann Biomed Eng. 2012;40(7):1586–96.**

29. Huiskes R. **The various stress patterns of press-fit, ingrown, and cemented femoral stems. Clin Orthop Relat Res. 1990;261:27–38.**

30. Van Rietbergen B, Huiskes R, Weinans H, Sumner DR, Turner TM, Galante JO. **The mechanism of bone remodeling and resorption around press-fitted THA stems. J Biomech. 1993;26:369–82.**

31. Bobyn JD, Mortimer ES, Glassman AH, Engh CA, Miller JE, Brooks CE. **Producing and avoiding stress shielding. Laboratory and clinical observations of noncemented total hip arthroplasty. Clin Orthop Relat Res. 1992;274:79–96.**

32. Cristofolini L. **A critical analysis of stress shielding evaluation of hip prostheses. Crit Rev Biomed Eng. 1997;25(4–5):409–83.**

33. Cristofolini L, Juszczyk M, Taddei F, Field R, Rushton N, Viceconti M. **Stress-shielding and stress-concentration of contemporary epiphyseal hip prostheses. Proc Inst Mech Eng H. 2009;223(1):27–44.**

34. Waide V, Cristofolini L, Stolk J, Verdonschot N, Toni A. **Experimental investigation of bone remodelling using composite femurs. Clin Biomech. 2003 Jul;18(6):523–36.**

35. Davis ET, Olsen M, Zdero R, Waddell JP, Schemitsch EH. **Predictors of femoral neck fracture following hip resurfacing: a cadaveric study. In 2007. p. 241–3.**

36. Cristofolini L, Affatato S, Erani P, Leardini W, Tigani D, Viceconti M. **Long-term implant-bone fixation of the femoral component in total knee replacement. Proc Inst Mech Eng H. 2008 Apr;222(3):319–31.**

37. Cristofolini L, Affatato S, Erani P, Tigani D, Viceconti M. **Implant fixation in knee replacement: preliminary in vitro comparison of ceramic and metal cemented femoral components. Knee. 2009;16:101–8.**

38. Skwara A, Figiel J, Knott T, Paletta JRJ, Fuchs-Winkelmann S, Tibesku CO. **Primary stability of tibial components in TKA: in vitro comparison of two cementing techniques. Knee Surg Sports Traumatol Arthrosc. 2009 Oct;17(10):1199–205.**

39. Figgie HE, Davy DT, Heiple KG, Hart RT. **Load-bearing capacity of the tibial component of the total condylar knee prosthesis. An in vitro study. :1.**

40. Ostermeier S, Hurschler C, Windhagen H, Stukenborg-Colsman C. **In vitro investigation of the influence of tibial slope on quadriceps extension force after total knee arthroplasty. Knee Surg Sports Traumatol Arthrosc. 2006 Sep 29;14(10):934–9.**

41. Finlay JB, Bourne RB, McLean J. **A technique for the in vitro measurement of principal strains in the human tibia. J Biomech. 1982 Jan;15(10):723–39.**

42. Gray HA, Zavatsky AB, Taddei F, Cristofolini L, Gill HS. **Experimental validation of a finite element model of a composite tibia. Proc Inst Mech Eng H. 2007 Apr;221(3):315–24.**

43. Gray HA, Taddei F, Zavatsky AB, Cristofolini L, Gill HS. **Experimental validation of a finite element model of a human cadaveric tibia. J Biomech Eng ASME. 2008 Jun;130(3-article 031016):1–9.**

44. Komistek RD, Kane TR, Mahfouz M, Ochoa JA, Dennis DA. **Knee mechanics: a review of past and present techniques to determine in vivo loads. J Biomech. 2005 Feb;38(2):215–28.**

45. Perillo-Marcone A, Barrett DS, Taylor M. **The importance of tibial alignment. J Arthroplasty. 2000 Dec;15(8):1020–7.**

46. Sawatari T, Tsumura H, Iesaka K, Furushiro Y, Torisu T. **Three-dimensional finite element analysis of unicompartmental knee arthroplasty–the influence of tibial component inclination. J Orthop Res. 2005 May;23(3):549–54.**

47. Jeffers JRT, Browne M, Roques A, Taylor M. **On the Importance of Considering Porosity When Simulating the Fatigue of Bone Cement. J Biomech Eng. 2005 Aug 1;127(4):563–70.**

48. Martelli S, Taddei F, Moindreau M, Cristofolini L, Viceconti M. **Pre-clinical validation of a new proximal epiphyseal replacement:design revision and optimisation by means of finite element models. J Biomech. 2008;41(S1):s34.**

49. Martelli S, Taddei F, Cristofolini L, Schileo E, Rushton N, Viceconti M. **A new hip epiphyseal prosthesis: design revision driven by a validated numerical procedure. Med Eng Phys. 2011 Dec;33(10):1203–11.**

50. Dar FH, Meakin JR, Aspden RM. **Statistical methods in finite element analysis. J Biomech. 2002;35:1155–61.**

51. Viceconti M, Affatato S, Baleani M, Bordini B, Cristofolini L, Taddei F. **Pre-clinical validation of joint prostheses: a systematic approach. J Mech Behav Biomed Mater. 2009;2(2):120–7.**

52. Martelli S, Taddei F, Cristofolini L, Viceconti M. **Preclinical validation of epiphyseal prostheses. In: Engineers and surgeons joined at the hip Conference Proceedings, IMechE publ: 161-165. 2007.**

53. Dickinson AS, Taylor AC, Ozturk H, Browne M. **Experimental validation of a finite element model of the proximal femur using digital image correlation and a composite bone model. J Biomech Eng. 2011;133(1):014504.**

54. Harvey AR, Barlow IW, Clarke NMP. **Early Femoral Loosening of the Titanium 3M Capital Cemented Hip: A Lesson Relearned. HIP Int. 2002 Jan;12(1):17–22.**

55. Reito A, Lehtovirta L, Lainiala O, Mäkelä K, Eskelinen A. **Lack of evidence—the anti-stepwise introduction of metal-on-metal hip replacements: A systematic review and a comparative assessment of the literature and registry data. Acta Orthop. 2017 Sep 3;88(5):478–83.**

56. Laaksonen I, Donahue GS, Madanat R, Makela KT, Malchau H. **Outcomes of the Recalled Articular Surface Replacement Metal-on-Metal Hip Implant System: A Systematic Review. J Arthroplasty. 2017 Jan;32(1):341–6.**

57. Montgomery DC. **Design and analysis of experiments. 6th edition. New York: J. Wiley; 2005. 660 p.**

58. ASTM F 04.22 draft WK 69860 Standard Practice for the Determination of Sample Sizes and the Corresponding Comparative Statistical Analysis of Test Data for Arthroplasty Medical Device Test Methods Which Generate Measurable or Quantitative Values Marc 19, 2021.

59. Maher SA, Prendergast PJ. **Discriminating the loosening behaviour of cemented hip prosthesis using measurements of migration and inducible displacement. J Biomech. 2002;35:257–65.**

60. Engh CA, O'Connor D, Jasty M, McGovern TF, Bobyn JD, Harris WH. **Quantification of implant micromotion, strain shielding, and bone resorption with porous-coated anatomic medullary locking femoral prostheses. Clin Orthop Relat Res. 1992;285:13–29.**

RESEARCH TOPIC

INTERFACE COMPATIBILITY/ INTERFACE GEOMETRY

Research Topic 3

Interface Compatibility of Joint Arthroplasty Devices – Geometrical, Dimensional and Functional Assessment of Implant/Instrument Interfaces

RESEARCH TOPIC 3

Interface Compatibility of Joint Arthroplasty Devices – Geometrical, Dimensional and Functional Assessment of Implant/Instrument Interfaces

Authors

Claude Rieker[1], Massimiliano Bernardoni[2], Christoph Schilling[3], Matthias Woiczinski[4], Joshua Bridgens[5]

[1]Zimmer Biomet EMEA, Scientific Affairs, Zug, Switzerland
[2]Medacta International SA, Castel San Pietro, Switzerland
[3]Aesculap AG Research & Development, Tuttlingen, Germany
[4]Ludwig Maximilians University Munich, Department of Orthopaedic and Trauma Surgery, Musculoskeletal University Center Munich (MUM), Campus Grosshadern, Munich, Germany
[5]DePuy Synthes, Leeds, United Kingdom

1. Abstract

Question: Interface compatibility / interface geometry

How can be confirmed that all interfaces (implant – instrument only) are geometrically / dimensionally compatible and fulfil the intended purpose, i.e. the interface is functional in clinical practice?

1. Which interfaces must be considered for hip and knee implants regarding compatibility?
2. Which relevance have function and actual geometry for the safety and compatibility of interfaces?
3. How should the implant – instrument interface be assessed regarding compatibility, safety and functional performance?

Summary / Recommendation:

The key aspect to guarantee that all implant – instrument interfaces are geometrically / dimensionally compatible and fulfil the intended clinical purpose is to optimize their resistance to wear / deformation as well as their corrosion resistance.

International standards (ISO and ASTM) give useful guidelines about the most suitable metallic alloys / stainless steels for orthopaedics instruments and these guidelines have to be followed.

Furthermore, as the life expectancy of these instruments is limited, the manufacturer must give information when the device must no longer be reused, e.g. signs of material degradation or the maximum number of allowable reuses. This requirement is now explicitly mentioned in the annex I (point 23.4. [n]) of the European Medical Device Regulation (MDR) 2017/745.

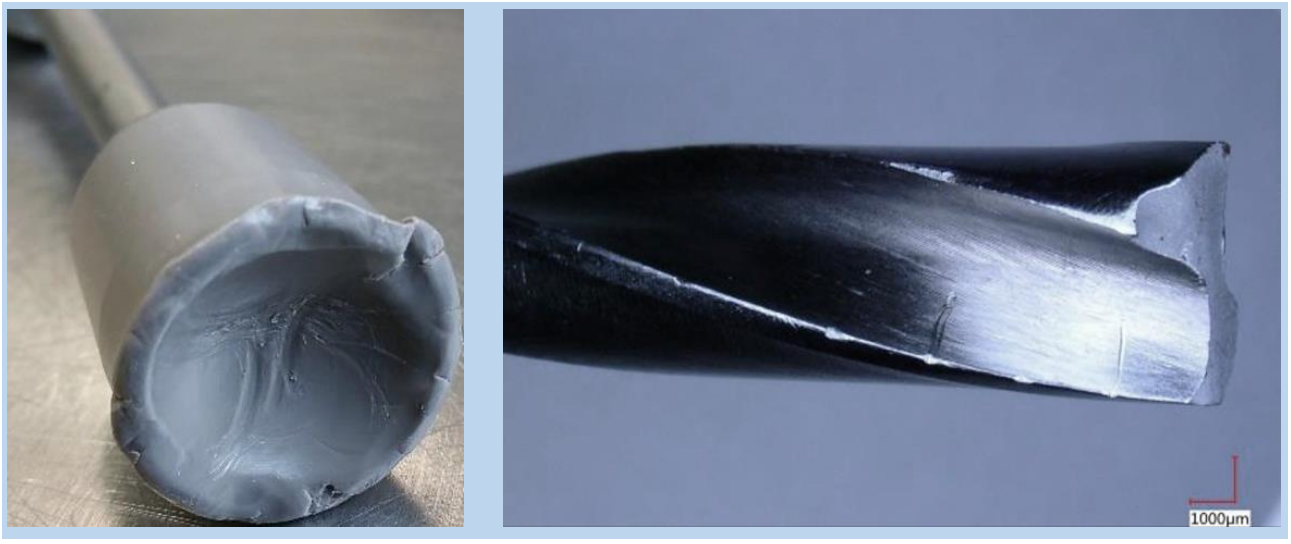
2. Level of Evidence

Medium to strong

3. Consensus Delegate Vote

98% - unanimous, strongest consensus (98% agree / 0% disagree / 2% abstain)

4. Graphical Abstract



Unacceptable worn instruments leading to improper implant – instrument interfaces [1].

5. Search Strategy

A non-systematic literature search was conducted end of March 2021 to gain an overview of the available published literature about orthopaedics instruments.

The two electronic databases MEDLINE using PubMed and "Web of Science" were explored using a combination of the following key words:

- THA, TKA, hip prosthesis, knee prosthesis
- Instruments
- (Reamer AND awl) OR (broach AND rasp) OR (impactor AND inserter) OR (cutting guide) OR (spanner AND wrench)

Article language was restricted to English or German or French languages. Case reports were also included.

The results of this literature search were disappointing as the number of publications exploring the implant – instrument interface is practically non-existent. Only one publication treating partially this interface could be identified [1]. All the other identified publications mainly explore the performance and / or the failure of orthopaedics instruments and the implant – instrument interface seems not to be considered as being an area of large interest.

As the new European Medical Device Regulation (MDR) 2017/745 is now fully implemented, the impact of this new regulation on reusable orthopaedics instruments was also considered.

6. Rationale

1. Which interfaces must be considered for hip and knee implants regarding compatibility?

Three different types of interface have been defined:

1. Indirect interfaces

Even if this type of interfaces does not have a direct contact between the implant and the instrument, these indirect interfaces do play an extremely important role for the clinical performance of hip and knee implants.

Typical instruments for these indirect interfaces are either instrument allowing to prepare the bony cavity such as broaches, rasps, reamers and awls or test implants.

Without a precise preparation of the bony cavity, the positioning of the implant may be problematic (e. g. implants in varus / valgus) and their primary fixation of (mainly uncemented) implants cannot be secured.

The position of the test implants must correspond to the final position of the definitive implants to assure for example the correct leg length in the case of a total hip arthroplasty.

2. Direct interfaces

This interface has a direct contact between the implant and the instrument.

Typical instruments for this type of interface are impactors, screwdrivers, repositioning levers and repositioning tops.

3. Direct interface with additional / complex functionality

This interface has a direct contact between the implant and the instrument and the instrument has a specific / complex functionality.

Typical instruments for this type of interface are measuring devices, impactors with a given impulse and torque wrenches.

2. Which relevance have function and actual geometry for the safety and compatibility of interfaces?

To assure a good functionality, reusable instruments must fulfil some key requirements independently from the type of interfaces defined in the first question.

Reusable instruments are defined in the annex VIII (point 2.3.) of the Medical Device Regulation (MDR) 2017/745:

- "Reusable surgical instrument means an instrument intended for surgical use in cutting, drilling, sawing, scratching, scraping, clamping, retracting, clipping or similar procedures, without a connection to an active device and which is intended by the manufacturer to be reused after appropriate procedures such as cleaning, disinfection and sterilisation have been carried out."

These reusable instruments have to keep their local and global geometry as intact as to assure the safety and compatibility of the different interfaces. The two key aspects for maintaining their geometry / functionality are their resistance to wear / deformation as well as their corrosion resistance.

Fortunately, for metallic instruments, international standards give clear guidelines for the choice for the metallic alloys / stainless steels which should be used for orthopaedics instruments:

1. ISO 7153-1 [2]

Surgical Instruments – Materials

2. ASTM F899 - 20 [3]

Standard Specification for Wrought Stainless Steels for Surgical Instruments

For each type of instruments (e.g. reamers and rasps), these standards defined the recommended alloys / stainless steels (including their specific heat treatment) giving the best possible performance (hardness [sharpness and wear resistance], corrosion resistance). The corrosion resistance of these metallic instruments can also be optimised by their passivation.

Two additional ASTM standards / guides give also useful indication for assuring the best possible performance of these metallic instruments in the long-term:

3. ASTM F565 – 21 [4]

Standard Practice for Care and Handling of Orthopedic Implants and Instruments

4. ASTM F1744 – 96 [5]

Standard Guide for Care and Handling of Stainless Steel Surgical Instruments

These standards / guides give clear directives for the cleaning and the storage of these reusable instruments, helping to keep their service life as long as possible.

As long as the local and global geometry of these metallic instruments is preserved, the safety and compatibility of the implant – instrument interfaces is maintained. But even if their geometry is preserved, these metallic instruments should be replaced earlier in case that the wear has lowered their performance (cutting performance as a typical example) below a critical threshold, impairing with their functionality. The threshold value should be defined by the manufacturer defining a fixed maximum number of cycles, by the judgement of the operating surgeon or by a functional control done for example by the sterilisation technicians / engineers.

For polymeric instruments (e.g. impactor tops, repositioning tops), as their resistance to plastic deformation and their resistance to wear is definitively much lower compared to metallic instruments, they should be considered as being disposable instruments and must regularly be replaced.

3. How should the implant – instrument interface be assessed regarding compatibility, safety and functional performance?

For the instruments having an interface with a functionality, the requirements to keep their local and global geometry to assure the safety and compatibility of the different interfaces is also mandatory.

The functionality of these instruments must be maintained during their whole service life. Taking torque wrenches as an example, three different strategies are possible:

1. Passive functionality

The required torque can be generated by the elastic deformation of a metallic rod.

2. "Sacrificial" functionality

The required torque is controlled by the "sacrifice / destruction" of a disposable element inside the wrench. This disposable element must be replaced each time before the wrench is use.

3. Active functionality

The required torque is obtained by a mechanical system inside the wrench. The manufacturer must define a maximum number of cycles before the instrument must be controlled and possibly be recalibrated.

4. Take home message

In conclusion, for the three types of interfaces defined in this document, all these general guidelines should be followed. The manufacturer must also give the information when the instruments must be replaced. This specific requirement is also mentioned in the annex I (point 23.4. [n]) of the European Medical Device Regulation (MDR) 2017/745:

- "Information shall be provided to identify when the device should no longer be reused, e.g. signs of material degradation or the maximum number of allowable reuses."

7. References

1. Ellenrieder M, Enz A, Kluess D, Mittelmeier W. **Ist der unbegrenzte Gebrauch von chirurgischen Instrumenten unbedenklich? Der Orthopäde 2020; 49(12):1049–1055.**

2. ISO 7153-1:2016. Surgical instruments – Materials. International Organization for Standardization. Geneva, Switzerland

3. ASTM F899 – 20. Standard Specification for Wrought Stainless Steels for Surgical Instruments. ASTM Headquarters. West Conshohocken, PA, USA

4. ASTM F565 – 21. Standard Practice for Care and Handling of Orthopedic Implants and Instruments. ASTM Headquarters. West Conshohocken, PA, USA

5. ASTM F1744 – 96. Standard Guide for Care and Handling of Stainless Steel Surgical Instruments. ASTM Headquarters. West Conshohocken, PA, USA

RESEARCH TOPIC

MECHANICAL COMPONENT TESTING

RESEARCH TOPIC 4

Methods in Total Hip Arthroplasty Mechanical Component Testing – Discussing Standard and Additional Test Protocols for Contemporary Product Development

Authors

Massimiliano Bernardoni[1], Luca Cristofolini[2], Jan Philippe Kretzer[3]

[1]Medacta International SA, Castel San Pietro, Switzerland
[2]Department of Industrial Engineering, School of Engineering and Architecture, Alma Mater Studiorum – Università di Bologna, Bologna, Italy
[3]Laboratory of Biomechanics & Implant Research, Department of Orthopaedic Surgery, Heidelberg University Hospital, Heidelberg, Germany

1. Abstract

Question: Mechanical component testing

Are there standard methods to establish that the implant will withstand the endurance habitual and peak loads that must reasonably be expected (i.e., single implant parts as well as complete arthroplasty combination)?

1 Which standard methods/tests according to norms are required for CE registration?
2 What validity do they have with regard to the clinical use?
3 Is it necessary to take into consideration the instrumentation in this process?

Are there additional test methods to establish that an implant will withstand the endurance habitual and peak loads that must reasonably be expected (i.e., single implant parts as well as complete arthroplasty combination)?

1 Which additional/new test methods can reasonably complement the tests according to norm?
2 What validity do they have with regard to the clinical use as opposed to the standard tests?
3 Which supplementary steps (methods) are imaginable to further close the gap between pre-clinical and clinical application (e.g., use of a numerical simulator or comparative testing with established products)?
4 What are methods that simulate the in vivo conditions for load testing of implants in order to gain information about biological impact on material properties?

Summary / Recommendation:

There are several International Standards (IS) that describe tests under simplified loading conditions to verify the ability of a THA device (single device or a combination of devices) to withstand repeated load application.

These tests can be static or dynamic and the number of cycles should be correlated to the expected lifespan of the device. It is generally assumed that 2 mio of cycles correspond to 1 year of time, but a specific patient could experience a much higher number of cycles in the same period of time and this aspect should be considered.

Nowadays, a minimum of 10 mio cycles seems to be adequate for dynamic testing, even if the definition of a strict correlation between the test protocol and the in vivo behavior of the component can be extremely difficult and useless.

Test protocols adapted from the IS have been designed to prove that a minimal mechanical performance of a device is achieved and that these mechanical performances are adequate to the load conditions that could be experienced in vivo.

To prove this consideration, a review of the national registry is relevant to understand the frequency of the implant failure related to mechanical breakage: according to all registries, component breakage is responsible for only 1.8-5.2% of all revisions.

The usage of specific instruments could lead to implant micro and macro damage that could cause early implant failure. For this reason, it could be important to consider the application of some specific instruments for mechanical testing of implant endurance / fatigue properties.

2. Level of Evidence

Moderate

3. Consensus Delegate Vote

93% - super majority, strong consensus (93% agree / 2% disagree / 5% abstain)

4. Graphical Abstract

		iso/astm	Norm	year	Description	Implant type
Dynamic Test Femoral Stem	1	ISO	7206-4	2010	Implants for surgery — Partial and total hip joint prostheses — Part 4: Determination of endurance properties and performance of stemmed femoral components	Stem
	2	ISO	7206-6	2013	Implants for surgery — Partial and total hip joint prostheses — Part 6: Endurance properties testing and performance requirements of neck region of stemmed femoral components	
	3	ASTM	F2580-18	2018	Evaluation of Modular Connection of Proximally Fixed Femoral Hip Prosthesis	
Modular Connection Head/Taper	4	ISO	7206-10	2018	Implants for surgery — Partial and total hip-joint prostheses — Part 10: Determination of resistance to static load of modular femoral heads	Head
	5	ISO	7206-13	2016	Determination of resistance to torque of head fixation of stemmed femoral components	
	6	ASTM	F2345-03	2013	Standard Test Methods for Determination of Static and Cyclic Fatigue Strength of Ceramic Modular Femoral Heads	
	7	ASTM	F2009-00	2011	Standard Test Method for Determining the Axial Disassembly Force of Taper Connections of Modular Prostheses	
Dynamic Test Acetabular Cup	8	ASTM	F3090-20	2020	Standard Test Method for Fatigue Testing of Acetabular Devices for Total Hip Replacement	Acetabular cup
Acetabular Cup Deformation and Modular connection	9	ISO	7206-12	2016	Implants for surgery — Partial and total hip joint prostheses — Part 12: Deformation test method for acetabular shells	Acetabular cup
	10	ASTM	F1820-13	2013	Standard Test Method for Determining the Forces for Disassembly of Modular Acetabular Devices	
ROM and Impingement	11	ASTM	F2582-20	2020	Standard Test Method for Dynamic Impingement Between Femoral and Acetabular Hip Components	Acetabular cup
Adesion	12	ASTM	F1147-05	2017	Standard Test Method for Tensile Testing of Calcium Phosphate Coatings and Metallic Coatings	stem/acetabular cup
	13	ISO	13779-4	2002	Implants for surgery – Hydroxyapatite – Determination of coating adhesion strength	
Head Burst	14	ISO	11491	2017	Determination of impact resistance of ceramic femoral heads for hip joint prostheses	Head

Figure 1

5. Search Strategy

No formal literature search was conducted on the basis that the question could be better answered on the basis of current registry data answering this topic. The impression was that a formal search will not give more evidence to this answer.

Therefore, a research of national registries was carried out, trying to extrapolate the failure rate directly linked to implant breakage.

	Range	RIPO Emilia Romagna		RIAP Italy*	Germany	Netherlands *	Norway **	Sweden		Uk	New Zealand	Australia
								cemented	uncemented			
Diagnosis in revision	**%**		*(% of number of revisions)*	*(% of number of revisions)*	*(% of number of revisions)*	*(% of number of revisions)*	*(% of number of revisions)*	*(% of number of revisions)*	*(% of number of revisions)*	*(% of number of revisions)*	*(% of number of revisions)*	*(% of number of revisions)*
Loosening of acetabulum componer	15.6-30.9		30,9	18,1	15,6	22,8	30,8				21,7	
Loosening of femur component	8.1-31.1		13,3	10,8	10,9	18,9	31,1	8,1	29,2		17,2	
Loosening total aseptic	3.3-18.3		18,3	7,0	3,3							24,0
Loosening septic	1.1-60.4		1,1	8,8	15,2	19,1	6,5	60,4	41,7	12,3	12,2	
Loosening aseptic	26,2									26,2		
Prosthesis breakage*	1.8-5.2	% of prosth. breakage	5,2	3,5	1,8					2,5		
Primary instability	0,7		0,7									
Prosthesis dislocation	9.2-19.0		9,2	13,9	11,7	19,0	8,3	16,2	16,7	12,0	18,6	21,6
Inlay wear	4.2-18.0		4,2			18,0						
Wear	3.1-9.2			8,2	8,1		3,1			9,2		
Metallosis	0,8		0,8									
Femur fracture	3.7-13.2			12,3	10,9	13,2	3,7	10,7	12,5	10,4	12,5	
Acetabulum fracture	0,2		0,2									21,2
Bone fracture	7,1		7,1									
Pain	1.7-15.2		1,7	5,6			6,4			3,6	15,2	
Trauma	0,2		0,2									
Osteolysis acetabular (no loosening	1,6						1,6					
Osteolysis femur	1,9						1,9					
Lysis	0.9-8.9			2,7	0,9					8,9		2,1
Condition after removal	9,2				9,2							
Girdlestone situation	2.1-5.5					5,5	2,1					
Two steps prosthesis removal	5,0		5,0									
Symptomatic mom bearing	3,7					3,7						
Previous prosthesis removal	1,6			1,6								
Malalignment	1.6-3.2				1,6					3,2		
Adverse reaction to particulate debri	7,8									7,8		
Peri-articular ossification	0.5-1.8		0,5			1,8						
Progression of arthrosis	0,3				0,3							
Disease progression	0,1			0,1								
Head-socket size mismatch	0,3									0,3		
Missing information	0,4						0,4					
Other	0.0-11.7		1,5	7,4	10,5	11,3	4,2	4,6	0,0	3,8		11,7

**= one patient may have more than one reason for revision or re-surgery. As such the total proportion is over 100%*

***= Diseases are not mutually exclusive*

According to the registry the mechanical failure of THA implants varies from 1.8 and 5.2%, but only the registry RIPO has detailed analysis of incidence of component failure and due to historical reason has been highly impacted by the presence on its territory by modular neck stem failure.

This is in contrast to published data on pooled worldwide arthroplasty register the incidence of implant fractures in total hip arthroplasty is very low at 304 fractures per 100,000 implants (1).

	Range	RIPO Emilia Romagna		RIAP Italy*	Germany	Netherlands *	Norway **	Sweden		Uk	New Zealand	Australia
								cemented	uncemented			
Diagnosis in revision	**%**		*(% of number of revisions)*	*(% of number of revisions)*	*(% of number of revisions)*	*(% of number of revisions)*	*(% of number of revisions)*	*(% of number of revisions)*	*(% of number of revisions)*	*(% of number of revisions)*	*(% of number of revisions)*	*(% of number of revisions)*
Prosthesis breakage*	1.8-5.2	% of prosth. breakage	5,2	3,5	1,8					2,5		
modular necks (% of Prosthesis breaka		*34,4*	*1,8*									
liners (% of Prosthesis breakage)		*22,2*	*1,2*									
heads (% of Prosthesis breakage)		*14,5*	*0,8*									
stems (% of Prosthesis breakage)		*12,3*	*0,6*									
cups (% of Prosthesis breakage)		*12,5*	*0,6*									
liner and head (% of Prosthesis breaka		*1,7*	*0,1*									
not specified (% of Prosthesis breakag		*2,3*	*0,1*									

6. Rationale

After the analysis of the principal (ISO and ASTM) existing International Standards (IS) that describe static and/or dynamic tests for THA, it is assumed that the actual ISO and ASTM IS, are adequate for this purpose and that they cover all aspects for static and dynamic testing under simplified loading conditions.

An exhaustive list of IS that describe tests under simplified loading conditions to verify the ability of a THA device (single device or a combination of devices) to withstand static or repeated load application has been listed in Figure 1.

A detailed analysis of the National registries has shown that only a minimal percentage (between 1.8% and 5.2%) failures is linked to implant breakage. This confirms that the actual standards seem to be adequate to test the mechanical properties of the implantable components and assemblies.

Unfortunately, most registries do not generally report a detailed implant failure categorization and for this reason it is impossible to extrapolate data that could be associated with a specific IS but globally it seems that breakage of implant does not represent a significant problem today, demonstrating the adequateness of the required tests.

In general, the applied load condition is a crucial aspect for mechanical tests. Specific load levels and acceptance criteria are specified by IS only in some Tests (i.e. ISO 7206-6).

For all static and Dynamic test, where the IS are not imposing a load value, we think that load condition has to be well justified and documented. If not possible of course a comparison test with other marked devices are fundamental to support adequate performance.

Even if for the hip, the analysis of Bergmann et al. provides a generally well accepted view as to whether the standards impose relevant and adequate loads (2), it is believed that most standards are only valid to a certain degree, regarding the clinical application of use.

However, historically, implant failures were used to adjust the standards in terms of static and dynamic loading conditions to evaluate the fatigue performance. For this reason, the current standards seem to address the clinical requirements sufficiently.

Another important point is represented by the number of cycles required for a fatigue test.

The number of applied cycles can be debatable and even if the majority of the IS required at least 10 mio cycles, considering 2 mio cycles represent a year of activity. Bergmann et al. have demonstrated that for an active patient 10 mio cycles represents only 3.9 years (2). Considering the fact that there is an increasing global tendency of operating younger patients, 10 mio cycles can be limited if compared to the expected lifetime of an implant related to young patients.

Probably these considerations should lead to an increase of the number of cycles above 10 mio for fatigue tests for some products that are designed specifically for young patients.

Two very important aspect that could influence the mechanical performance of an implant is linked to the tolerance of coupled components, and its surface finishing. For this reason, all surgical instruments that could influence this characteristic should be considered within the mechanical test protocol.

As an extreme example, which is well documented by R. Sonntag et al, the usage of an electrocautery in proximity of an implant can cause damages that can reduce implant fatigue strength (3).

Clearly, it is debated whether the electrocautery is to be considered an integral part of the surgical instrumentation, nevertheless with this example it can be demonstrated how an alteration of the implant surface can alter the mechanical characteristics.

Because the instrument could adversely compromise the mechanical performance of the implant, a risk analysis must be conducted to evaluate its potential influence to compromise the mechanical performance of the implant.

If there is a minimal probability that the usage of a surgical instrument could influence the performance of the device, the test protocol should consider this aspect.

It is believed that the existing standards are covering the mechanical evaluation of an implant well and for this reason no different methods are suggested. Sometime a non-standardized method is more adequate to evaluate a particular characteristic, but this should represent only an add-on to the existing IS.

In general, the environment prescribed by the standards (saline, solution, bovine serum etc.) creates a relevant condition, regarding the fretting or corrosion induced fatigue failures. More information regarding this aspect can be found in EFORT Consensus Statement 13 - Modularities / Interfaces.

7. References

1. Sadoghi P, Pawelka W, Liebensteiner MC, Williams A, Leithner A, Labek G. **The incidence of implant fractures after total hip arthroplasty. International Orthopaedics (SICOT) 2014; 38(1):39–46.**

2. Bergmann G, Bender A, Dymke J, Duda G, Damm P. **Standardized Loads Acting in Hip Implants. PLOS ONE 2016; 11(5):e0155612.**

3. Sonntag R, Gibmeier J, Pulvermacher S, Mueller U, Eckert J, Braun S et al. **Electrocautery Damage Can Reduce Implant Fatigue Strength: Cases and in Vitro Investigation. J Bone Joint Surg Am 2019; 101(10):868–78.**

RESEARCH TOPIC 5

Standard and Additional Methods in Total Knee Arthroplasty to Establish that the Implant will Withstand the Endurance Habitual and Peak Loads that Must Reasonably be Expected

Authors

Christian Kaddick[1], Christoph Schilling[2], Dennis Janssen[3], Jan Philippe Kretzer[4]

[1]EndoLab Mechanical Engineering GmbH, Riedering, Germany
[2]Aesculap AG Research & Development, Tuttlingen, Germany
[3]Radboud University Medical Center, Orthopaedic Research Lab, Research Institute for Medical Innovation, Nijmegen, the Netherlands
[4]Laboratory of Biomechanics & Implant Research, Department of Orthopaedic Surgery, Heidelberg University Hospital, Heidelberg, Germany

1. Abstract

Question: Mechanical component testing TKA

Part 1

Are there standard methods to establish that the implant will withstand the endurance habitual and peak loads that must reasonably be expected (i.e. single implant parts as well as complete arthroplasty combination)?

Q1: Which standard methods/tests according to norms are required for CE registration?
Q2: What validity do they have with regard to the clinical use?
Q3: Is it necessary to take into consideration the instrumentation in this process?

Part 2

Are there additional test methods to establish that an implant will withstand the endurance habitual and peak loads that must reasonably be expected (i.e. single implant parts as well as complete arthroplasty combination)?

Q1: Which additional/new test methods can reasonably complement the TKA tests according to norm?
Q2: What validity do they have with regard to the clinical use as opposed to the standard tests?

Part 3

Are there additional numerical test methods to establish that an implant will withstand the endurance habitual and peak loads that must reasonably be expected (i.e. single implant parts as well as complete arthroplasty combination)?

Q1: Which supplementary steps (methods) are imaginable to further close the gap between pre-clinical and clinical application (e.g. use of a numerical simulator or comparative testing with established products)?

Q2: What are methods that simulate the in vivo conditions for load testing of implants in order to gain information about biological impact on material properties?

Summary / Recommendation:

Part 1

The existing standards for testing knee joint implants were found to be fully technically feasible and clinically relevant. In addition, the urgent need for further standards was identified (see also Part 2 and 3 of this Statement).

There is a principled need to include instrumentation in pre-clinical testing. Several aspects are important here: On the one hand, the direct interaction of instrument and implant, such as the potential damage to gliding surfaces by impact instruments. On the other hand, indirect influences such as the alignment or the generation of third-body particles on the wear of the implant. Last but not least, the stability of the surgical instruments in the operating room and their correct applicability is of decisive importance for long-term success.

Part 2

Various types of clinical failure modes can be addressed in pre-clinical testing by tests not standardized as well as by improvements of standardized tests.

Standardized wear testing of TKR simulates level walking under clean lubrication conditions. Testing under increased and varying loads derived from daily living activities as well as third body particles such as bone cement is known to enhance wear and shall be regarded for pre-clinical testing. The distinction between backside wear and regular wear can be crucial in understanding particle sources.

The primary strength of cemented and uncemented TKRs is crucial for the long-term anchorage of the implant and thus for clinical success. Micromotion, as well as bone-cement adhesion strength, shall be investigated pre-clinically. Please see EFORT Consensus Statements 15 and 16 for detailed information.

The selection of the correct implant sizes or size combinations is usually not specified in the standards and requires an in-depth worst-case analysis. Please see also EFORT research topic 11 and 12 – Size range and anatomical design.

Fretting corrosion as well as biological degradation and corrosion are known clinical failures and also need to be investigated in preclinical testing.

Part 3

Q1: In silico methods (e.g. Finite Element Analysis – FEA) are an important complement to established preclinical testing methods. In terms of patient safety, FEA can make an important contribution to the selection of the worst-case implant size or its worst-case load application. Each FEA shall be performed under the conditions of the procedure described in ASME V&V40. The calculation by means of FEA does not replace final physical testing in any case.

Testing of a reference device is a proven method to establish implant safety. The lack of generally accepted acceptance criteria as well as the lack of transferability of test results between different test laboratories in some cases makes reference testing sometimes unavoidable. Since the procurement of reference products is fraught with great difficulties, the definition of general acceptance criteria must be accelerated and the comparability of test results increased.

Q2: Most of the materials used for knee arthroplasty are bioinert. Aspects of corrosion are described in detail under EFORT research topic 13 – Modularities/interfaces.

No scientifically generally accepted method is currently available for the aging of polyethylene in vivo. The most promising approach at present consists of squalene conditioning before oxygen aging according to ASTM F2003.

2. Level of Evidence

NA/Moderate

3. Consensus Delegate Vote

Part 1

98% – unanimous, strongest consensus (98% agree / 2% disagree / 0% abstain)

Part 2

91% – super majority, strong consensus (91% agree / 3% disagree / 6% abstain)

Part 3

95% – unanimous, strongest consensus (95% agree / 2.5% disagree / 2.5% abstain)

4. Part 1

a. Graphical Abstract

b. Search Strategy

For the third sub question of Part 1, the following search strategies were applied:

A pubmed (MEDLINE) literature review with the keywords ((knee) AND (instrument)) AND ((particle) OR (debris)) did reveal 133 results. A total of four papers have been found relevant.

A pubmed (medline) literature review with the keywords ((hip) OR (knee) AND ((instrument*) OR (tool*)) AND (intraoperative*) AND ((scratch*) OR (surface damage) OR (crack*))) did reveal 86 results. None of the papers has been found relevant for the topic. Therefore, the expert group has been asked to provide references known to be relevant to the topic.

c. Rationale

Which standard methods/tests according to norms are required for CE registration and what validity do they have with regard to the clinical use?

The group did not identify any superfluous test standards intended for pre-clinical testing of TKR. The upcoming revision of ISO 21536:2007 will provide a summary of ISO as well as ASTM standards available and regarded as essential for pre-clinical testing of TKR.

Test standard	Title	Implant type applicable	Clin	Notes
ISO 21536:2007	Non-active surgical implants – Joint replacement implants – Specific requirements for knee-joint replacement implants	ALL	NA	Includes summary of the tests listed in this table
ISO 14879:2000 (ASTM F1800-19e1)	Implants for surgery – Total knee-joint prostheses – Part 1: Determination of endurance properties of knee tibial trays	ALL but not UC or monobloc tibial component	[1,2]	Tibial tray fracture has rarely been reported in modern TKA. The pre-clinical acceptance level of 900N (see ISO 21536) seems to be too low for stem type designs and too high for peg type designs[3].
ASTM F3140-17	Standard Test Method for Cyclic Fatigue Testing of Metal Tibial Tray Components of Unicondylar Knee Joint Replacements	UC	[4]	Fatigue fracture of the femoral component has been reported also[5].
ISO 14243-1:2009	Implants for surgery -- Wear of total knee-joint prostheses -- Part 1: Loading and displacement parameters for wear-testing machines with load control and corresponding environmental conditions for test	ALL	[6]	Simulated ligament stiffness has been modified in 2009. Wear data measured by the previous standard can not be used for comparison.
ASTM F3141-14	Standard Guide for Total Knee Replacement Loading Profiles	ALL	NA	To be used in combination with TKR wear testing.
ISO 14243-2:2016	Implants for surgery -- Wear of total knee-joint prostheses -- Part 2: Methods of measurement	NA	NA	Supplemental information to ISO 14243-1 and -3
ISO 14243-3:2014	Implants for surgery -- Wear of total knee-joint prostheses -- Part 3: Loading and displacement parameters for wear-testing machines with displacement control and corresponding environmental conditions for test	ALL but not FA	[6]	The sign convention for ap-displacement was changed in the 2014 version of this standard. Wear data measured by the previous standard can not be used for comparison.
ASTM F2723-21	Standard Test Method for Evaluating Mobile Bearing Knee Tibial Baseplate/Bearing Resistance to Dynamic Disassociation	ALL	[7,8]	Standard for MB but useful for FB also.
ASTM F2722-21	Standard Test Method for Evaluating Mobile Bearing Knee Tibial Baseplate Rotational Stops	MB with rotational stop	[9]	none
ASTM F2724-21	Standard test Method for Evaluating Mobile Bearing knee Dislocation	MB	[7]	Often referred to as “spin-out” or “spit-out” test
ASTM F1814-15	Standard Guide for Evaluating Modular Hip and Knee Joint Components	ALL with modular connections	[10]	See research question 8 also.
ASTM F1223-20	Standard Test Method for Determination of Total Knee Replacement Constraint	ALL	NA	none
ASTM F2083-21	Standard Specification for Knee Replacement Prosthesis	ALL	NA	Standard includes test parameters for contact pressure also.
ASTM F1672-14	Standard Specification for Resurfacing Patellar Prosthesis	ALL TKRs with patella surface replacement	NA	Standard includes test parameters for contact pressure also.
ISO 14243-5:2019	Implants for surgery – Wear of total knee prostheses – Part 5: Durability performance of the patellofemoral joint	PA	[11]	none
ASTM F2777-16	Standard Test Method for Evaluating Knee Bearing (Tibial Insert) Endurance and Deformation Under High Flexion	ALL	[12]	none

The question of the mandatory application of a standard was linked to its clinical relevance and presented in tabular form. In this context, the associated clinical damage cases were included as examples via a literature reference. The referenced literature does not claim completeness and serves as an introduction to further analysis.

Several additional tests such as fatigue loading of femoral condyles, third body particle wear, tibial post fatigue, and contact pressure are under development. These tests are already available as state-of-the-art but have not yet been implemented in test standards.

All testing standards are based on clinical failure mechanisms and were mostly developed after the fact. For this reason, the question is not one of clinical relevance but of the predictive power of the procedures, especially in view of future designs and materials. Standardization authorities are urged to increase the speed of development in order to convert procedures known in the literature into test standards.

Abbreviations used

TKR	Total knee replacement,	MB	Mobile bearing
FA	Fixed axis	FB	Fixed bearing
UC	Unicondylar	PA	Patella replacement

Is it necessary to take into consideration the instrumentation in this process?

The influence of instruments on the clinical outcome of a TKR is not described in the literature in an overarching manner. Therefore, a more focused literature review based on known failure modes has been performed but also found not expedient (see literature search).

Surgical instruments are known to generate metal particles which can increase the wear rate. Three papers[13–15] do describe the generation of metal particles by oscillating saws and cutting blocks. The authors point out that the metal particles generated by the instruments can enter the articulation and thus enhance the wear rate of the implant leading to aseptic loosening. One paper[16] describes a corresponding effect found for the revision of an uncemented trabecular metal implant. Further literature13–15 describes the generation of metal particles by oscillating saws and cutting blocks. The authors point out that the metal particles generated by the instruments can enter the articulation and thus enhance the wear rate of the implant leading to aseptic loosening. One paper16 describes a corresponding effect found for the revision of an uncemented trabecular metal implant.

In conclusion, the generation of metal particles by oscillating instruments is to be expected in general but is not specific for certain implant/instrument combinations. The sensitivity of a TKR to particle contaminations shall be investigated in general (see EFORT Consensus research topic 7 to 10 - Biotribology).

Surgical instruments are also known to damage implant surfaces when not used carefully. A pubmed search was performed to identify relevant references (see literature search strategy). Vanlommel et al.[17] did simulate the clinical impaction of CoCr as well as an Oxidized Zirconium

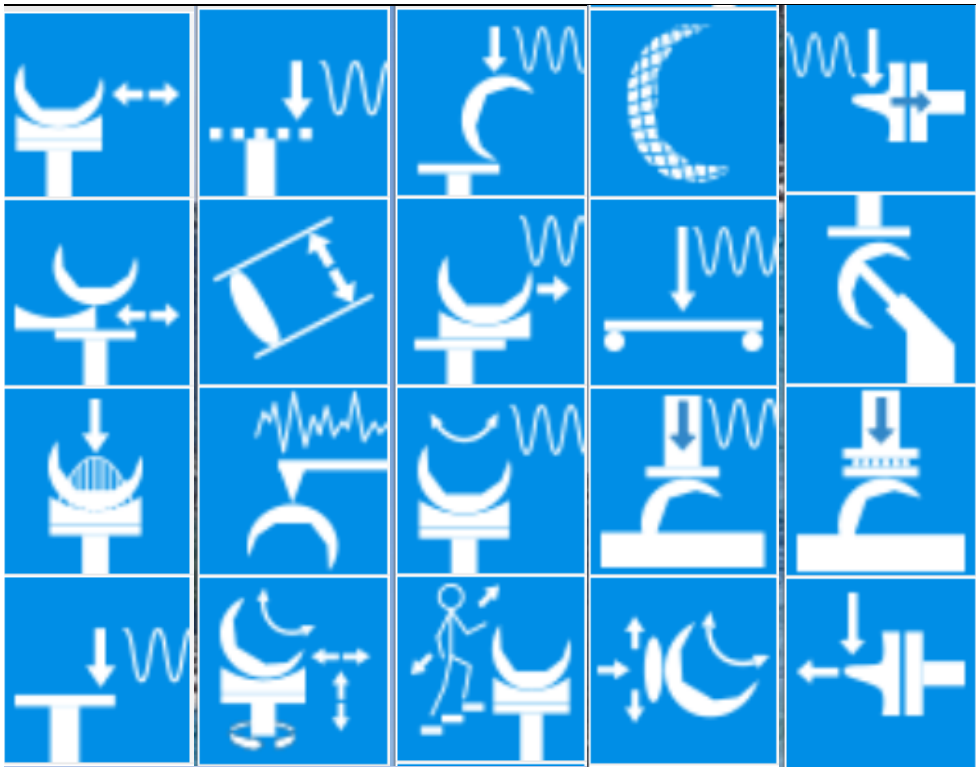
(OxZr) femoral components and found no increase in surface roughness at the impaction area. Burnell et al.[18] did analyze posterior surface damage of OxZr femoral components and found surface damage due to intraoperative contact between the femoral component and the tibial tray and concludes that manufacturers should supply the proper instrumentation to prevent this type of damage. In a meta-study performed by van Hove[19], the sensitivity of TiN coatings to surface damage caused by improper use of instruments is highlighted. In conclusion: No evidence for the failure of a specific instrument/implant combination has been found. In general, damage to surface-modified implants must be avoided by proper instrumentation and training.

The aspect of implant alignment and wear is discussed in Part 2 and 3. Generally speaking, it should be noted that the instruments provided have a significant impact on the accuracy of implantation and thus implant wear[20,21] and may need to be the target of a preclinical investigation.

The same applies to the use of an electrocauter which may lead to damage of implant surfaces causing implant fracture or increase of the wear rate[22,23].

5. Part 2

a. Graphical Abstract



b. Search Strategy

NA

c. Rationale

Wear

Heavy-duty load cycles

Wear testing of TKR according to ISO standards based on level walking is an appropriate method to compare the wear behavior of implant designs. However, different clinically reported failure mechanisms could not be replicated by simulating just level walking[24].

Taking daily activities into account in the wear loading protocol by alternating different high demanding activities (see ASTM F3141 also), combined with artificial aging of the polymers (see ASTM F2003 and squalene aging25), it could be shown that the level of wear characterization can be increased substantially. A replication of clinically relevant failure modes like delamination of the polymeric components is possible[26,27]. This method can also be used for the evaluation of the cement-implant interface.

A detailed rationale for wear testing under high demanding activities is given under EFORT Consensus research topic 7 to 10 – Biotribology.

Third body wear debris

Third body particles released from bone cement or surface coatings are known to increase the in-vivo wear rate and thus reduce the implant life[28]. The influence of third bodies on the wear rate can be determined in the laboratory[29]. There are two methods for pre-clinical testing: The addition of particles to the test fluid and roughening of the components prior to testing. The advantage of the former method is that it can also be used to investigate coatings and new materials.

ASTM is currently developing a standard for in vitro testing in which alumina particles[30] of size 100µm and 300µm are proposed at a rate of 2mg every 0.5 million cycles.

Misalignment

The alignment of TKA components could vary due to presurgical alignment, the accuracy of the surgical instrumentation and due to patient factors, such as soft tissue balance. However, this alignment conditions, which are different from ideal are more often than not clinical reality, could have a significant effect on the kinematics and wear and thus on the anterior knee pain, instability or aseptic loosening.

In terms of evaluation for robustness of TKA designs a wider range of component alignment conditions could be investigated by experimental[21] or computational approaches[31].

Backside wear of tibial inserts

25-30% of the total wear may be related to backside wear with smaller particles compared to articular wear particles 32. However, the analytical approach is challenging and as backside wear is part of the overall wear (including the backside wear particles) a separate analysis of backside wear does not appear to be mandatory. Nevertheless, if backside wear should be evaluated in particular, appropriate methods are available[33].

Primary stability

Tibial and femoral cement adhesion

Primary stability of cemented implants is discussed in more detail in the EFORT Consensus Statement 15, but a brief summary is given here. Debonding of the femoral and tibial implant-cement interface debonding may lead to early failure, even when this may not be evident from radiographs[34]. Test methods exist to evaluate the bonding strength between bone cement and the implant surface. Evaluation of interface debonding in full reconstructions is mainly performed through physical testing in which the pull-off strength of tibial[35] and femoral[36] reconstructions is compared. Although pull-off or push-off testing configurations are not representative for in vivo loading conditions, they do allow for qualitative comparisons between implant systems. Contamination of the interface with marrow, fat, or blood have a significant effect on cement adhesion to the implant[37]. Computational modeling using finite element analysis can provide insights into implant-cement interface debonding38, but the validity of such simulations depends on the quality of the input parameters characterizing the strength of the bond between the implant and the cement.

Primary stability of cementless implants

Primary stability of cementless implants is discussed in more detail in the EFORT Consensus Statement 16, but a brief summary is given here. Primary stability of cementless implants is a prerequisite for obtaining secondary stability through growth of bone on and into the implant surface. Animal studies have demonstrated that osseointegration of the implant depends on the micromotions at the interface, with micromotions exceeding 150 µm leading to the formation of a soft tissue layer that may obstruct secondary fixation[39]. Mechanical testing therefore mainly focuses on the measurement of implant-bone interface micromotions under loading conditions replicating activities of daily living. Reconstructions are made either in cadaveric[40] or synthetic[41] bone, the former providing conditions that closely approximate the clinical setting, while the latter minimizes inter-specimen variability. Finite element modeling allows for evaluation of micromotions over the full implant surface, including locations that are not visible in experimental testing[42]. Important input parameters to these models are the coefficient of friction, the amount of press fit (or interference fit), and (non-linear) material properties of the implant and bone.

Implant sizes and worst-case analysis

Testing of TKR is typically based on the worst-case approach. A distinction must be made between the testing of individual components and the combination of components. For individual components, maximum stress is the dominating factor for most tests. The correlation between implant size and patient weight is described as only moderate [43] which means that a reduction of the load for small sizes is not permissible. Finite Element Analysis (FEA) is a efficient tool to pre-select the implant size to be physically tested. Currently there are two specific FEA knee standards available: ASTM F3334 to simulate tibial tray fatigue (ISO 14879 and ASTM F1800) and ASTM F3161 to simulate knee femoral components under closing conditions (ASTM physical test standard upcoming).

Wear testing is historically performed on the mean size of the implant system as wear test standards such as ISO 14243-1 are based on loads and soft tissue constraints of average patients. Under this aspect, the combination of implant sizes might be important: In some systems, for example, a tibial component can be used with several femoral components. This also applies to the use of a patella replacement with multiple femoral implant sizes. In this case, physical testing of multiple combinations is necessary.

Fretting corrosion, degradation and biological corrosion

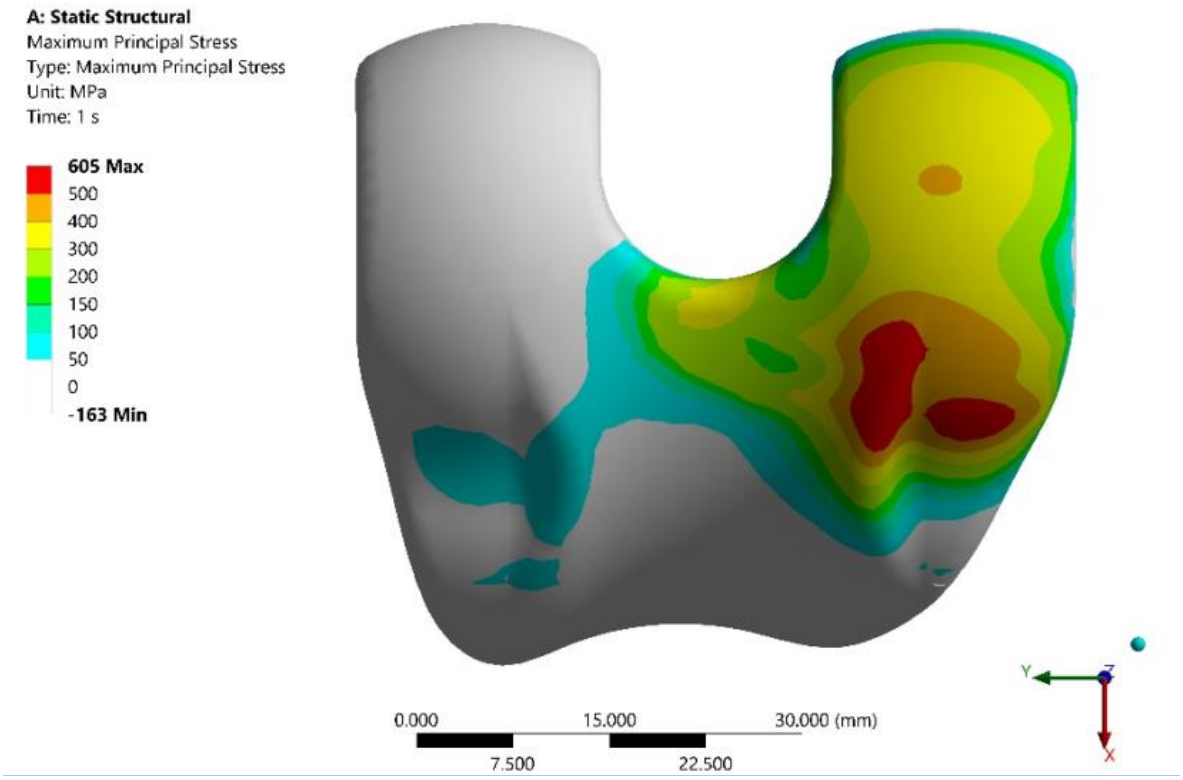
Fretting corrosion, degradation and biological corrosion are not unique for TKR. Please see EFORT Consensus Statement 4 as well as 13 and 14 for further information.

Static and fatigue fracture

Fatigue fracture of femoral components Uni and bicondylar) has been observed clinically and is not currently captured in preclinical testing via standardized testing[44,45]. Currently, a corresponding standardization procedure exists at ASTM. Static fractures have been observed during impaction of ceramic femoral components and are also the target of ASTM standardization currently underway[46.22,23]. Currently, a corresponding standardization procedure exists at ASTM. Static fractures have been observed during impactation of ceramic femoral components and are also the target of ASTM standardization currently underway[46].

6. Part 3

a. Graphical Abstract



b. Search Strategy

NA

c. Rationale

Which supplementary steps (methods) are imaginable to further close the gap between pre-clinical and clinical application (e.g. use of a numerical simulator or comparative testing with established products)?

Numerical simulation of implants ("in-silico") is of increasing importance for pre-clinical testing. On the one hand, it can provide knowledge about the complex load application in vivo[47]; on the other hand, it allows direct comparison between different implant sizes or implant combinations[48,49]. One of the fundamental problems of FEA is that the calculation leads to a result in almost any case, but the reliability of the results is often difficult to judge. To increase the reliability of an FEA, it must be performed based on ASME V&V40[50] which describes the principles of verification and validation. Based on this standard, the ASTM F3334 and F3161 test specifications provide applications for components of a TKR.

The in-silico method has two major limitations in the field of knee arthroplasty: On the one hand, the calculation of the implant fatigue strength is not aggravated by the lack of known material properties for modified implant surface (coatings, structured surfaces) or manufacturing depending on variations in the alloy microstructures, on the other hand, the complex interactions of multiple parameters that define the wear behavior of a TKR could not be predicted satisfactorily so far[51]. In summary, FEA has proven to be a valuable tool for determining the worst-case implant but is not a substitute for subsequent physical testing.

One method that has been used by the FDA for some time already, is the testing of a "reference device", i.e. a reference product that has proven itself insufficient and safe clinical use. This approach has the advantage that the comparability of measured values between different testing laboratories is not of importance. On the other hand, the procurement of comparative implants occasionally presents smaller manufacturers in particular with insurmountable obstacles. In the medium term, the quality of test results must be improved by improving the standardization, training, and technical equipment of testing laboratories to make multiple (reference) tests obsolete.

What are methods that simulate the in vivo conditions for load testing of implants in order to gain information about biological impact on material properties?

Metallic and ceramic implant materials are largely bioinert. Inflammatory cell-induced corrosion (ICIC) has been detected on metal retrievals[52] but the impact on implant structural stability and loosening is unknown. Further information on implant corrosion can be found in EFORT research topic 13 and 14 – Modularities/interfaces.

Polyethylene materials are known to reduce their material properties under in-vivo conditions. In particular, standard UHMWPE gamma sterilized in air shows significantly reduced fatigue strength and wear resistance[53]. Modern polyethylene such as vitamin E polyethylene show clinically significantly better aging behavior54 but the pre-clinical methods to investigate the aging behavior are still under discussion:

The test standard ASTM F2003 describes oxygen aging under elevated temperature and pressure for two weeks. A study performed by Grupp et al.[55] did show reduced wear rates for modern polyethylene but oxidation bands have not been found to correlate to retrievals. Corresponding results have been published by Oral et al. even when extending the aging period[56].

Accelerated aqueous (PBS or distilled water) aging at a temperature between 37 and 60°C has been found to correlate to oxidative patterns encountered in vivo[57] for some polyethylene materials.

Accelerated aging methods incorporating lipid absorption (squalene) combined with the previously mentioned oxygen aging method58 are a promising method to replicate the in-vivo aging of new polyethylene materials[59,60,25] and are expected to become a standard procedure in the future.

The current standardized method for aging according to ASTM F2003 is considered insufficient for polyethylene in clinical use. An improvement of the standard is currently under development.

7. References

1. Chatterji U, Ashworth MJ, Smith AL, et al. **Retrieval study of tibial baseplate fracture after total knee arthroplasty. J Arthroplasty 2005; 20: 101–107.**

2. Callaghan JJ, DeMik DE, Bedard NA, et al. **Tibial tray fracture in a modern prosthesis with retrieval analysis. Arthroplast Today 2018; 4: 143–147.**

3. Stoller A (ed). **Is the Standardized Performance Requirement for Tibial Tray Cantilever Fatigue Strength Appropriate?: Poster 0940, 2013.**

4. Manzotti A, Chemello C, Pullen C, et al. **An uncommon cause of cemented unicompartmental knee arthroplasty failure: fracture of metallic components. Knee Surg Sports Traumatol Arthrosc 2013; 21: 2518–2522.**

5. Cameron HU and Welsh RP. **Fracture of the femoral component in unicompartmental total knee arthroplasty. J Arthroplasty 1990; 5: 315–317.**

6. Chakravarty R, Elmallah RDK, Cherian JJ, et al. **Polyethylene Wear in Knee Arthroplasty. J Knee Surg 2015; 28: 370–375.**

7. Fisher DA, Bernasek TL, Puri RD, et al. **Rotating platform spinouts with cruciate-retaining mobile-bearing knees. J Arthroplasty 2011; 26: 877–882.**

8. Voskuul T_**Acta Orthop Belg_2015_ Insert dissociation after fixed bearing PS constrained Genesis II total knee arthroplasty. A case series of nine patients.**

9. Sudanese A, Castiello E and Affatato S. **Mobile-bearing total knee arthroplasty: a full traumatic rotation of 180°. Int J Artif Organs 2013; 36: 421–425.**

10. Kahan M and Estes C. **Failure at the femoral stem extension-condylar interface in a rotating hinge knee without radiographic evidence of loosening. Arthroplast Today 2020; 6: 36–40.**

11. Hambright DS, Watters TS, Kaufman AM, et al. **Fracture of highly cross-linked all-polyethylene patella after total knee arthroplasty. J Knee Surg 2010; 23: 237–240.**

12. **Polyethylene failure in New Jersey low-contact stress total knee arthroplasty.**

13. Hooper J, Schwarzkopf R, Fernandez E, et al. **Feasibility of single-use 3D-printed instruments for total knee arthroplasty. Bone Joint J 2019; 101-B: 115–120.**

14. Sydney SE, Pickering SA, Bell CGR, et al. **Reducing metal debris generation during total knee arthroplasty. ORTHO 2007; 30: 999–1000.**

15. Gotterson PR, Nusem I, Pearcy MJ, et al. **Metal debris from bony resection in knee arthroplasty--is it an issue? Acta Orthop 2005; 76: 475–480.**

16. Sanchez Marquez JM, Del Sel N, Leali A, et al. **Case reports: Tantalum debris dispersion during revision of a tibial component for TKA. Clin Orthop Relat Res 2009; 467: 1107–1110.**

17. Vanlommel J, Porteous AJ, Hassaballa MA, et al. **Iatrogenic surface damage during femoral component impaction in total knee arthroplasty. Knee 2010; 17: 43–47.**

18. Burnell CDC, Brandt J-M, Petrak MJ, et al. **Posterior condyle surface damage on retrieved femoral knee components. J Arthroplasty 2011; 26: 1460–1467.**

19. van Hove RP, Sierevelt IN, van Royen BJ, et al. **Titanium-Nitride Coating of Orthopaedic Implants: A Review of the Literature. Biomed Res Int 2015; 2015: 485975.**

20. D'Lima DD, Hermida JC, Chen PC, et al. **Polyethylene wear and variations in knee kinematics. Clin Orthop Relat Res 2001: 124–130.**

21. Johnston H, Abdelgaied A, Pandit H, et al. **The effect of surgical alignment and soft tissue conditions on the kinematics and wear of a fixed bearing total knee replacement. J Mech Behav Biomed Mater 2019; 100: 103386.**

22. Sonntag R, Gibmeier J, Pulvermacher S, et al. **Electrocautery Damage Can Reduce Implant Fatigue Strength: Cases and in Vitro Investigation. J Bone Joint Surg Am 2019; 101: 868–878.**

23. Miller KC, Morrow BR, Sorrels JH, et al. **Electrocautery Induced Damage of Total Knee Implants. J Arthroplasty 2021; 36: 1126–1132.**

24. Grupp TM, Kaddick C and Baxmann M. **Biomechanische Testung in der Orthopädie und Wirbelsäulenchirurgie. Osteologie 2015; 24: 152–157.**

25. Sakoda H, Okamoto Y and Haishima Y. **In vitro estimation of reduction in strength and wear resistance of UHMWPE for joint prostheses due to lipid-induced degradation. J Biomed Mater Res B Appl Biomater 2020; 108: 3155–3161.**

26. Schwiesau J, Schilling C, Kaddick C, et al. **Definition and evaluation of testing scenarios for knee wear simulation under conditions of highly demanding daily activities. Med Eng Phys 2013; 35: 591–600.**

27. Zietz C, Reinders J, Schwiesau J, et al. **Experimental testing of total knee replacements with UHMW-PE inserts: impact of severe wear test conditions. J Mater Sci Mater Med 2015; 26: 134.**

28. Malikian R, Maruthainar K, Stammers J, et al. **In vivo roughening of retrieved total knee arthroplasty femoral components. Knee 2014; 21: 278–282.**

29. Schroeder C, Grupp TM, Fritz B, et al. **The influence of third-body particles on wear rate in unicondylar knee arthroplasty: a wear simulator study with bone and cement debris. J Mater Sci Mater Med 2013; 24: 1319–1325.**

30. Villiers D de, Traynor A, Collins SN, et al. **The increase in cobalt release in metal-on-polyethylene hip bearings in tests with third body abrasives. Proc Inst Mech Eng H 2015; 229: 611–618.**

31. Popescu R, Haritinian EG and Cristea S. **Relevance of Finite Element in Total Knee Arthroplasty - Literature Review. Chirurgia (Bucur) 2019; 114: 437–442.**

32. Braun S, Uhler M, Bormann T, et al. **Backside wear in total knee replacement: A new quantitative measurement method and a comparison of polished cobalt-chromium tibial trays with titanium tibial trays. Wear 2021; 466-467: 203552.**

33. Ngai V, Kunze J, Cip J, et al. **Backside wear of tibial polyethylene components is affected by gait pattern: A knee simulator study using rare earth tracer technology. J Orthop Res 2020; 38: 1607–1616.**

34. Sadauskas A, Engh C, Mehta M, et al. **Implant Interface Debonding After Total Knee Arthroplasty: A New Cause for Concern? Arthroplast Today 2020; 6: 972–975.**

35. Gebert de Uhlenbrock A, Püschel V, Püschel K, et al. **Influence of time in-situ and implant type on fixation strength of cemented tibial trays - a post mortem retrieval analysis. Clin Biomech (Bristol, Avon) 2012; 27: 929–935.**

36. Bergschmidt P, Dammer R, Zietz C, et al. **Adhesive strength of total knee endoprostheses to bone cement - analysis of metallic and ceramic femoral components under worst-case conditions. Biomed Tech (Berl) 2016; 61: 281–289.**

37. Billi F, Kavanaugh A, Schmalzried H, et al. **Techniques for improving the initial strength of the tibial tray-cement interface bond. Bone Joint J 2019; 101-B: 53–58.**

38. Zelle J, Janssen D, van Eijden J, et al. **Does high-flexion total knee arthroplasty promote early loosening of the femoral component? J. Orthop. Res. 2011; 29: 976–983.**

39. Kienapfel H, Sprey C, Wilke A, et al. **Implant fixation by bone ingrowth. J Arthroplasty 1999; 14: 355–368.**

40. Yang H, Bayoglu R, Renani MS, et al. **Validation and sensitivity of model-predicted proximal tibial displacement and tray micromotion in cementless total knee arthroplasty under physiological loading conditions. J Mech Behav Biomed Mater 2020; 109: 103793.**

41. Morwood MP, Guss AD, Law JI, et al. **Metaphyseal Stem Extension Improves Tibial Stability in Cementless Total Knee Arthroplasty. J Arthroplasty 2020; 35: 3031–3037.**

42. Berahmani S, Janssen D and Verdonschot N. **Experimental and computational analysis of micromotions of an uncemented femoral knee implant using elastic and plastic bone material models. J Biomech 2017; 61: 137–143.**

43. Blevins JL, Rao V, Chiu Y-F, et al. **Predicting implant size in total knee arthroplasty using demographic variables. Bone Joint J 2020; 102-B: 85–90.**

44. Chithartha K, Thilak J, Sukesh AN, et al. **Fatigue fracture of the femoral component in total knee replacement. Knee 2020; 27: 1439–1445.**

45. Gilg MM, Zeller CW, Leitner L, et al. **The incidence of implant fractures after knee arthroplasty. Knee Surg Sports Traumatol Arthrosc 2016; 24: 3272–3279.**

46. Bergschmidt P, Lohmann C, Ganzer D, et al. **Knieendoprothetik mit keramischen Femurkomponenten: Nationale prospektive Multicenterstudie zum klinischen und radiologischen Outcome. Orthopade 2011; 40: 224–230.**

47. Sintini I, Fitzpatrick CK, Clary CW, et al. **Computational evaluation of TKR stability using feedback-controlled compressive loading. J Orthop Res 2018; 36: 1901–1909.**

48. Woiczinski M, Maas A, Grupp T, et al. **Realitätsnahe Finite-Elemente-Simulation in der präklinischen Testung von Knie- und Hüftimplantaten. Orthopade 2020; 49: 1060–1065.**

49. Dharia MA, Snyder S and Bischoff JE. **Computational Model Validation of Contact Mechanics in Total Ankle Arthroplasty. J Orthop Res 2020; 38: 1063–1069.**

50. ASME V&V 40-2018. **Assessing credibility of computational modeling through verification and validation: Application to medical devices.: New York, NY: American Society of Mechanical Engineers.**

51. Burchardt A, Abicht C and Sander O. **An efficient and robust simulator for wear of total knee replacements. Proc Inst Mech Eng H 2020; 234: 921–930.**

52. Arnholt CM, MacDonald DW, Malkani AL, et al. **Corrosion Damage and Wear Mechanisms in Long-Term Retrieved CoCr Femoral Components for Total Knee Arthroplasty. J Arthroplasty 2016; 31: 2900–2906.**

53. MCGOVERN TF, Ammeen DJ, Collier JP, et al. **Rapid polyethylene failure of unicondylar tibial components sterilized with gamma irradiation in air and implanted after a long shelf life. J Bone Joint Surg Am 2002; 84: 901–906.**

54. Lambert B, Neut D, van der Veen HC, et al. E**ffects of vitamin E incorporation in polyethylene on oxidative degradation, wear rates, immune response, and infections in total joint arthroplasty: a review of the current literature. Int Orthop 2019; 43: 1549–1557.**

55. Grupp TM, Fritz B, Kutzner I, et al. **Vitamin E stabilised polyethylene for total knee arthroplasty evaluated under highly demanding activities wear simulation. Acta Biomater 2017; 48: 415–422.**

56. Oral E, Neils AL, Doshi BN, et al. **Effects of simulated oxidation on the in vitro wear and mechanical properties of irradiated and melted highly crosslinked UHMWPE. J Biomed Mater Res B Appl Biomater 2016; 104: 316–322.**

57. Mazzucco DC, Dumbleton J and Kurtz SM. **Can accelerated aqueous aging simulate in vivo oxidation of gamma-sterilized UHMWPE? J Biomed Mater Res B Appl Biomater 2006; 79: 79–85.**

58. Oral E, Fung M, Rowell SL, et al. **In-vitro oxidation model for UHMWPE incorporating synovial fluid lipids. J Orthop Res 2018; 36: 1833–1839.**

59. Reinitz SD, Currier BH, van Citters DW, et al. **Oxidation and other property changes of retrieved sequentially annealed UHMWPE acetabular and tibial bearings. J Biomed Mater Res B Appl Biomater 2015; 103: 578–586.**

60. Rowell SL, Reyes CR, Malchau H, et al. **In Vivo Oxidative Stability Changes of Highly Cross-Linked Polyethylene Bearings: An Ex Vivo Investigation. J Arthroplasty 2015; 30: 1828–1834.**

RESEARCH TOPIC 6

Endurance Habitual and Peak Loads in Hip and Knee Arthroplasty – Review and Summary of Aspects to be Considered

Authors

Michael Morlock[1], Jan Philippe Kretzer[2], Ronja A. Schierjott[3], Francesco Traina[4,5], Ricardo Larrainzar-Garijo[6], Georg N. Duda[7], Christian Kaddick[8]

[1]Hamburg University of Technology, Hamburg, Germany
[2]Laboratory of Biomechanics & Implant Research, Department of Orthopaedic Surgery, Heidelberg University Hospital, Heidelberg, Germany
[3]Aesculap AG Research & Development, Tuttlingen, Germany
[4]Ortopedia-Traumatologia e Chirurgia protesica e dei reimpianti d'anca e di ginocchio, IRCCS Istituto Ortopedico Rizzoli, Bologna, Italy
[5]DIBINEM, Università di Bologna, Bologna, Italy
[6]Orthopaedics and Trauma Department, Medicine School, Universidad Complutense, Hospital Universitario Infanta Leonor, Madrid, Spain
[7]Julius Wolff Institute and Center for Musculoskeletal Surgery, Berlin Institute of Health at Charité - Universitätsmedizin Berlin, Berlin, Germany
[8]EndoLab Mechanical Engineering GmbH, Riedering, Germany

1. Opening Comments

Mechanical implant failure has clinically become a minor concern. This was made possible by the improvement of the pre-clinical testing methods and the awareness in the industry that testing beyond the "necessary" standards is required to prevent mechanical failures after the introduction of a new implant. It has to be emphasized that the evolved testing methods have allowed to prevent many implants from entering the market.

Most problems occurring clinically are related to the interface between the implant and bone or implant and cement. This is difficult to address sufficiently by pre-clinical testing. In this field further improvements are needed, as are improvements needed simulating the variations in the surgical processes between users.

2. Abstract

Question: Mechanical component testing

How can be established from a clinical perspective that the implant will withstand the endurance habitual and peak loads that must reasonably be expected (i.e. single implant parts as well as complete arthroplasty combination)?

Summary/Recommendation:

Presently catastrophic implant failure as the origin of clinical failure is rare as long as the implant is loaded according to its designed purpose, especially with respect to the load transfer between implant and bone. Aspects to be considered are:

- The execution of established and standardized pre-clinical testing methods for the determination of strength and endurance properties are the essential basis for successful pre-clinical assessment. These have to be performed under physiological conditions (i.e. considering the pH-value in the body and potential corrosion processes, especially for implants with interfaces such as modular implants – see also Consensus Statement 13: Modularities / Interfaces) and worst-case size combinations (e.g. based on FEA).
- The combination matrix for implant systems (size range of component 1 compatible with size range of component 2) has to be respected in the clinical application to avoid failure due to overloading (e.g. for TKA components to find suitable size combinations and avoid failure of PE inserts due to mechanical overloading).
- Systematic analysis of retrieved implants from clinical studies and specific cases (even if the reason for retrieval was not mechanical failure, the retrievals could be analyzed for potential signs of fatigue)
- Careful analysis of registry results (especially for combinations) and their association with patient characteristics and the respective results from pre-clinical testing.
- If implants are used in a mix&match combination, which is typically not tested pre-clinically, the potential influence of this combination on overall survival is difficult to assess since frequently not both manufacturers are notified of adverse events. This is one more justification for public registries.

3. Level of Evidence

High

4. Consensus Delegate Vote

90% - super majority, strong consensus (90% agree / 3% / 7%)

5. Graphical Abstract

N/A

6. Search Strategy

N/A

7. Rationale

1. Which clinical examples of systematic mechanical implant failure are known from the past which could have been eliminated today by applying modern / up-to-date pre-clinical testing methods?

- The reason for clinical implant failure is predominantly due to failure of the interface between bone and implant or cement and implant or cement and bone associated with corrosion or wear products or mechanical impingement (1).
- Catastrophic implant failure is also seen at modular junctions of implants due to corrosion (2-4).
- Early / mid- / long-term catastrophic mechanical implant failure such as component fracture due to long term repetitive loading without detectable biological or corrosive cause are rare (e.g. mobile bearing tibial insert fracture (5)) due to the pre-clinical testing capabilities, which identify most problematic components.

- Meta-analyses regarding the duration in vivo until failure and the rate of failure are extremely rare (4). The incidence of implant fractures in total hip arthroplasty in pooled worldwide arthroplasty register datasets is very low at 304 fractures per 100,000 implants (6,7). This is the consequence of the good established pre-clinical testing methods. However, this rate might be higher for specific designs.

2. How is it possible in clinical applications to identify insufficient mechanical characteristics, i.e. is it possible to identify clinical implant failure as undeniably associated with poor endurance properties of the implant?

- Only if the implant fails catastrophically in the patient, which could also be due to the loading situation in the patient and not the poor characteristics of the implant (e.g. diaphysal fixation of a stem intended for proximal fixation).
- Only if the implant fails catastrophically and "catastrophic" loading situations / e.g. trauma or car accidents can be excluded
- Potential surgery or implant combination related factors (e.g. damage to the stem taper (8), damage to an implant by spark transition of high frequency instruments (9-11), insufficient assembly force (12), mismatch of sizes (13, 14) can significantly reduce the fatigue fracture toughness.
- Clinical mechanical implant failures of approved implants undeniably associated with the implant itself can only be identified by registries which eliminate the surgeon bias.

3. How to differentiate between mix and match or mismatch?

- Mix&match: combination of nominally fitting components but not approved by the respective manufacturers – can be clinically successful as shown in the registries (15).
- Mismatch: combination of nominally NOT fitting components (non-compatible) not approved by the manufacturer. Does not always lead to immediate failure. If failure occurs it can be directly related to the combination (e.g. clear geometrical or size mismatch) (13)
- Working group 2 is addressing this issue in full detail.

8. References

1. Hirakawa, K., Jacobs, J. J., Urban, R., Saito, T., (2004). **Mechanisms of failure of total hip replacements: lessons learned from retrieval studies. Clin.Orthop Relat.Res. 10-17.**

2. Mertl, P., Dehl, M., (2020). **Femoral stem modularity. Orthop Traumatol.Surg.Res. 106, S35-S42.**

3. Arnholt, C. M., MacDonald, D. W., Tohfafarosh, M., Gilbert, J. L., Rimnac, C. M., Kurtz, S. M., Klein, G., Mont, M. A., Parvizi, J., Cates, H. E., (2014). **Mechanically assisted taper corrosion in modular TKA. The Journal of Arthroplasty 29, 205-208.**

4. Krueger, D. R., Guenther, K. P., Deml, M. C., Perka, C., (2020). **Mechanical failure of 113 uncemented modular revision femoral components. Bone Joint.J. 102-B, 573-579.**

5. Yoon, J. R., Jeong, H. I., Oh, K. J., Yang, J. H., (2014). **Bilateral condyle fracture of tibial insert in mobile bearing total knee arthroplasty. Knee. 21, 318-321.**

6. Sadoghi, P., Pawelka, W., Liebensteiner, M. C., Williams, A., Leithner, A., Labek, G., (2014). **The incidence of implant fractures after total hip arthroplasty. Int.Orthop 38, 39-46.**

7. Gilg, M. M., Zeller, C. W., Leitner, L., Leithner, A., Labek, G., Sadoghi, P., (2016). **The incidence of implant fractures after knee arthroplasty. Knee.Surg.Sports.Traumatol.Arthrosc. 24, 3272-3279.**

8. Hothi, H. S., Panagiotopoulos, A. C., Whittaker, R. K., Bills, P. J., McMillan, R. A., Skinner, J. A., Hart, A. J., (2017). **Damage Patterns at the Head-Stem Taper Junction Helps Understand the Mechanisms of Material Loss. J.Arthroplasty 32, 291-295.**

9. Almansour, H., Sonntag, R., Pepke, W., Bruckner, T., Kretzer, J. P., Akbar, M., (2019). **Impact of Electrocautery on Fatigue Life of Spinal Fusion Constructs-An In Vitro Biomechanical Study. Materials (Basel.) 12.**

10. Zobel, S. M., Morlock, M. M., Huber, G., (2020). **Fatigue strength reduction of Ti-6Al-4V titanium alloy after contact with high-frequency cauterising instruments. Med.Eng.Phys. 81, 58-67.**

11. Sonntag, R., Gibmeier, J., Pulvermacher, S., Mueller, U., Eckert, J., Braun, S., Reichkendler, M., Kretzer, J. P., (2019). **Electrocautery Damage Can Reduce Implant Fatigue Strength: Cases and in Vitro Investigation. J.Bone Joint.Surg.Am. 101, 868-878.**

12. Rehmer, A., Bishop, N. E., Morlock, M. M., (2012). **Influence of assembly procedure and material combination on the strength of the taper connection at the head-neck junction of modular hip endoprostheses. Clin.Biomech.(Bristol., Avon.) 27, 77-83.**

13. Koper, M., Mathijsen, N., van Ravensway Claasen, H., Witt, F., Morlock, M. V. C., (2015). **Severe wear and pseudotumor formation due to taper mismatch in a total hip artroplasty.**

14. Hanks, G. A., Foster, W. C., Cardea, J. A., (1986). **Total hip arthroplasty complicated by mismatched implant sizes. Report of two cases. J.Arthroplasty 1, 279-282.**

15. Tucker, K., Pickford, M., Newell, C., Howard, P., Hunt, L. P., Blom, A. W., (2015). **Mixing of components from different manufacturers in total hip arthroplasty: prevalence and comparative outcomes. Acta Orthop. 86, 671-677.**

RESEARCH TOPIC

BIOTRIBOLOGY

RESEARCH TOPIC 7

Biotribological Methodologies in Wear Simulation and Wear Debris Characterisation in Hip and Knee Arthroplasty – Review and Introduction of Thresholds for Standard Testing

Authors

Thomas M. Grupp[1,2], Christian Kaddick[3], Claude Rieker[4], Jan Philippe Kretzer[5], John Fisher[6]

[1]Aesculap AG Research & Development, Tuttlingen, Germany
[2]Ludwig Maximilians University Munich, Department of Orthopaedic and Trauma Surgery, Musculoskeletal University Center Munich (MUM), Campus Grosshadern, Munich, Germany
[3]EndoLab Mechanical Engineering GmbH, Riedering, Germany
[4]Zimmer Biomet EMEA, Scientific Affairs, Zug, Switzerland
[5]Laboratory of Biomechanics & Implant Research, Clinic for Orthopedics & Trauma Surgery, Heidelberg University Hospital, Germany
[6]Institute of Medical & Biological Engineering, School of Mechanical Engineering, University of Leeds, United Kingdom

1. Abstract

Question: Biotribology

Can standard test methods in total hip and knee arthroplasty (THA, TKA) show that the planned articulations enable the function of the joint replacement throughout the expected implant lifetime without producing a critical amount of wear?

1. Which standard test methods according to standards (ISO, ASTM) are required for registration of an implant?
How meaningful are they for the actual clinical application?

Summary / Recommendation:

Standards for pre-clinical testing of wear and biotribological behaviour are particularly useful in comparing performance and function of new implants (under development) with existing substantially equivalent implants with clinical history. Appropriate clinically established devices should be chosen with similar articulation materials and same basis of similarity in design (i.e. cruciate retaining, sacrificing or posterior stabilized TKA implants), with a clinical history of between 5 and 10 years recorded through registers or trials should be used to demonstrate equivalence or improved performance of the new device. Current ISO 14242 and ISO 14243 standard test series are applied over two decades, are in continuous use in numerous laboratories and have been validated for their prediction of clinical behaviour.

Performance data from n = 257 wear simulation studies from four different laboratories has been reviewed, including the majority of today's clinically long-term successful THA and TKA implant designs and bearing materials. Considering the upper 75% percentile level for THA articulations with conventional polyethylene (CPE n = 112 studies) and highly cross-linked polyethylene (XLPE n = 53 studies), a reasonable threshold of 30 mg/million cycles for CoCPE, of 40 mg/million cycles for MoCPE and of 5 mg/million cycles for CoXLPE as well as for MoXLPE can be recommended (Fig. 1). For TKA CR and PS implant designs with CPE gliding surfaces, a reasonable threshold is given at 25 mg/ million cycles, performing knee wear simulation according to ISO 14243-1:2009 under load control (Fig. 2). Measuring average wear rates in THA ISO ISO 14242-1:2014 or TKA ISO 14243-1:2009 wear testing above the recommended thresholds in the four laboratories from where the data was collected, an optimisation in implant design, bearing material, manufacturing tolerances and/ or surface quality should be undergone. Considering the current unsatisfactory reproducibility and repeatability of wear results between some laboratories, these thresholds cannot be applied to tests undertaken in other laboratories, where differences in methods, test equipment or staff experience in biotribology may produce different wear rates. In other laboratories, comparison to data from a substantial equivalent device with clinical history in the same laboratory is needed for comparison.

2. Level of Evidence

Moderate

3. Consensus Delegate Vote

98% – unanimous, strongest consensus (98% agree / 0% disagree / 2% abstain)

4. Graphical Abstract

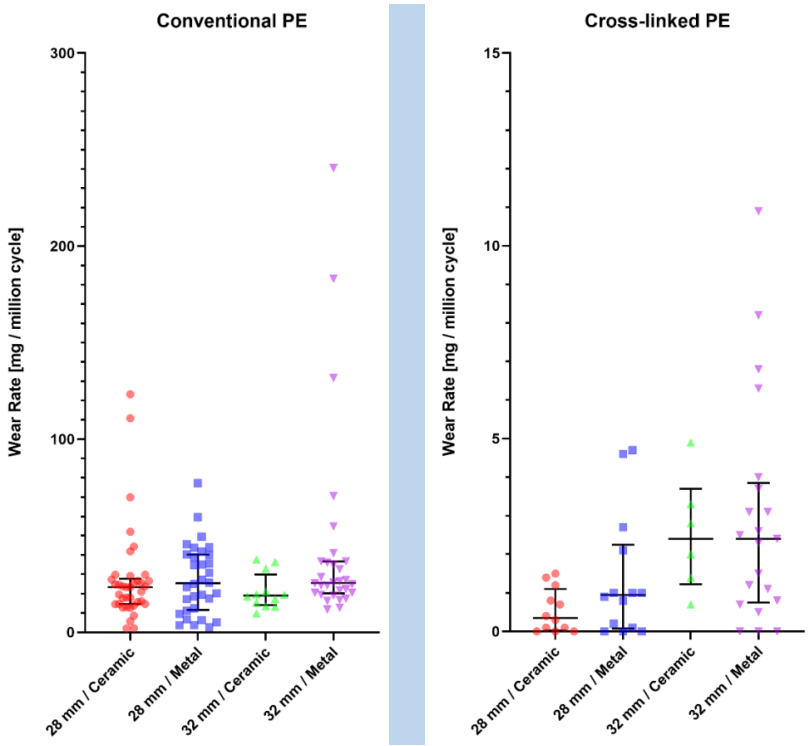


Figure 1: Wear rates (median, 25% & 75% percentile) of n=165 THA wear simulation studies according to ISO 14242-1

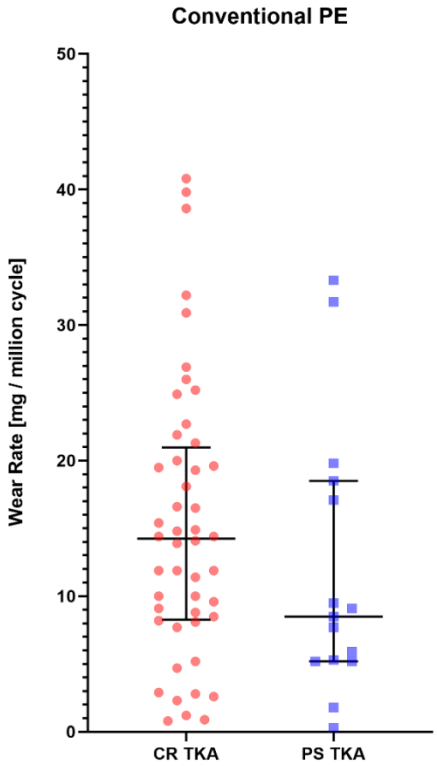


Figure 2: Wear rates (median, 25% & 75% percentile) of n=63 TKA wear simulation studies according to ISO 14243-1 & ISO 14243-3

5. Search Strategy

Manual search was performed focused on ISO and ASTM standards related to biotribology by some expert delegates, being part of the responsible ISO/TC 150 SC1, SC4 & SC5 technical committee's and the ASTM Committee F04 "Medical & Surgical Materials & Devices".

By the expertise given in the group additional publications dealing on ISO & ASTM standard wear testing and aspects of their clinical validation were selected.

Potentially eligible articles were screened in detail for applicability and 61 articles were included in the references.

6. Rationale

Can standard test methods in total hip and knee arthroplasty (THA, TKA) show that the planned articulations enable the function of the joint replacement throughout the expected implant lifetime without producing a critical amount of wear?

Standards are helpful and useful for industry, researchers, healthcare providers, health regulators, competent authorities, notified bodies, litigation bodies, the legal systems, the patients and the wider population and society. Standards are becoming increasingly important, not just in regulation and approval for use but also in litigation, health and safety product liability and consumer protection laws. There is an expectation that international standards are used by many industry sectors and many parts of society. Expert guidance and advice on use of existing standards and development of new standards is critical.

How specific standards are used by manufacturers, industry and regulators will depend on the individual implant design approach, the types of individual implant designs. The rational and how standards are used should form part of the implant design dossier and portfolio submitted to the regulator and should inform the instructions and indications for clinical use and patient selection.

Standards for pre-clinical testing particularly wear and biotribology are particularly useful in comparing performance and function of new implants (under development) with existing substantially equivalent implants with clinical history. Appropriate clinically established devices should be chosen with similar articulation materials and same basis of similarity in design (i.e. cruciate retaining, sacrificing or posterior stabilized TKA implants), with a clinical history recorded through registers or trials should be used to demonstrate equivalence or improved performance of the new device to the reference device. A clinical history of between 5 and 10 years might be expected of a reference device (substantially equivalent), unless the new device is so radically different to any existing devices in clinical use. The use of pre-clinical testing standards in new product developments can substantially shorten the product development (design loops) and evaluation time line, reduce the dependency and duration on having to undertake clinical studies or generate clinical evidence prior to regulatory approval. They do not however reduce the value or need of post market surveillance.

Many standard tends to evaluate performance or function under one standard or average (most commonly found) set of conditions. Under that standard set of conditions the pre-clinical standard test or simulation might be expected to predict the function of the average clinical implant (implant function in the average or most common conditions patients and surgeons). *The pre-clinical standard test does not predict the variation found in clinical implant functions associated with variation in clinical, patient or surgical conditions.* The variation in function found in standard tests does not reflect the variation in function found in the wider clinical population. Variation in function found in the clinical population is frequently associated with variation in combinations of multiple variables and conditions. Variables that influence clinical performance and the variation in clinical function include:

- Patient activity (level and type)
- Patient anatomy and gait and disease state
- Implant size
- Surgical positioning (rotation and translation) and relative alignment
- Soft tissue reconstruction
- Changes in soft tissue function post operatively.
- Imaging (weight bearing /non weight bearing) and changes in functional positioning post operatively and post rehabilitation
- Material ageing and degradation over time
- Changes in bone geometry and properties over time
- Pseudo synovial fluid in the joint space, volume and composition

As well as combinations of all of above.

Standard tests cannot currently accommodate the variation in these variables, or combinations in the variation of these variables. Therefore average values need to be set in standard tests (i.e. medium size TKA implants, level walking activity, duration 5 million cycles, well aligned THA & TKA components). Where a specific average value cannot be agreed in the standard then a range for the input variable should be set and a specific value chosen within the range by the manufacturer with which to compare the new product device with the substantially equivalent device. The rational for the choice of the input variable should be provided in the test report and design dossier.

1 Which standard test methods according to standards (ISO, ASTM) are required for registration of an implant?

While there may be concerns about some existing standards, the value of current and appropriate new standards in wear and biotribology cannot be underestimated. However, not all standards are of equal value and expert guidance and advice is crucially needed on the use and interpretation of today's international standards.

ISO 14242-1:2014 & AMD 1:2018 Implants for surgery — Wear of total hip-joint prostheses — Part 1: Loading and displacement parameters for wear-testing machines and corresponding environmental conditions for test. After two decades of widespread use of hip wear simulation, validation against retrievals [1,2,3,4,5], in vivo measured wear [6,7] and a broad variety of clinically long-term successful materials and designs [8,9,10,11], this standard is of high clinical relevance for all types of Ceramic-on-Polyethylene (CoP), Metal-on-Polythelene (MoP) and Oxidized-Zirkonium-on- Polyethylene (OxZroP) THA articulations. It might be not really clinically relevant to test a Ceramic-on-Ceramic (CoC) THA bearing under standard level walking conditions [12,13,14,15]. For CoC THA articulations wear simulation according to ISO 14242-4 shall be performed. Large diameter Metal-on-Metal (MoM) hip articulations have not been considered within the consensus statement, because it is much more complex to mimic their clinical behavior in wear simulation compared to CoP, MoP, OxZroP, CoC articulations.

ISO 14242-2:2016 Implants for surgery — Wear of total hip-joint prostheses — Part 2: Methods of measurement. This standard describes the methodology of wear measurement, for THA wear simulation acc. to ISO 14242 Part 1.

ISO 14242-3:2009 & AMD 1:2019 Implants for surgery — Wear of total hip-joint prostheses — Part 3: Loading and displacement parameters for orbital bearing type wear testing machines and corresponding environmental conditions for test. Orbital bearing type wear simulation may be used for comparison to historical data, but is not fitting for the purpose to mimic in vivo wear behaviour and should be withdrawn.

ISO 14242-4:2018 Implants for surgery — Wear of total hip-joint prostheses — Part 4: Testing hip prostheses under variations in component positioning which results in direct edge loading. This standard is of high clinical relevance for all types of hard-on-hard THA bearings [16,17,39] and considers a wider and more realistic set of clinical variables and conditions. For Ceramic-on-Ceramic (CoC) THA articulations it is mandatory to simulate wear according to ISO 14242-4:2018 to include more severe in vivo conditions such as edge loading, micro-separation etc. [12,13,14,15,38,39]. The clinical validation and clinical evidence to test a Ceramic-on-Polyethylene (CoP) or Metal-on-Polythelene (MoP) THA bearing under edge loading conditions is not fully complete. As such this test may be considered as a comparative test comparing the device under development to a predicate device with proven clinical history. Further advice on the use of this test in combination with artificial ageing protocols is given in section 6B.

ISO 14243-1:2009 & AMD 1:2020 Implants for surgery — Wear of total knee-joint prostheses — Part 1: Loading and displacement parameters for wear-testing machines with load control and corresponding environmental conditions for test. More than two decades ago in vitro wear simulation was introduced to assess the biotribological mechanisms of total knee replacements under level walking conditions [18,19,20,21] and on clinically proven knee implants of different design it has been demonstrated that articulation wear similar to in vivo wear modes can be generated [22]. Experimental wear studies were carried out to optimise implant designs, contact mechanics and articulation materials [23,24,25,26]. After two decades of widespread use of knee wear simulation under load control, validation against retrievals [22] and a broad variety of clinically long-term successful materials and designs [10,11,43,46], this standard is of high

clinical relevance for different types of knee replacements (UKA, TKA (CR, DD, UC, MS, PS, PS+), PS Revision varus-valgus stabilised, Rotating hinge knees). Regarding the effect of knee joint laxity on wear and kinematics, Kretzer et al. [27] found that the relatively high linear motion restraint given in the previous version ISO 14243-1:2002 does not represent adequately the in vivo conditions. They proposed the use of an asymmetric non-linear ligament and soft tissue restraint model based and reported increased AP translation and IE rotation in good agreement with clinical findings [28]. This has been considered in the current version ISO 14243-1:2009 by a more realistic soft tissue restraint [29]. A limitation may be given that current in vitro knee wear testing is mainly focused on abrasive-adhesive surface wear based on level walking test conditions and does not reproduce fatigue and delamination wear an essential clinical failure mode [29,30].

ISO 14243-2:2016 Implants for surgery — Wear of total knee-joint prostheses — Part 2: Methods of measurement. This standard describes the methodology of wear measurement, for TKA, UKA wear simulation acc. to ISO 14243 Part 1 & Part 3.

ISO 14243-3:2014 & AMD 1:2020 Implants for surgery — Wear of total knee-joint prostheses — Part 3: Loading and displacement parameters for wear-testing machines with displacement control and corresponding environmental conditions for test. After two decades of extensively use of knee wear simulation under displacement control [20,21,24,30,31,32] and validation against a broad variety of clinically long-term successful implant designs [10,11,47,50] this standard is of high clinical relevance and allows for knee bearing material optimisations under defined test conditions (displacement control). It has also importance for suitable wear testing of some mobile bearing knee designs not showing a relevant internal-external rotational movement pattern under load control (Part 1) test conditions.

ISO 14243-5:2019 Implants for surgery — Wear of total knee prostheses — Part 5: Durability performance of the patellofemoral joint. The requirement is the completion of 50,000 cycles without failure, but it remains questionable if 50,000 cycles are enough to perform a realistic destructive testing in regard to the clinical relevance. The polyethylene patellar component shall be artificially aged ahead testing. Interpretation of the test results is somehow demanding as this is a destructive test and some authorities are not aware on this. In the future it might be useful to develop a wear simulation for the patella-femoral joint.

ISO 17853:2011 Wear of implant materials — Polymer and metal wear particles — Isolation and characterization. Can only be used to compare similar materials and does not predict or present evidence in relation to biological reactions. For new or substantially modified materials it is of limited value and needs to be revisited and updated.

ASTM F2003 Standard Practice for Accelerated Aging of Ultra-High Molecular Weight Polyethylene after Gamma Irradiaton in Air. The method has also shown to be validated for in vivo oxidation of CPE after gamma irradiation under nitrogen [7,33,34], but is not representative of shelf or in vivo ageing of highly cross-linked or vitamin E stabilized polyethylene (see Consensus Statements regarding Biotribology). For current polymers it is of limited value and needs to be revisited and updated.

ASTM F3047M Standard Guide for High Demand Hip Simulator Wear Testing of Hard-on-hard Articulations This standard describes the specific test procedures for metal/metal, ceramic/ceramic and coated bearing couplings not covered by routine wear testing such as ISO 14242-1. High inclination angles, third body particles, enhanced loading, stop-dwell-start and microseparation have been identified as critical test conditions for hard-on-hard articulations. Further literature is referenced for the respective damage mechanisms.

General Advice on the application and use of these standards

The following general advice can be given for the input conditions in the standard tests, and which can be applied across most of the standard tests.

- Clinically established reference devices, an existing approved device of similar nature (substantial equivalent) with minimum 5 and preferably 10 years of clinical history as found in registries or clinical studies.
- Replicates: to obtain a consistent result considering the surrounding conditions of implant wear testing, a minimum of n = 3 + 1 (loaded soak control) specimen should be tested. For new materials, thin THA inserts or thin knee gliding surfaces n = 6 specimen shall be tested.
- Positioning: in standard wear simulation one defined position shall be tested. Going forward standards for a wider range of implant positions will be needed, as exemplified by ISO 14242-4.
- Test duration: five million cycles with preference to increase duration to 10 million cycles for new materials or thin THA inserts or thin knee gliding surfaces (see ISO 21535 & ISO 21536).
- Frequency: preference one Hertz with a range of 0.8 to 1.2 Hz.
- Protein concentration: range of 25 % to 50% bovine serum as specified by manufacturer or test laboratory (concentrations above 50% should not be used).
- Activity: standard walking
- Size: One size as defined by manufacturer or test house and design rational. In THA a comparative worst case testing (i.e. largest intended diameter of a system) should be tested according to the ISO 14242-1 standard protocol. In TKA it is suitable to perform a medium size in comparative testing.
- Ageing: devices tested as prepared in packaging and shelf life. Going forward accelerated artificial aged components should also be incorporated into future standard tests (see Consensus Statement regarding Biotribology).
- Measurements/ outputs: wear, deformation (creep) and fatigue need to be determined as outputs and outcomes of standard tests

2 How meaningful are they for the actual clinical application?

In Hip and Knee Arthroplasty the biological response to wear particles released from polymer, metal and ceramic articulation bearing materials is a key factor in peri-prosthetic osteolysis and subsequent implant loosening [35,36,37,38,39,40,41,42,43]. In the mid- to long-term (5- to 25-years service in vivo) wear debris induced osteolysis and subsequent aseptic loosening is the main reason for revision [10,11,44,45,46,47,48,49,50].

Current ISO 14242 and ISO 14243 standard test series are applied over two decades [1,2,3,12] and are now in widespread use in numerous laboratories and have been validated for their prediction of clinically outcome [6,8,9,10,11,14,15,27,28,29,31,51,52,53,54,55, 56,57]. Pre-clinical testing results (wear volume and rate, head penetration, released particle sizes, shapes and distributions) have been compared to clinical retrievals, in vivo wear measurements, clinical studies and registry data by quite a number of research groups in Europe and worldwide. The authors are fully aware that a comparison of standard wear testing results between different laboratories, wear testing machine types and procedures, wear measurement and particle analysis protocols is limited. For the evaluation of a new device a comparison to a substantial equivalent clinical established reference device is necessary.

Having this in mind we reviewed and compiled performance data from n = 257 wear simulation studies from four different laboratories, including diverse types and generations of wear simulators (Leeds, Endolab, Prosim, AMTI) for this consensus statement. Based on the high number of completed wear studies (5 million cycles each, n = 3,4,5 or 6 specimen) the individual wear rates were plotted with median, 25% and 75% percentile for THA (Figure 1) and TKA (Figure 2). Including the majority of today's clinically long-term successful THA and TKA implant designs and bearing materials the following recommendations can be given. Considering the upper 75% percentile level for THA articulations with conventional polyethylene (CPE n = 112 studies) and highly cross-linked polyethylene (XLPE n = 53 studies), an approximated reasonable threshold of 30 mg/million cycles for CoCPE, of 40 mg/million cycles for MoCPE and of 5 mg/million cycles for CoXLPE as well as for MoXLPE can be recommended (Figure 1). In THA the thresholds have been based on performance data for diameter 28 & 32 bearings, but in view to further reduce the risk for osteolysis it might be suitable to use these thresholds also for diameter 36 and 40 articulations as an orientation.

For TKA CR and PS implant designs with CPE gliding surfaces a reasonable threshold is given at 25 mg/ million cycles, performing knee wear simulation according to **ISO 14243-1:2009** under load control (Figure 2) [58,59,60]. For other material combinations such as TKA with XLPE gliding surfaces there was not enough performance data (CR type n = 26 & PS type n = 1 studies) and clinical long-term experience to propose a reasonable threshold.

Measuring average wear rates in THA ISO **ISO 14242-1:2014** or TKA **ISO 14243-1:2009** wear testing above the recommended thresholds in the four laboratories from where the data was collected an optimisation in implant design, bearing material, manufacturing tolerances and/ or surface quality should be undergone.

Considering the current unsatisfactory reproducibility and repeatability of wear results between some laboratories, these thresholds cannot be applied to tests undertaken in other laboratories, which may due to differences in simulation methods, use of test equipment or less staff experience in biotribology result in systematic differences in the simulated wear rates. In other laboratories, comparison to data from a substantial equivalent device with clinical history carried out in the same laboratory is needed for comparison. Today an additional concern is that a single laboratory may set or generate conditions which may produce artificially low wear rates [23,61], as such the comparison of the product under development to a substantial equivalent device with clinical history is not as strong and sensitive. In such a case these wear rates of a substantial equivalent device with clinical history have to be validated against average clinical wear rates or have to be validated with comparison to other laboratories (i.e. round robin tests etc.).

7. References

1 Barbour PSM, Stone MH, Fisher J: **A hip joint simulator study using simplified loading and motion cycles generating physiological wear paths and rates. Proc Instn Mech Engrs Part H 1999;214:455-67.**

2 Barbour PSM, Stone MH, Fisher J. **A hip joint simulator study using new and physiologically scratched femoral heads with ultra-high molecular weight polyethylene acetabular cups. Proc Instn Mech Engrs, Part H 2000;214:569-76.**

3 Kaddick C, Wimmer MA: **Hip simulator wear testing according to the newly introduced standard ISO 14242. Proceedings of the Institution of Mechanical Engineers. Part H, Journal of engineering in medicine 2001;215(5):429-42.**

4 Puente Reyna AL, Jäger M, Floerkemeier T, Frecher S, Delank KS, Schilling C, Grupp TM: **Backside wear analysis of retrieved acetabular liners with a pressfit locking mechanism in comparison to wear simulation in vitro. BioMed Res Int 2016; Article ID 8687131.**

5 Puente Reyna AL, Holderied M, Jäger M, Schilling C, Grupp TM: **Articulation and backside wear analysis after long-term in vitro wear simulation of vitamin E stabilized polyethylene acetabular liners with a press-fit locking mechanism. Traumatology & Orthopedics of Russia 2018;24(2):29-40.**

6 Busch A, Jaeger M, Klebingat S, Baghdadi J, Floerkemeier T, Huetter F, Grupp TM, Haversath M: **Vitamin E blended highly cross-linked polyethylene liners in total hip arthroplasty: a randomised, multicenter trial using virtual CAD-based wear analysis at 5-year follow-up. Archives of Orthopaedic and Trauma Surgery. 2020;140:1859-66.**

7 Grupp TM, Holderied M, Mulliez MA, Streller R, Jäger M, Blömer W, Utzschneider S: **Biotribology of a vitamin E stabilised polyethylene for hip arthroplasty – influence of artificial ageing and third-body particles on wear. Acta Biomaterialia. 2014;10:3068-78.**

8 Galvin AL, Tipper JL, Jennings LM, Stone MH, Jin ZM, Ingham E, Fisher J. **Wear and biological activity of highly crosslinked polyethylene in the hip under low serum protein concentrations. Proc IMechE Part H: J Engineeering in Medicine Part H, 2007;221:1-10.**

9 Wang A, Essner A, Cooper J: **The Clinical Relevance of Hip Simulator Testing of High Performance Implants. Seminars in Arthroplasty 2006;17(2):49-55.**

10 Reed M, Howard P, Brittain R, Lawrence S, Stonadge J, Wilkinson M, Wilton T: **National Joint Registry England, Wales, Northern Ireland & Isle of Man. 16th Annual Report 2019.**

11 Graves S, de Steiger R, Lewis P, Harris I: **Australian Orthopaedic Association National Joint Replacement Registry for Hip, Knee & Shoulder Arthroplasty. Annual Report 2019.**

12 Nevelos JE, Ingham E, Doyle C, Nevelos AB, Fisher J: **Wear of HIPed and non-HIPed alumina-alumina hip joints under standard and severe simulator testing conditions. Biomaterials 2001;22:2191-7.**

13 Nevelos J, Ingham E, Doyle C, Streicher R, Nevelos A, Walter W, Fisher J: **Microseparation of the centers of alumina-alumna artificial hip joints during simulator testing produces clinically relevant wear and patterns. J Arthroplasty 2000;15:793-5.**

14 Al-Hajjar M, Leslie IJ, Tipper J, Williams S, Fisher J, Jennings LM. **Effect of cup inclination angle during microseparation and rim loading on the wear of BIOLOX® delta ceramic-on-ceramic total hip replacement. J Biomed Mater Res B Appl Biomater 2010;95:263-8.**

15 Al-Hajjar M, Fisher J, Tipper JL, Williams S, Jennings LM. **Wear of 36-mm BIOLOX(R) delta ceramic-on-ceramic bearing in total hip replacements under edge loading conditions. Proc Inst Mech Eng H 2013;227:535-42.**

16 Williams S, Al-Hajjar M, Issac GH, Fisher J: **Comparison of ceramic-on-metal and metal-on-metal hip replacements under adverse conditions. Journal of Biomedical Materials Research Part B: Applied Biomaterials 2013;101B:770–5.**

17 Leslie IJ, Williams S, Isaac G, Ingham E, Fisher J: **High Cup Angle and Microseparation Increase the Wear of Hip Surface Replacements. Clin Orthop Relat Res 2009;467(9):2259-65.**

18 Walker PS, Blunn GW, Broome DR, Perry J, Watkins A, Sathasivam S, Dewar ME, Paul JP: **A knee simulating machine for performance evaluation of total knee replacements. J Biomechanics 1997;30:83-9.**

19 DesJardins JD, Walker PS, Haider H, Perry J: **The use of a force-controlled dynamic knee simulator to quantify the mechanical performance of total knee replacement designs during functional activity. J Biomechanics 2000;33:1231-42.**

20 McEwen HMJ, Fisher J, Goldsmith AAJ, Auger DD, Hardaker C, Stone MH: **Wear of fixed bearing and rotating platform mobile bearing knees subjected to high levels of internal and external tibial rotation. J Materials Science: Materials in Medicine 2001;12:1049-52.**

21 Barnett PI, Fisher J, Auger DD, Stone MH, Ingham E: **Comparison of wear in a total knee replacement under different kinematic conditions. J Materials Science: Materials in Medicine 2001;12:1039-42.**

22 Walker PS, Blunn GW, Perry JP, Bell CJ, Sathasivam S, Andriacchi T, Paul JP, Haider H, Campbell PA: **Methodology for long-term wear testing of total knee replacements. Clin. Orthop. Rel. Res. 2000;372:290-301.**

23 Muratoglu OK, Bragdon CR, Jasty M, O´Connor DO, von Knoch RS, Harris WH: **Knee simulator testing of conventional and cross-linked polyethylene tibial inserts. J Arthroplasty 2004;19:887-97.**

24 McEwen HMJ, Barnett PI, Bell CJ, Farrar R, Auger DD, Stone MH, Fisher J: **The influence of design, materials and kinematics on the in vitro wear of total knee replacements. J Biomechanics 2005;38:357-65.**

25 Grupp TM, Stulberg DS, Kaddick C, Maas A, Fritz B, Schwiesau J, Blömer W: **Fixed bearing knee congruency – influence on contact mechanics, abrasive wear and kinematics. Int J Artificial Organs 2009;32:213-23.**

26 Grupp TM, Kaddick C, Schwiesau J, Maas A, Stulberg SD: **Fixed and mobile bearing total knee arthroplasty – Influence on wear generation, corresponding wear areas, knee kinematics and particle composition. Clinical Biomechanics 2009;24:210-7.**

27 Kretzer JP, Jakubowitz E, Sonntag R, Hofmann K, Heisel C, Thomsen M: **Effect of joint laxity on polyethylene wear in total knee replacement. J Biomechanics 2010;43:1092-6.**

28 DesJardins JD, Banks SA, Benson LC, Pace T, LaBerge M: **A direct comparison of patient and force-controlled simulator total knee replacement kinematics. J Biomechanics 2007;40:3458-66.**

29 Johnston H, Abdelgaied A, Pandit H, Fisher J, Jennings LM. **Representing the Effect of Variation in Soft Tissue Constraints in Experimental Simulation of Total Knee Replacements. J Mechanical Behavior of Biomedical Materials 2018;87:87-94.**

30 Galvin A, Jennings LM, McEwen HM, Fisher J: **The influence of tibial tray design on the wear of fixed-bearing total knee replacements. Proc. IMechE Part H: J. Engineering in Medicine 2008;222:1289-93.**

31 Brockett CL, Jennings LM, Hardaker C, Fisher J: **Wear of moderately cross-linked polyethylene in fixed bearing total knee replacement. IMechE H 2012;226:529-35.**

32 Brockett C, Abdelgaied A, Haythornthwaite T, Hardaker C, Fisher J, Jennings LM: **The influence of simulator input conditions on the wear of total knee replacements: An experimental and computational study. Proceedings of the Institution of Mechanical Engineers, Part H: J Engineering in Medicine, 2016;230:429-39.**

33 Schwiesau J, Schilling C, Utzschneider S, Jansson V, Fritz B, Blömer W, Grupp TM: **Knee wear simulation under conditions of highly demanding daily activities - Influence on an unicompartmental fixed bearing knee design. Medical Engineering & Physics. 2013;35:1204-11.**

34 Grupp TM, Fritz B, Kutzner I, Schilling C, Bergmann G, Schwiesau J: **Vitamin E stabilised Polyethylene for Total Knee Arthroplasty evaluated under Highly Demanding Activities Wear Simulation. Acta Biomaterialia 2017;48:415-22**.

35 Amstutz HC, Campbell P, Kossovsky N, Clarke IC: **Mechanism and clinical significance of wear debris induced osteolysis. Clin. Orthop. Rel. Res. 1992;276:7-18.**

36 Algan SM, Purdon M, Horowitz SM: **Role of tumor necrosis factor alpha in particulate-induced bone resorption. J Orthop Res 1996;14(1):30-5.**

37 Revell PA, Al-Saffar N, Kobayashi A: **Biological reaction to debris in relation to joint prostheses. Proceedings of the Institution of Mechanical Engineers. Part H, J Engineering in Medicine 1997;211:187-97.**

38 Green TR, Fisher J, Stone M, Wroblewski BM, Ingham E: **Polyethylene particles of a `critical size` are necessary for the induction of cytokines by macrophages in-vitro. Biomaterials 1998;19:2297-302.**

39 Ingham E, Fisher J: **Biological reactions to wear debris in total joint replacement. Proceedings of the Institution of Mechanical Engineers. Part H, J Engineering in Medicine 2000;214:21-37.**

40 Tipper JL, Hatton A, Nevelos JE, Ingham E, Doyle C, Streicher R, Nevelos AB, Fisher J: **Alumina–alumina artificial hip joints. Part II: Characterisation of the wear debris from in vitro hip joint simulations.Biomaterials 2002;23:3441-8.**

41 Bertrand J, Delfosse D, Mai V, Awiszus F, Harnisch K, Lohmann CH: **Ceramic prosthesis surfaces induce an inflammatory cell response and fibrotic tissue changes. Bone Joint J 2018;100B:882-90.**

42 Van der Straeten C: **Hip resurfacing arthroplasty in young patients: international high-volume centres' report on the outcome of 11,382 metal-on-metal hip resurfacing arthroplasties in patients <50 years at surgery. Hip International 2020;1-10.**

43 Seppänen M, Laaksonen I, Pulkkinen P, et al.: **High revision rate for large-head metal-on-metal THA at a mean of 7.1 years: a registry study. Clin Orthop Relat Res 2018;476:1223–30.**

44 Sadoghi P, Liebensteiner M, Agreiter M, Leithner A, Böhler N, Labek G: **Revision surgery after total joint arthroplasty: a complication-based analysis using worldwide arthroplasty registers. J Arthroplasty. 2013;28:1329-1332.**

45 Sharkey PF, Lichstein PM, Shen C, Tokarski AT, Parvizi J: **Why are total knee arthroplasties failing today? – Has anything changed after 10 years? J Arthroplasty. 2014;29:1774-1778.**

46 Paxton EW, Inacio MCS, Namba RS, Love R, Kurtz SM: **Metal-on-conventional polyethylene total hip arthroplasty bearing surfaces have a higher risk of revision than metal-on-highly crosslinked polyethylene: Results from a US registry. Clin Orthop Relat Res. 2015;473:1011-1021.**

47 Paxton EW, Inacio MCS, Kurtz SM, Love R, Cafri G, Namba RS: **Is there a difference in total knee arthroplasty risk of revision in highly crosslinked versus conventional polyethylene? Clin Orthop Relat Res. 2015;473:999-1008.**

48 Kurtz SM, Gawel HA, Patel JD: **History and systematic review of wear and osteolysis outcomes for first-generation highly crosslinked polyethylene. Clin Orthop Relat Res. 2011;469:2262-77.**

49 Dyrhovden GS, Lygre SHL, Badawy M, Gothesen O, Furnes O: **Have the causes of revision for total and unicompartmental knee arthroplasties changed during the past two decades? Clin Orth Rel Res 2017;475(7):1874-87.**

50 **Regional Registry of Orthopaedic Prosthetic Implantology. Hip, Knee & Shoulder Arthroplasty in Emilia Romagna Region, Italy**. Annual Report 2018.

51 Nebergall AK, Greene ME, Rubash H, Malchau H, Troelsen A, Rolfson O: **Thirteen-Year Evaluation of Highly Cross-Linked Polyethylene Articulating With Either 28-mm or 36-mm Femoral Heads Using Radiostereometric Analysis and Computerized Tomography. J Arthroplasty 2016;31:S269-76.**

52 Goyal P, Howard JL, Yuan X, Teeter MG, Lanting BA: **Effect of Acetabular Position on Polyethylene Liner Wear Measured Using Simultaneous Biplanar Acquisition. J Arthroplasty 2017;32:1670-4.**

53 Haw JG, Battenberg AK, Huang DCT, Schmalzried TP: **Wear Rates of Larger-Diameter Cross-Linked Polyethylene at 5 to 13 Years: Does Liner Thickness or Component Position Matter? J Arthroplasty 2017;32:1381-6.**

54 Bookman JS, Kaye ID, Chen KK, Jaffe FF, Schwarzkopf R: 1**0-Year Follow-Up Wear Analysis of Marathon Highly Cross-Linked Polyethylene in Primary Total Hip Arthroplasty. J Arthroplasty 2017;32:2587-9.**

55 O'Dwyer Lancaster-Jones O, Williams S, Jennings LM, Thompson J, Isaac GH, Fisher J, Al-Hajjar M: **An in vitro simulation model to assess the severity of edge loading and wear, due to variations in component positioning in hip joint replacements. J Biomedical Materials Research - Part B Applied Biomaterials 2018;106(5):1897-906.**

56 Al-Hajjar M, Jennings LM, Begand S, Oberbach T, Delfosse D, Fisher J: **Wear of novel ceramic-on-ceramic bearings under adverse and clinically relevant hip simulator conditions. J Biomedical Materials Research - Part B Applied Biomaterials 2013;101(8):1456-62.**

57 Okazaki Y, Hosoba M, Miura S, Mochizuki T: **Effects of knee simulator control method and radiation dose on UHMWPE wear rate, and relationship between wear rate and clinical revision rate in National Joint Registry. J Mechanical Behavior of Biomedical Materials 2019;90:182–90.**

58 Grupp TM, Saleh KJ, Mihalko WM, Hintner M, Fritz B, Schilling C, Schwiesau J, Kaddick C: **Effect of anterior-posterior and internal-external motion restraint during knee wear simulation on a posterior stabilised knee design. Journal of Biomechanics. 2013;46:491-7.**

59 Grupp TM, Schröder C, Kim TK, Miehlke RK, Fritz B, Jansson V, Utzschneider S: **Biotribology of a mobile bearing posterior stabilised knee design - Effect of motion restraint on wear, tibio-femoral kinematics and particles. Journal of Biomechanics 2014;47:2415-23.**

60 Haider H, Weisenburger JN, Kurtz SM, Rimnac CM, Freedman J, Schroeder DW, Garvin KL: **Does vitamin E-stabilized ultrahigh-molecular-weight polyethylene address the concerns of cross-linked polyethylene in total knee arthroplasty? J Arthroplasty. 2012;27:1-9.**

61 Bragdon CR, Jasty M, Muratoglu OK, O'Connor DO, Harris WH: **Third-body wear of highly cross-linked polyethylene in a hip simulator. J Arthroplasty 2003;18(5):553-561**.

RESEARCH TOPIC 8

Biotribological Methodologies in Wear Simulation, Evaluation of Wear Debris Release and biological Response in Hip and Knee Arthroplasty – Review and Guidance on Additional Methodologies to Complement Standard Testing

Authors

Thomas M. Grupp[1,2], Christian Kaddick[3], Claude Rieker[4], Jan Philippe Kretzer[5], John Fisher[6]

[1]Aesculap AG Research & Development, Tuttlingen, Germany
[2]Ludwig Maximilians University Munich, Department of Orthopaedic and Trauma Surgery, Musculoskeletal University Center Munich (MUM), Campus Grosshadern, Munich, Germany
[3]EndoLab Mechanical Engineering GmbH, Riedering, Germany
[4]Zimmer Biomet EMEA, Scientific Affairs, Zug, Switzerland
[5]Laboratory of Biomechanics & Implant Research, Clinic for Orthopedics & Trauma Surgery, Heidelberg University Hospital, Germany
[6]Institute of Medical & Biological Engineering, School of Mechanical Engineering, University of Leeds, United Kingdom

1. Abstract

Question: Biotribology

What kind of proof apart from traditional standard test methods (each for TKA and THA) can be applied to show that the planned articulations (i.e. also including patella-trochlea) enable the function of the implant / the joint throughout the expected implant lifetime without producing a critical amount of wear?

1. Which kind of additional/ new test methods exist to complement the standard tests in a useful way? How meaningful are they to further close the gap between pre-clinical assessment and actual clinical application?
2. Which additional methods dedicated for research and not for routine use have been described?
3. What are methods that simulate the in vivo conditions for wear testing of implants in order to gain information about biological impact on biotribological material properties and wear debris release?

Summary / Recommendation:

Additional biotribological methodologies are described, which are of importance to complement the given ISO and ASTM standard testing in the context of their clinical relevance (Question 1). Based on literature review and expertise on biotribological testing methodologies, engineering experiences and history of clinical performance guidance notes are provided as recommendations as to how they should be used and under which conditions they should be applied.

To answer question 2 additional research methodologies in biotribology are discussed based on suitable examples detected within the systematic review and wider published literature, which are of importance for tribological research but not intended for routine use.

The average local and systemic biological response in the patient population to wear debris released from today's long-term clinically established CoC, CoP and MoP articulations in THA and TKA/ UKA are reasonably well understood. The variation in the biological response found in the patient population is less well understood or documented. For implant material modifications, which are very similar to a clinically established material (i.e. Vitamin E stabilization of XLPE), then comparative volume wear from simulators and comparative particle analysis (particle size, size distribution, shape) may be used to demonstrate substantial equivalence (Question 3). For substantial material changes or new bearing materials, particle analysis and distributions cannot predict whole body system response or the system level intensity of adverse reactions in vivo. The body system response to debris from a new type of material cannot be easily determined pre-clinically and requires prolonged clinical studies to determine safety and efficacy.

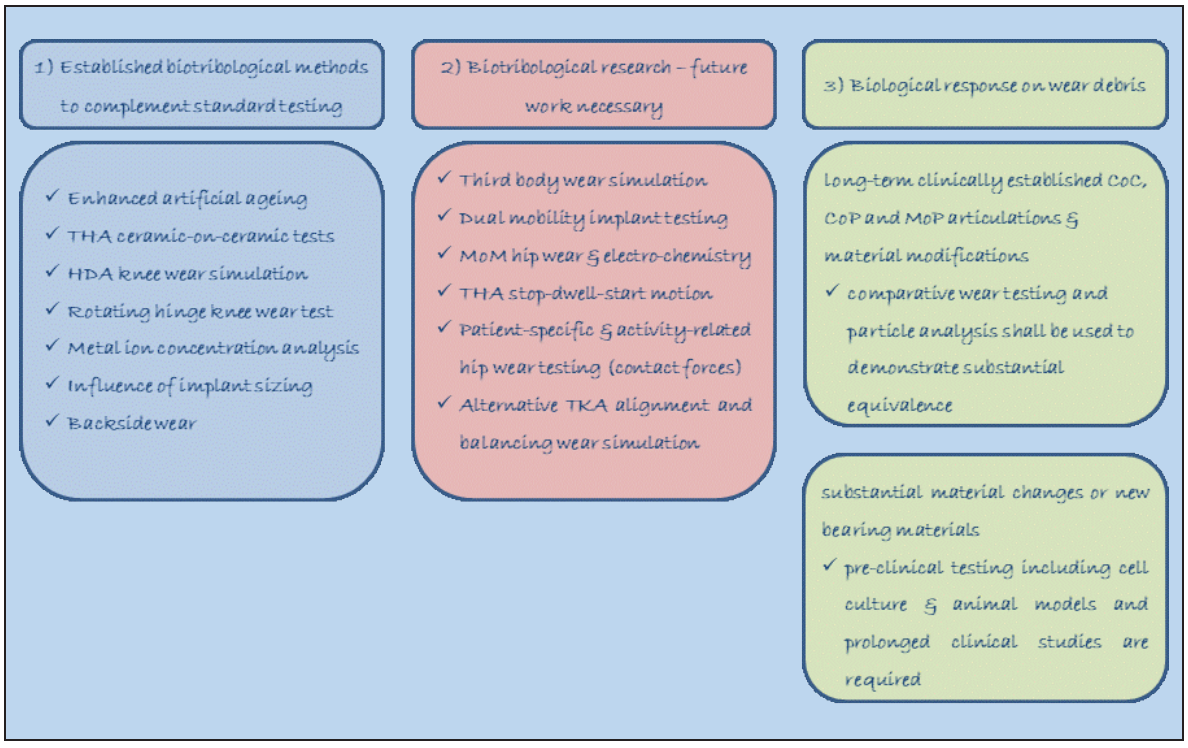
2. Level of Evidence

Strong

3. Consensus Delegate Vote

95% – unanimous, strongest consensus (95% agree / 0% disagree / 5% abstain)

4. Graphical Abstract



5. Search Strategy

A systematic literature search was conducted in March 2021 to identify literature with the potential to contribute to the answer of the above mentioned research question.

The electronic databases EMBASE incl. Medline were explored using a combination of the following search-terms which were combined with the Boolean operator "AND":

- TKA/THA, knee/hip prosthesis, knee/hip replacement
- implant testing/simulation, in vitro testing/simulation, wear testing/simulation, articulation testing/simulation, material testing/simulation, hip/knee simulator
- (biotribolog*, tribolog*, tribocorrosion, articulation) OR (wear debris, wear particle, wear behaviour, wear resistance, third body wear, hip wear, knee wear, prosthetic wear, patella wear, trochlea wear, surface wear, delamination wear, implant wear, wear simulation, particulate debris) OR (ion release, ion measurement, ion analysis, biologic* response, inflammatory response)

Search words were adapted to the specific terminology of the database and were searched including different spelling and synonyms.

Article language was restricted to English or German language. Because of the aim of the process studies from 2000 to 2021 were included to provide insight into current practice and state of the art procedures. Abstracts from scientific meetings were not included in the review as sufficient data cannot be extracted. Case reports were also excluded due to the nature of the contained information.

The title and abstract of all articles (n = 164) were screened to select articles contributing to the topic. Potentially eligible articles were obtained for detailed assessment and were screened for applicability in detail. References of used publications were searched for additional studies meeting the inclusion criteria. The selection was not further restricted regarding basic study design. Duplicate articles were removed.

164 publications met the search criteria. After the selection process, 53 articles were included in the review.

An additional manual search yielded further 77 publications which were included in the literature review after verification of eligibility based on topic and quality. Additional relevant literature was also identified and included (Figure 1).

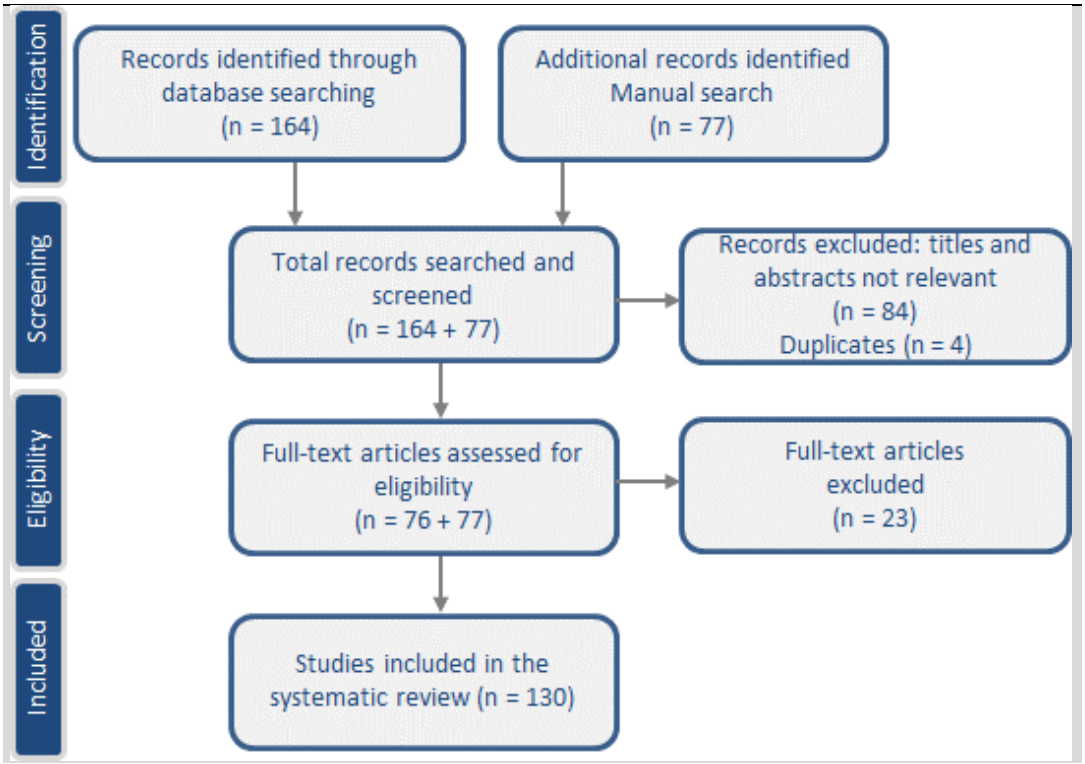


Figure 1: Flow chart literature selection process

6. Rationale

1 Which kind of additional/ new test methods exist to complement the standard tests in a useful way? How meaningful are they to further close the gap between pre-clinical assessment and actual clinical application?

In this section additional biotribological methodologies are described, which are of importance to complement the given ISO and ASTM standard testing currently used in the context of their clinical relevance. Based on literature review and expertise on biotribological testing methodologies and clinical experiences guidance notes are made how they should be used and under which conditions they should be applied.

To evaluate the long-term behaviour and possibly up-coming clinical failure mechanism of highly cross-linked polyethylene (XLPE), polyethylene stabilised by antioxidants (i.e. Vitamin E, AOX) or substantial modifications of currently established polyethylene for hip and knee arthroplasty (THA & TKA), an enhanced artificial ageing has to be considered. It has been previously shown that prolonged artificial ageing of 4 to 6 weeks using a procedure described in ASTM F2003 leads also in polyethylene with higher oxidative resistance (i.e. XLPE re-melted) as conventional (CPE 25-40 kGy y N2), to a progressive increase in wear and substantial increased oxidation [1,2,3]. To determine the oxidation status of polyethylene inserts after artificial ageing slices shall be cut by a microtome and oxidation index measurement shall be performed by Fourier transform infrared spectroscopy. Based on retrieved XLPE re-melted tibia inserts it has been shown that the oxidation index after 5 to 6 weeks of artificial ageing corresponds well to 3 to 7 years in vivo [2,4]. Furthermore Muratoglu et al. [5] described a mechanism of combined cyclic loading and absorption of lipids, which can further reduce the oxidative stability of irradiated and re-melted XLPE in vivo. For non-oxidative stabilised re-melted XLPE a combination of 3 to 4 weeks artificial ageing and doping in a lipid emulsion has been proposed by Suhardi et al. [3]. Oral et al. [6] described an in vitro oxidation model for XLPE re-melted incorporating synovial fluid lipids. They combined diffusion of emulsified single and mixed lipids into XLPE re-melted and accelerated ageing (2 weeks) and found that mixed lipids doping was able to reproduce average and maximum oxidation values and products observed in vivo in longterm retrievals (96 to 190 months). The oxidation depth profile and the cross-link density was different. To predict the long-term in vivo performance of polyethylene materials, it is crucial to understand the material specific oxidative processes and to apply a clinically relevant enhanced artificial ageing protocol [6]. Due to the variation in materials and in vivo oxidation mechanism, it is obvious that the procedure cannot be generalised and the choice of an appropriate method in relation to the specific wear tested material is important. In case of new or substantially modified polyethylene materials for THA inlays or TKA gliding surfaces, a complete physical, chemical and mechanical material characterization in context to biotribological properties has to be considered [7,8,9,10]. It is worth noting that full validation of clinical relevance of the wear produced by these ageing method has yet to be established. Full validation could be achieved by comparative simulator wear testing of real time in the body ageing of retrieved prostheses with accelerated aged material of equivalent age. Validation can be made through comparison of wear rates, wear mechanisms and of wear debris produced. Until such validation is completed, then confidence levels predictions of aged wear rates based on changed physical/chemical/mechanical property changes remains low to medium.

THA ceramic-on-ceramic (CoC) articulations clinical retrievals show evidence of edge loading stripe wear on head and rim wear on the cup insert [11,12]. This is not replicated by standard simulator walking cycle tests according to ISO 14242-1, which show very low surface wear < 0.1 mm3/million cycles. Simulator edge loading tests replicate clinical edge loading and stripe wear and wear rates of about 1 mm3/million cycles with alumina-on-alumina ceramic (Biolox Forte) [13]. Edge loading tests demonstrate reduced wear with zirconia toughened alumina (ZTA) ceramic composite (Biolox Delta) compared Biolox forte under edge loading [14]. In THA ceramic-on-ceramic (CoC) articulations simulating subluxation as a result of low tissue constraint and/ or a translational mismatch due to variations in medio-lateral component positioning and cup inclination angles in hip wear simulation has been introduced, to examine the occurrence and magnitude of dynamic separation, edge loading severity and wear increase [15]. Any new ceramic composite materials or substantial design changes in CoC bearings should be tested and compared to an established clinical reference such as a ZTA ceramic composite (i.e. Biolox Delta) under edge loading conditions [14], as specified in ISO 14242- 4.

Increased metal wear debris from metal-on-metal (MoM) articulations in hip resurfacing arthroplasty or in THA with large metal heads (> 36 mm) has been found in some patients and can be the cause for adverse local tissue reactions, elevated systemic metal ion levels and related diseases [16,17,18]. Under standard simulator walking cycle tests ISO 14242-1 metal metal bearings show low wear typically < 1 mm3/million cycles, with larger heads showing less wear than smaller heads, under these ideal laboratory conditions. This low level of wear does not usually cause adverse reactions in patients, except for in the small number of patients who may be susceptible and pre disposed to metal ion hypersensitivity reactions.

However, these standard tests (ISO 14242-1) do not reflect the much wider range and much higher wear rates found clinically particularly in large diameter (> 36 mm) bearings. Clinical variations in angle of implantation, (inclination and version), combined with deficiency in restoring the offset (medial lateral and anterior posterior) causes edge loading and substantial deterioration in performance. In smaller diameter bearings this can lead to increases in wear, subluxation and also dislocation, whereas in larger diameter bearings dislocation rate is lower, but edge loading and increases in wear becomes much more severe. Retrievals of large diameter bearings above 36mm show wear rates with edge loading, commonly exceeding 10 mm3/million cycles, sometimes reaching as high as 50 mm^3/million cycles, which cause adverse tissue and systemic reactions and need for revision. Laboratory simulations with simulated edge loading with both increased inclinations angles and offset deficiency have shown wear rates as high as 10 mm^3/million cycles [19], ten times higher than found in standard tests. Even higher rates may occur with more severe edge loading conditions as specified in ISO 14242-4. If any future design or development of MoM bearings or derivatives such as surface engineered bearings [20] are to be considered, these would need to be tested under severe edge loading conditions as specified in ISO 14242- 4. Edge loading tests can reveal potential clinical failure of surface engineered bearings [21] which is not found in standard tests [19]. Future edge loading tests would need to be conducted comparatively with a metal on metal device with a long clinical history, an established reference. Given the clinical performance of metal on metal bearings are typically considered unsatisfactory, a substantial improvement in performance compared to the clinical predicate device would be needed to justify a clinical study of any new device. The clinical study would need tight controls over implant positioning, inclination version and restoration of appropriate offset. Long term clinical follow up of 5 to 10 years would be needed in such a clinical study.

Current ISO 14243 series knee wear testing is able to simulate abrasive-adhesive surface wear behaviour in vitro. Due to the applied level walking test conditions only, a major limitation is given that an essential clinical failure mode with subsurface degradation, delamination and structural material fatigue [22,23,24,25] is not reflected, unless additional artificial or real time ageing is also introduced into the test. The structural material fatigue of TKA gliding surfaces caused by oxidation-induced embrittlement, degradation and crack concentration below the articulating surface has been reported for standard γ-irradiation under inert atmosphere [21], for sequentially irradiated and

annealed XLPE [22,23] and also for re-melted XLPE [24]. As total knee arthroplasty is being increasingly performed on heavier, more active and younger patients with higher life-time expectancy [26], it appears desirable to test for degradation and delamination wear. It has been previously demonstrated performing a highly demanding activities (HDA) wear simulation including stair climbing, and high flexion activities like chair raising and knee bending in combination with clinical relevant artificial ageing is able to create in vitro subsurface delamination and structural material fatigue TKA failure modes [27,28,29,30,31,32,33,34]. In addition to that it has been shown that moderate activities of daily living using in vivo measured knee joint loading, produces the highest abrasive-adhesive wear for walking based on in vivo data [35]. Based on the given evidence it is recommended to perform a highly demanding activities knee wear simulation for polyethylene, which are sensitive to reduce their mechanical strength due to in vivo oxidation (CPE, XLPE). In comparison to a clinically established reference knee implant the gliding surface degradation and fatigue behaviour should be evaluated. A HDA wear testing procedure shall be also applied for substantial polyethylene material modifications or new articulation material combinations (i.e. surface coatings, polymers), which have not been clinically established. In addition to that HDA knee wear simulation shall be performed for posterior stabilised (PS) implant designs to evaluate surface wear and potential structural material fatigue at the post/cam interface. As quite the majority of today's PS knee designs does not engage their post/cam mechanism during level walking test conditions (ISO 14243 series) or in the swing phase only [36,37], it seems to be mandatory to test PS implant designs under HDA knee wear simulation in combination with artificial ageing [26,28,33].

Rotating hinge knee (RHK) prostheses are a viable clinical option for severe deformities, complex primary and revision cases, accompanied by unstable ligaments and they constrain anterior-posterior (AP) shear and adduction-abduction moments [38,39,40]. To simulate the absence of both cruciate and collateral ligaments in RHK wear testing a substantial reduction in AP motion restraint and internal-external rotation restraint or alternatively HDA wear testing shall be considered [41]. In relation to retrieved RHK implants, it has been observed that the main wear damage modes of flanges and axis bushings were comparable between retrievals and in vitro tested specimen [42].

During knee wear simulation metal ion and particle release originating from femur and tibia components has been described [43], which might be of clinical relevance for patients with metal ion hypersensitivity [44,45]. Additionally the implantation of a Cobalt chrome implant and release of metal ions, without release of particles may also have clinical relevance to patients who are hypersensitive and should be avoided if possible. For the biotribological evaluation of implant solutions with barrier function to prevent ion release (i.e. mono- or multilayer ceramic coatings) as well as for articulation material combinations, which have not been clinically established, wear simulation (ISO 14243 series or HDA) in combination with metal ion concentration analysis of the lubricant by inductively coupled plasma mass spectroscopy (acc. to ISO 17294-2) shall be performed [42,46,47]. However it is not known what are considered acceptable levels of ion release in simulator tests and how this compares with ion levels detected in the body, recognizing that in simulator tests ions are contained and collected, while in body, ions are distributed widely creating at approximately 1000 fold dilution. Unlike wear volume, there is not calibration or direct comparison able to be made between ion release in simulators and ion release in body.

In regard to the influence of sizing on wear simulation the wear performance of THA and TKA depends on the size of the bearing articulation in the hip and the implant size in the knee. In THA a larger diameter (i.e. 36 or 40 mm) CoP or MOP bearing produces more surface wear than a diameter 28 mm bearing, due to a greater sliding distance and a greater surface contact area. The hip diameter is a design choice by the surgeon and a comparative worst case testing (i.e. largest intended diameter of a system) should be tested according to the ISO 14242-1 standard protocol. In TKA the implant size depends on the anatomy of the patient it is suitable to perform a medium size comparative testing. A large knee size might result in increased abrasive-adhesive surface wear (greater sliding distance and surface contact area), but this can be covered by the comparison to a clinically established reference system with similar dimensions (design and size range) based on a design rationale not requiring additional testing. For a new design principle or design modifications substantially changing dynamic surface and sub-surface contact mechanics (i.e. by FEA simulation), additional testing of a small worst case size combined with ageing shall be considered in regard to fatigue wear.

In hip and knee arthroplasty the articulation surface is the main source of particulate debris. However backside wear produced at the interface between polyethylene inlay and acetabular cup has been associated with retro-acetabular osteolysis and aseptic cup loosening and previously detected on retrieved inlays [48,49,50]. Similarly a backside wear phenomenon has been observed for polyethylene gliding surfaces on fixed bearing tibial trays revised for osteolysis and aseptic loosening [51] with a rotational pattern of scratching and burnishing on the inferior polyethylene surface and the superior titanium tray surface. McEwen et al. [52] demonstrated that implants tested in knee simulator studies showed consistent backside wear pattern to retrieved implant components. For a press-fit cone locking mechanism of hip inlays in a titanium alloy cup Puente Reyna et al. [53,54] found a good correlation between retrievals and simulator tested inlays. They also determined that for this fixation principle most of the backside wear occurred during inlay insertion into the cup and even performing a 20 million cycle hip wear simulation backside wear does not increase. In conclusion it can be stated, that for known material combinations & fixation design principles backside hip or knee polyethylene wear is well integrated in todays in vitro wear simulation procedures [51,52,53].

The evidence level of these additional test methods and the confidence levels by which they may predict clinical performance and potential variation in clinical performance are lower than for the standard test methods described in Consensus Statement 7/Biotribology. This does not mean they should not be used in pre-clinical assessments, but it does mean that caution in needed in terms of the confidence level by which they are able to predict the average clinical performance. Or give an indication of the variability of clinical performance.

2 Which additional methods dedicated for research and not for routine use have been described?

In this section additional research methodologies in biotribology are described based on suitable examples detected within the systematic review, which maybe of importance for tribological research but not intended for routine use at the present time.

For new hip or knee insert fixation design concepts, other polymer inlay materials (non-polyethylene) or new material combinations a test method allowing to separately quantify articulation and backside wear in regard to particle volume and particle morphometry may be considered in the future [55].

Third body wear, caused by bone cement particles containing zirconium dioxide or barium sulphate as radiopaque or titanium/ hydroxyapatite particulate debris originated from coatings for cementless fixation, leads clinically to increased wear and accelerates particle-induced aseptic loosening [56,57,58,59,60]. The term third body wear has been used to describe two different mechanisms which increase wear. One mechanism by which the third body particles themselves generate increased wear, but are subsequently ejected from the contact with the wear debris. The second whereby the third body particles enter the contact and damage the hard counterface, The damage to the hard conterface, typically the femoral surface then increases the wear of the other surface (polyethylene). The amount of increase in wear depends on the relative motions and kinematics and can last for prolonged periods as the damage can last for prolonged periods. The relative importance of these two mechanisms is different in different situations and for different bearing types and designs. This makes the development and validation of a standard test extremely difficult, and indicates that such test (and their validation) may have to be implant and material specific. Third body wear mechanism have been simulated in vitro for different THA bearing combinations [2,56,61,62,63], for TKA [64], as well as in unicompartmental knee arthroplasty (UKA) [65,66]. Third body wear testing may be valid to understand clinical failure modes and eventually to analyse the resistance of diverse bearing material couplings against hard third body particulate debris [2,45,56]. ASTM is currently developing a standard for in vitro testing in which alumina particles of size 100µm and 300µm are proposed at a rate of 2mg every 0.5 million cycles [67]. However, the published testing conditions differ substantially in regard to the used particle sizes (< 30 µm to 640 µm), the applied particle concentrations and the method of contamination of the articulation [2,61,62,63,64]. Due to that no consistent methods to analyse the influence of third body wear are available today [31]. Extreme caution is needed in the use of such tests currently.

Hip instability and dislocation can be caused by suboptimal component position, impingement or soft tissue deficiency and is a common complication in primary and revision THA [68,69,70,71]. Dual mobility implants have demonstrated to significantly reduce dislocation rates in patients at risk for hip instability [67,68,69]. Modular Dual Mobility implants are usually incorporate a cobalt-chromium inlay anchored in situ with a taper connection in a titanium alloy acetabular cup [72,73,74]. Due to the cobalt-chromium/ titanium alloy material combination, mechanically assisted crevice corrosion can take place with release of cobalt and chromium debris and ions [70,71,75]. Kolz et al. [70] examined 12 retrieved modular dual mobility inlays with a mean time in vivo of 26 months by a visual (Goldberg criteria) and a quantitative dimensional analysis, and found in 2 (17%) severe and in 5 (42%) inlays moderate corrosion with an average maximum depth of 35.5 µm at the inlay taper interface. Due to that the cobalt and chromium particle and ion release of modular dual mobility inlay/ cup interfaces can be evaluated in a demanding hip test setup (5340 N, 10 million cycles) under a steep cup inclination of 55° [72] or alternatively in an adapted ISO 14242-4 hip wear simulation. Due to limited evidence (i.e. it is even not known if the movement/ micro-movement in the simulations represent realistic movements in patients) future research is necessary to create some clinical validation. This is an important focus of future research work. It is important to recognise the role of patient selection in clinical use and clinical evidence for dual mobility bearings. While some evidence of clinical performance has been obtained from elderly patients. It needs to be recognised that to use these bearings in a wider age range of patients is a new application and pre-clinical test methods need to be developed for these patients, and at this point in time there is no clinical evidence available to validate these methods for younger and more active patients. Clinical studies are needed in this younger age group of patients.

MoM hip articulations produce abrasive-adhesive wear and a surface degradation due to a combination of tribological and corrosion processes and their interactions [76,77,78]. To study these complex tribo-corrosion mechanism, a hip simulator with integrated in situ electro-chemical real-time measurements has been proposed [74,75,79]. For the interaction of wear and corrosion processes the sources for the release of metal ions were described as tribology driven depassivation of the MoM articulation surfaces and in addition corrosion of nano-sized wear particles [74,75,76]. In research driven studies the methodology allows to characterize the mass balance of ions relative to metal-protein complexes and released cobalt-chromium particles during tribo-corrosion in hip simulators [76]. The transfer of real-time electro-chemical evaluation of tribo-corrosion to the clinical situation has not been demonstrated so far.

Simulating average THA patient movement pattern, Hadley et al. [80] proposed a hip wear simulation with adverse stop-dwell-start motion activities, to determine the impact on wear of different bearing combinations. Testing the robustness to stop-dwell-start motion alternating with walking activities a significant increase in wear was observed for the MoM combination attributed to depletion of lubricant in the hip bearing during the dwell period, whereas the relative wear increase was lower for metal- and ceramic-on-polyethylene (MoP, CoP) combinations, and the CoC articulation was not sensitive to these adverse conditions [78]. This research shows that the increase in wear rate is dependent on the stop dwell start duty run cycle based on the relative length of the dwell time compared to the run time. The wear rate per step, increases if the dwell time increases and run time decreases. However in terms of the wear rate per day this may not be the case, as a patient with long dwell times and short run times would typically take fewer steps per day. Today there is no consensus about what the preferred duty cycle should be and no definition of the test duration. As of today this test cannot be applied as a routine pre-clinical test. If it is performed for research purpose, it needs to be used as a comparative test with a well selected reference device of similar material, with careful consideration and justification of the test conditions and duty cycle.

To represent better in vivo loads in THA patient groups, Lunn et al. [81] examined patient-specific and activity-related variation of hip contact forces via musculo-skeletal modeling, in a cohort of 132 THA patients undergoing motion capture analysis during activities of daily living. They observed systematic differences in hip contact forces between current ISO 14242 hip wear testing and their measurements of patient's daily activities. Consequently, they advocated for more clinically relevant and more demanding loading conditions in THA wear testing. In the current stage their research is based on patient activity measurements and musculo-skeletal modeling. No information is given, if the different pattern have any relevance in relation to wear generation and so far the clinical relevance has not been demonstrated

(i.e. by retrieval studies etc.). Future work is necessary how to adapt their patient specific hip contact force pattern for THA wear simulation including the clinical validation.

In TKA new alignment and knee balancing approaches as an alternative to the established mechanical axis alignment, like kinematic alignment, anatomic alignment [82,83], constitutional varus [84], dynamic coronal femoro-tibial mechanical angle [85] or functional phenotypes in coronal limb alignment [86] coming more in consideration and growing interest by knee arthroplasty surgeons.

The majority of current TKA systems is designed and tested for implantation perpendicular to the mechanical axis of the knee joint (180° ± 3°). To consider alternative alignment and knee balancing principles in wear simulation there is a lack of knowledge which impact these alignment techniques have on the biomechanical loading conditions and related wear behaviour of total knee implants. As an example Maag et al. [87] performed a wear testing in a six-degree-of-freedom knee joint simulator based on implant- and alignment-specific loading conditions estimated by a previously validated finite element analysis (FEA). Based on their approach they reported for a cruciate retaining fixed, mobile bearing and posterior stabilized knee design, no significant differences in wear rates between anatomic and mechanical alignment [85]. Further work is also needed to understand how different knee balancing principles interact with different variations in positioning that are inevitable in surgery. This is a complex systems problem, with many different combinations of different variables. Our understanding of these interactions is not currently complete. Caution is needed in making changes to individual variable in such complex systems, when the impact of the interactions with other variables is not known.

Taking into account quite a number of influencing factors like implant size combination, internal-external rotation of femur and tibia component, tibial slope and coronal plane alignment principle and their entire clinical relevant range, there is today a lack of data, knowledge and evidence how to perform an appropriate wear simulation to reflect the clinical conditions abroad of the mechanical axis alignment [88,89]. It is obvious that in the next years a significant progress in FEA based contact mechanics for various TKA alignment approaches [90,91,92,93], validated dynamic FEA knee kinematic models [86,87] and advanced musculo-skeletal modeling [94] is necessary to create suitable knee loading and kinematic input parameters for an experimental knee wear simulation dedicated to alternative alignment and balancing principles.

At the current state of research it is impossible to replace experimental wear testing by FEA simulation and it is impossible to predict total joint replacement wear by computer simulation models. But modern in silico methods have a future role to combine the analysis of articulation contact mechanics with musculo-skeletal alignment approaches to deliver reasonable input parameters for experimental wear testing methods (i.e. ligament constrain or loading & kinematic conditions influenced by knee alignment).

Future wear testing for the hip, knee and patella-femoral arthroplasty need to consider variation in positioning (6 DOF), variation in activity, variation in sizes, variation in age, and on biological activity and reactions. A strategic framework as to how to define the choice of these input variables, both as single variables and as combination of variables needs to be defined.

3 What are methods that simulate the in vivo conditions for wear testing of implants in order to gain information about biological impact on biotribological material properties and wear debris release?

Wear debris generated from different bearing materials in hip and knee arthroplasty from can range from 1nm to 100 µm in size, and often the lowest size detected reflects the sensitivity of the measurement system used for different materials. Characterisation is typically in 2D only (except for very few studies). Particle distributions are highly skewed towards the smallest detected size. Trying to determine absolute numbers of particles is not easy or a helpful way to quantify or compare debris. Particles can be more usefully characterised as volume distributions as a function of particle size. Biological reactions to wear particles are dependent on their volume dose, number, size, shape and material chemistry. The biological reactions cannot be determined or inferred from traditional bulk material biocompatibility standard tests.

The biological response to polyethylene wear particles was described as a key factor in inducing periprosthetic osteolysis and subsequent implant loosening [95,96,97]. This complex mechanism involves activated macrophages and inflammatory cytokine release depending on the amount, morphology, material and size of the wear particles [98,99,100]. Periprosthetic osteolysis is stimulated by the macrophages activity which is, in particular, dependent on the volume of particulate debris in the submicron size range [101,102,103,104,105,106,107,108].

The local and systemic biological response to wear debris released from today's long-term clinically established CoC, CoP and MoP articulations in THA and TKA/ UKA are reasonably well understood and validation of pre-clinical wear testing and clinical behaviour based on retrievals [11], particle isolation and characterisation from peri-prosthetic tissue [104,109,110] and in vivo obtained synovial fluid [111,112,113] is given.

Also alternative articulation materials with more than a decade of clinical experience, such as Oxinium, ZrN-multilayer coating or CFR-PEEK as flanges and axis bushings in a RHK, have demonstrated their suitability based on well documented clinical studies and registry data [114,115,116,117,118,119,120,121].

For implant material modifications, which are very similar to a clinically established material (i.e. Vit. E stabilization of XLPE), then comparative volume wear from simulators and comparative particle analysis (particle size, size distribution, shape) may be used to demonstrate substantial equivalence [41,103,105,122].

But for substantial material changes, changes in material chemistry or new type bearing materials, then particle size analysis cannot be used to predict adverse biological reactions. Particle analysis and distributions cannot predict whole body system response or the system level intensity of adverse reactions in vivo. While comparative cell screening tests for cytotoxicity, inflammatory cytokines released by macrophages and for adverse immunological lymphocyte cell reactions may be conducted, the body system response to debris from a new type of material cannot be easily determined pre-clinically and requires prolonged clinical studies to determine safety and efficacy.

There are several types of adverse reactions to wear particles [98,123]. These include cytotoxic reactions, inflammatory reactions such as release of osteolytic cytokines from macrophages and also adverse immunological reactions involving lymphocyte cells. However, the particle-

induced biological response is a complex mechanism which is so far not completely understood in detail [124]. Although macrophages are thought to play an important role in particle-induced inflammation, some other potential cells such as T-lymphocytes, mast cells, histiocytes, fibroblasts, plasma cells, neutrophils, endothelial cells, osteoblasts and osteoclasts, are also involved in this process [125,126].

Different intensities of different types of adverse reactions occur with different materials, so it is not possible to compare different types of reactions produced by different materials. Comparative Cell culture studies may be used to look at individual adverse reactions from particles of similar materials, such as two types of polyethylene, but this is only a comparison, not a prediction of body system response [98,108,109]. Cell cultures cannot reflect the complex cellular interactions in the particle-induced immune response [111]. Due to that different animal models have been established to evaluate the patho-mechanism of this process. Zysk et al. [127,128] introduced an animal model to examine the biological response against wear particles via leucocyte-endothel-cell interactions in the murine knee joint [110,111]. To induce osteolysis Neuerburg et al. [129] applied a high concentration of Vitamin E stabilized XLPE wear particles to provoke an inflammatory response in an established murine calvaria model and analysed morphometric bone osteolytic changes (µCT) and inflammatory markers. Using extracted wear particles from a 100 million cycles THA wear testing of vitamin E blended XLPE in comparison to CPE Popoola et al. [130] injected particle suspensions in knee capsules of white new Zealand rabbits and examined the local inflammatory response and potential systemic effects by histological analysis of liver, spleen & lymph nodes.Animal studies, while having an important role in research, are very difficult to interpret clinically, and care needs to be taken in how they are used in pre human clinical studies and in consideration of approvals for use in humans.

7. References

1 Haider H, Weisenburger JN, Kurtz SM, Rimnac CM, Freedman J, Schroeder DW, Garvin KL: **Does vitamin E-stabilized ultrahigh-molecular-weight polyethylene address the concerns of cross-linked polyethylene in total knee arthroplasty? J Arthroplasty. 2012;27:1-9**.

2 Grupp TM, Holderied M, Mulliez MA, Streller R, Jäger M, Blömer W, Utzschneider S: **Biotribology of a vitamin E stabilised polyethylene for hip arthroplasty – influence of artificial ageing and third-body particles on wear. Acta Biomaterialia. 2014;10:3068-78.**

3 Suhardi JV, Oral, Muratoglu OK: **The effects of oxidation on highly cross-linked UHMWPE's. Abstract 7th UHMWPE International Meeting Philadelphia, PA Oct. 22-23, 2015.**

4 Currier BH, van Citters DW, Currier JH, Carlson EM, Tibbo MA, Collier JP: **In vivo oxidation in retrieved highly crosslinked tibial inserts. Journal of Biomedical Materials Research Part B: Applied Biomaterials Vol. 2013;101B:441-8.**

5 Muratoglu OK, Wannomae KK, Rowell SL, Micheli BR, Malchau H: **Ex vivo stability loss of irradiated and melted ultra-high molecular weight polyethylene. J Bone Joint Surg 2010;92A:2809-16.**

6 Oral E, Fung M, Rowell SL, Muratoglu OK: **In-vitro oxidation model for UHMWPE incorporating synovial fluid lipids. J Orthopaedic Res. 2018;36:1833-9.**

7 Akagi M, Asano T, Clarke IC, Niiyama N, Kyomoto M, Nakamura T, Hamanishi C: **Wear and toughness of crosslinked polyethylene for total knee replacements – A study using a simulator and small-punch testing. J Orthopaedic Res. 2006;24:2021-2027.**

8 Turner A, Okubo Y, Teramura S, Niwa Y, Ibaraki K, Kawasaki T, Hamada D, Uetsuki K, Tomita N: **The antioxidant and non-antioxidant contributions of vitamin E in vitamin E blended ultra-high molecular weight polyethylene for total knee replacement. J Mechanical Behavior of Biomedical Materials 2014;31:21-30.**

9 Mulliez MA, Schilling C, Grupp TM: **Equivalent mechanical properties of X-ray and E-beam cross-linked vitamin E blended ultrahigh molecular weight polyethylene. Journal of Biomedical Materials Research Part B – Applied Biomaterials. 2020;1-10 DOI: 10.1002/jbm.b.34552.**

10 Mulliez MA, Schilling C, Grupp TM: **Influence of Irradiation Temperature on Oxidative and Network Properties of X-Ray Cross-linked Vitamin E Stabilized Polyethylene for the Hip Arthroplasty BioMed Research International – Section Biomaterials. 2020; Article ID 2568428.**

11 Nevelos JE, Ingham E, Doyle C, Nevelos AB, Fisher J: **Analysis of retrieved alumina ceramic components for Mittelmeier total hip prostheses. Biomaterials 1999;20:1833-40.**

12 Nevelos J, Ingham E, Doyle C, Streicher R, Nevelos A, Walter W, Fisher J: **Microseparation of the centers of alumina-alumna artificial hip joints during simulator testing produces clinically relevant wear and patterns. J Arthroplasty 2000;15:793-5.**

13 Stewart T, Tipper JL, Streicher R, Ingham E, Fisher J: **Long-term wear or HIPed alumina on alumina bearings for THR under microseparation conditions. J of Materials Science: Materials in Medicine 2001;12:1053-6.**

14 Stewart TD, Tipper JL, Insley G, Streicher RM, Ingham E, Fisher J: **Long-term wear of ceramic matrix composite materials for hip prostheses under severe swing phase microseparation. J Biomedical Materials Research, Applied Biomaterials, Part B 2003;66B:567-73.**

15 O'Dwyer Lancaster-Jones O, Williams S, Jennings LM, Thompson J, Isaac GH, Fisher J, Al-Hajjar M: **An in vitro simulation model to assess the severity of edge loading and wear, due to variations in component positioning in hip joint replacements. J Biomedical Materials Research - Part B Applied Biomaterials 2018;106(5):1897-906.**

16 Van der Straeten C: **Hip resurfacing arthroplasty in young patients: international high-volume centres' report on the outcome of 11,382 metal-on-metal hip resurfacing arthroplasties in patients <50 years at surgery. Hip International 2020;1-10.hhttps://doi.org/10.1177/11207000209573tthttps://doi.org/10.1177/1120700020957ps://doi.org/10.1177/1120700020957**

17 Seppänen M, Laaksonen I, Pulkkinen P, et al.: **High revision rate for large-head metal-on-metal THA at a mean of 7.1 years: a registry study. Clin Orthop Relat Res 2018;476:1223–30.**

18 Campbell P, Ebramzadeh E, Nelson S, et al.: **Histological features of pseudotumor-like tissues from metal-on-metal hips. Clin Orthop Relat Res 2010; 468: 2321–7.**

19 Leslie IJ, Williams S, Isaac G, Ingham E, Fisher J: **High Cup Angle and Microseparation Increase the Wear of Hip Surface Replacements. Clin Orthop Relat Res 2009;467(9):2259-65.**

20 Leslie IJ, Williams S, Brown C, Anderson J, Isaac G, Hatto P, Ingham E, Fisher J: **Surface engineering: A low wearing solution for metal-on-metal hip surface replacements. J Biomed Mater Res B Appl Biomater, 2009;90B:558-65:**

21 Leslie I, Williams S, Isaac G, Hatto P, Ingham E, Fisher J: **Wear of surface-engineered metal-on-metal bearings for hip prostheses under adverse conditions with the head loading on the rim of the cup. Proceedings of the Institution of Mechanical Engineers, Part H - J Engineering in Medicine 2013;227(4):345-9.**

22 Berry DJ, Currier BH, Mayor MB, Collier JP: **Gamma-irradiation sterilization in an inert environment – A partial solution. Clin Orthop Relat Res. 2012;470:1805-13.**

23 Kop AM, Dip PG, Pabbruwe MB, Keogh C, Swarts E: **Oxidation of second generation sequentially irradiated & annealed highly cross-linked X3 polyethylene tibial bearings. J Arthroplasty. 2015;30:1842-6.**

24 MacDonald DW, Higgs GB, Chen AF, Malkani AL, Mont MA, Kurtz SM: **Oxidation, damage mechanism, and reasons for revision of sequentially annealed highly crosslinked polyethylene in total knee arthroplasty. J Arthroplasty. 2018;33:1235-41.**

25 Diamond OJ, Howard L, Masri B: **Five cases of tibial post fracture in posterior stabilised total knee arthroplasty using Prolong highly cross-linked polyethylene. Knee 2018;5:657-62**

26 Carr AJ, Robertsson O, Graves S, Price AJ, Arden NK, Judge A, Beard DJ: **Knee Replacement. The Lancet. 2012;379:1331-40.**

27 Popoola OO, Yao JQ, Johnson TS, Blanchard CR: **Wear, delamination, and fatigue resistance of meltannealed highly crosslinked UHMWPE cruciate-retaining knee inserts under activities of daily living. J Orthopaedic Research 2010;28(9):1120-6.**

28 Schwiesau J, Schilling C, Utzschneider S, Jansson V, Fritz B, Blömer W, Grupp TM: **Knee wear simulation under conditions of highly demanding daily activities - Influence on an unicompartmental fixed bearing knee design. Medical Engineering & Physics. 2013;35:1204-11.**

29 Schwiesau J, Fritz B, Kutzner I, Bergmann G, Grupp TM: **CR TKA UHMWPE wear tested after artificial aging of the Vitamin E treated gliding component by simulating daily patient activities. BioMed Research International – Biomaterials. 2014; Article ID 567374.**

30 Battaglia S, Belvedere C, Abdel-Jaber S, Affatato S, d'Angeli V, Leardini A: **A new protocol from real joint motion data for wear simulation in total knee arthroplasty: Stair climbing. Medical Engineering & Physics. 2014;36:1605-10.**

31 Abdel-Jaber S, Belvedere C, Leadini A, Affatato S: **Wear simulation of total knee prostheses using load and kinematic waveforms from stair climbing. J Biomechanics. 2015;48:3830-6.**

32 Zietz C, Reinders J, Schwiesau J, Paulus A, Kretzer JP, Grupp TM, Utzschneider S, Bader R: **Experimental testing of knee replacements with UHMW-PE inserts: impact of severe wear test conditions. J Mater Science: Mater in Medicine. 2015;26:134-47.**

33 Abdel-Jaber S, Belvedere C, de Mattia JS, Leardini A, Affatato S: **A new protocol for wear testing of total knee prostheses from real joint kinematic data: Towards a scenario of realistic simulations of daily living activities. J Biomechanics. 2016;49:2925-31.**

34 Grupp TM, Fritz B, Kutzner I, Schilling C, Bergmann G, Schwiesau J: **Vitamin E stabilised Polyethylene for Total Knee Arthroplasty evaluated under Highly Demanding Activities Wear Simulation. Acta Biomaterialia 2017;48:415-22.**

35 Reinders J, Sonntag R, Vot L, Gibney C, Nowack M, Kretzer JP: **Wear testing of moderate activities of daily living using in vivo measured knee joint loading. PLoS ONE. 2015; 10:Article ID e123155.**

36 Fitzpatrick C.K., Clary C.W., Cyr A.J., Maletsky L.P., Rullkoetter P.J.: **Mechanics of post-cam engagement during simulated dynamic activity. Journal of Orthopaedic Research 2013;31(9):1438-46.**

37 Arnout N., Vanlommel L., Vanlommel J., Luyckx J.P., Labey L., Innocenti B., Victor J., Bellemans J.: **Post cam mechanics and tibiofemoral kinematics: a dynamic in vitro analysis of eight posterior-stabilized total knee designs. Knee surgery, sports traumatology, arthroscopy. 2015;23(11):3343-53.**

38 Hossain F, Patel S, Haddad FS: **Midterm assessment of causes and results of revision total knee arthroplasty. Clin. Orthop. Rel. Res. 2010;468(5):1221-8.**

39 Giurea A, Neuhaus H, Miehlke R, Schuh R, Lass R, Kubista B, Windhager R: **Early results of a new rotating hinge knee implant. BioMed Research International 2014; Article ID 948520.**

40 Boehler C, Kolbitsch P, Schuh R, Lass R, Kubista B, Giurea A: **Midterm results of a new rotating hinge knee implant: A 5-year follow-up BioMed Research International 2017; Article ID 7532745.**

41 Grupp TM, Giurea A, Miehlke RK, Hintner M, Gaisser M, Schilling C, Schwiesau J, Kaddick C: **Biotribology of a new bearing material combination in a rotating hinge knee articulation. Acta Biomaterialia 2013;9:7054-63.**

42 Schierjott RA, Giurea A, Neuhaus HJ, Schwiesau J, Pfaff A, Utzschneider S, Tozzi G, Grupp TM: **Analysis of carbon fibre reinforced PEEK hinge mechanism articulation components in a rotating hinge knee design – A comparison of in vitro and retrieval findings. BioMed Research International – Biomaterials 2016; Article ID 7032830.**

43 Kretzer JP, Reinders J, Sonntag R, Hagmann S, Streit M, Jaeger S, Moradi B: **Wear in total knee arthroplasty – just a question of polyethylene?: Metal ion release in total knee arthroplasty. International Orthopaedics 2014;38:335–40.**

44 Mitchelson AJ, Wilson CJ, Mihalko WM, Grupp TM, Manning BT, Dennis DA, Goodman SB, Tzeng TH, Vasdev S, Saleh KJ: **Biomaterial hypersensitivity: is it real? Supportive evidence and approach considerations for metal allergic patients following total knee arthroplasty. BioMed Research International 2015; Article ID 137287.**

45 Thomas P, Weik T, Roider G, Summer B, Thomsen M: **Influence of Surface Coating on Metal Ion Release: Evaluation in Patients With Metal Allergy. Orthopedics 39, S24–30.**

46 Reich J, Hovy L, Lindenmaier H-L, Zeller R, Schwiesau J, Thomas P, Grupp TM: **Präklinische Ergebnisse beschichteter Knie-Implantate für Allergiker. Orthopäde 2010;39(5):495–502.**

47 Puente Reyna AL, Fritz B, Schwiesau J, Schilling C, Summer B, Thomas P, Grupp TM: **Metal ion release barrier function and biotribological evaluation of a zirconium nitride multilayer coated knee implant under highly demanding activities wear simulation. J Biomechanics. 2018;79:88–96.**

48 Nieuwenhuis JJ, Malefijt de JW, Hendriks JC,Gosens T, Bonnet M: **Unsatisfactory results with the cementless Omnifit acetabular component due to polyethylene and severe osteolysis. Acta Orthop Belg 2005;71(3):294–302.**

49 Wasielewski RC, Jacobs JJ, Arthurs B, Rubash HE: **The acetabular insert-metal backing interface: an additional source of polyethylene wear debris. J Arthroplasty 2005;20(7):914–922.**

50 Powers CC, Ho H, Beykirch SE, Huynh C, Hopper RH Jr, Engh CA Jr, Engh CA: **A comparison of a second- and a third-generation modular cup design: is new improved? J Arthroplasty 2010;25(4):514–521.**

51 Mikulak SA, Mahoney OM, Dela Rosa MA, Schmalzried TP: **Loosening and osteolysis with the press-fit condylar posterior-cruciate-substituting total knee replacement. J Bone Joint Surg Am 2001;83(3):398–403.**

52 McEwen HMJ, Barnett PI, Bell CJ, Farrar R, Auger DD, Stone MH, Fisher J: **The influence of design, materials and kinematics on the in vitro wear of total knee replacements. J Biomechanics Vol. 38, pp. 357–365, 2005.**

53 Puente Reyna AL, Jäger M, Floerkemeier T, Frecher S, Delank KS, Schilling C, Grupp TM: **Backside wear analysis of retrieved acetabular liners with a pressfit locking mechanism in comparison to wear simulation in vitro. BioMed Res Int 2016; Article ID 8687131.**

54 Puente Reyna AL, Holderied M, Jäger M, Schilling C, Grupp TM: **Articulation and backside wear analysis after long-term in vitro wear simulation of vitamin E stabilized polyethylene acetabular liners with a press-fit locking mechanism. Traumatology & Orthopedics of Russia 2018;24(2):29–40.**

55 Braun S, Sonntag R, Schroeder S, Mueller U, Jaeger S, Gotterbarm T, Kretzer JP: **Backside wear in acetabular hip joint replacement. Acta Biomaterialia 2019;83:467–476.**

56 Davidson JA, Poggie RA, Mishra AK: **Abrasive wear of ceramic, metal, and UHMWPE bearing surfaces from third-body bone, PMMA bone cement, and titanium debris. Biomed Mater Eng 1994;4(3):213–29.**

57 Wang A, Schmidig G. **Ceramic femoral heads prevent runaway wear for highly crosslinked polyethylene acetabular cups by third body bone cement particles. Wear 2003;255(7–12):1057–63.**

58 Jones VC, Williams IR, Auger DD, Walsh W, Barton DC, Stone MH, Fisher J: **Quantification of third body damage to the tibial counterface in mobile bearing knees. Proc Inst Mech Eng H 2001;215(2):171–9.**

59 Morscher EW, Hefti A, Aebi U. **Severe osteolysis after third body wear due to hydroxyapatite particles from acetabular cup coating. J Bone Joint Surg Br 1998;80(2):267–72.**

60 Hauptmann SM, Weber P, Glaser C, Birkenmaier C, Jansson V, Muller PE: **Free bone cement fragments after minimally invasive unicompartmental knee arthroplasty: an underappreciated problem. Knee Surg Sports Traumatol Arthrosc 2008;16(8):770–5.**

61 Minakawa H, Stone MH, Wroblewski BM, Lancaster JG, Ingham E, Fisher J: **Quantification of third-body damage and its effect on UHMWPE wear with different types of femoral head. J Bone Joint Surg Br 1998;80(5):894–9.**

62 Bragdon CR, Jasty M, Muratoglu OK, O'Connor DO, Harris WH: **Third-body wear of highly cross-linked polyethylene in a hip simulator. Journal of Arthroplasty 2003;18:553–61.**

63 Netter J.D., Hermida J.C., Chen P.C., Nevelos J.E., D'Lima D.D: **Effect of microseparation and third-body particles on dual-mobility crosslinked hip liner wear. J Arthroplasty 2014;29(9):1849–53.**

64 Zietz C, Bergschmidt P, Lange R, Mittelmeier W, Bader R: **Third-body abrasive wear of tibial polyethylene inserts combined with metallic and ceramic femoral components in a knee simulator study. Int J Artif Organs 2013;36(1):47–55.**

65 Schroeder C, Grupp TM, Fritz B, Schilling C, Chevalier Y, Utzschneider S, Jansson V. **The influence of third-body particles on wear rate in unicondylar knee arthroplasty: a wear simulator study with bone and cement debris. J Mater Sci Mater Med 2013;24(5):1319–25.**

66 Paulus AC, Franke M, Kraxenberger M, Schroeder C, Jansson V, Utzschneider S: **PMMA third-body wear after unicondylar knee arthroplasty decuples the UHMWPE wear particle generation in vitro. BioMed Res Int 2015, Article ID 575849**

67 Villiers D de, Traynor A, Collins SN, et al. **The increase in cobalt release in metal-on-polyethylene hip bearings in tests with third body abrasives. Proc Inst Mech Eng H 2015;229:611–8.**

68 Patel PD, Potts A, Froimson MI: **The dislocating hip arthroplasty: prevention and treatment. J Arthroplasty 2007;22:86–90.**

69 Reina N, Pareek A, Krych AJ, Pagnano MW, Berry DJ, Abdel MP: **Dual-mobility constructs in primary and revision total hip arthroplasty: A systematic review of comparative studies. J Arthroplasty 2019;34:594-603.**

70 Batailler C, Fary C, Verdier R, Aslanian T, Caton J, Lustig S: **The evolution of outcomes and indications for the dual-mobility cup: a systematic review. Int. Orthopaedics (SICOT) 2017;41:645-59.**

71 Hartzler MA, Abdel MP, Sculco PK, Taunton MJ, Pagnano MW, Hanssen AD: **Otto aufranc award: dual-mobility constructs in revision THA reduced dislocation, re-revision, and reoperation compared with large femoral heads. Clin Orthop Relat Res 2018;476:293-301.**

72 Kolz JM, Wyles CC, van Citters DW, Chapman RM, Trousdale RT, Berry DF: **In vivo corrosion of modular dual-mobility implants: A retrieval study. J Arthroplasty 2020;35:3326-9.**

73 Abdelaal MS, Zachwieja E, Sharkey PF: **Severe corrosion of modular dual mobility acetabular components identified during revision total hip arthroplasty. Arthroplasty Today 2021;8:78-83.**

74 Bader C, Pfaff A, Reina N, Heller KD, Haas H, Schneider J, Schwiesau J, Grupp TM: **Differences In The Ion-Release Of Modular Dual Mobility Designs In A High-Demand Test Setup. Abstract 22st Virtual EFORT Congress June 30- July 2, 2021.**

75 Grupp TM, Baxmann M, Jansson V, Henning Windhagen H, Heller KD, Morlock MM, Knaebel HP: **How to proceed with asymptomatic modular dual taper hip stems in case of acetabular revision. Materials Manuscript ID: Materials 2020;13,1098; doi:10.3390/ma13051098.**

76 Neville A, Hesketh J, Beadling AR, Bryant MG, Dowson D: **Incorporating corrosion measurement in hip wear simulators: An added complication or a necessity? Proceedings of the Institution of Mechanical Engineers. Part H, J Engineering in Medicine 2016;230(5):406-20.**

77 Yan Y, Dowson D, Neville A: **In-situ electrochemical study of interaction of tribology and corrosion in artificial hip prosthesis simulators. J Mechanical Behavior of Biomedical Materials 2013;18:191-9.**

78 Simoes TA, Bryant MG, Brown AP, Milne SJ, Ryan M, Neville A, Brydson R: **Evidence for the dissolution of molybdenum during tribocorrosion of CoCrMo hip implants in the presence of serum protein. Acta Biomaterialia 2016;45:410-18.**

79 Yan Y, Neville A, Dowson D, Williams S, Fisher J: **Electrochemical instrumentation of a hip simulator: a new tool for assessing the role of corrosion in metal-on-metal hip joints.Proceedings of the Institution of Mechanical Engineers. Part H, J Engineering in Medicine 2010;224(11):1267-73.**

80 Hadley M, Hardaker C, Isaac G, Fisher J: **Wear of different materials for total hip replacement under adverse stop-dwell-start in vitro wear simulation conditions. Proceedings of the Institution of Mechanical Engineers. Part H, J Engineering in Medicine 2018;232(12):1261-70.**

81 Lunn DE, De Pieri E, Chapman GJ, Lund ME, Redmond AC, Ferguson SJ: **Current Preclinical Testing of New Hip Arthroplasty Technologies Does Not Reflect Real-World Loadings: Capturing Patient-Specific and Activity-Related Variation in Hip Contact Forces. J Arthroplasty 2020;35(3):877-85.**

82 Rivière C, Lazic S, Boughton O, Wiart Y, Villet L, Cobb J: **Current concepts for aligning knee implants – Patient specific or systematic? EFORT Open Reviews 2018;4:**

83 Calliess T, Bauer K, Stukenborg-Colsman C, Windhagen H, Budde S, Ettinger M: **PSI kinematic versus non-PSI mechanical alignment in total knee arthroplasty a prospective Randomised study. Knee Surgery, Sports Traumatology, Arthroscopy 2016;24:**

84 Bellemans J, Colyn W, Vandenneucker H, Victor J: **The Chitranjan Ranawat award: is neutral mechanical alignment normal for all patients? The concept of constitutional varus. Clin Orthop Relat Res 2012;470:45–53**

85 Larrainzar-Garijo R, Murillo-Vizuete D, Garcia-Bogalo R, Escobar-Anton D, Horna-Castneiras L, Peralta-Molero J: **Dynamic alignment analysis in the osteoarthritic knee using computer simulation. J Knee Surgery 2017;30(9):909-15.**

86 Hirschmann MT, Moser LB, Amsler F, Behrend H, Leclerq V, Hess S: **Functional knee phenotypes: a novel classification for phenotyping the coronal lower limb alignment based on the native alignment in young non-osteoarthritic patients. Knee Surgery, Sports Traumatology, Arthroscopy 2019;27:doi.org/10.1007/s00167-019-05509-z**

87 Maag C, Cracaoanu I, Langhorn J, Heldreth M: **Total knee replacement wear during simulated gait with mechanical and anatomic alignments. Proceedings of the Institution of Mechanical Engineers. Part H, J Engineering in Medicine 2021;**

88 Maas A, Kutzner I, Larrainzar-Garijo R, Grupp TM: **Development of a simplified method to estimate implant design performance under dynamic loading conditions in total knee replacement systems. Abstract 21. Congress of the European Society of Biomechanics Prague July 5-8, 2015.**

89 Grupp TM, Taylor WR, Hosseini Nasab SH, Larrainzar-Garijo R, Maas A: **Patient musculoskeletal phenotypes, real time biomechanical modeling and predicted individualised Functional outcomes in total knee arthroplasty. Abstract 28. Conference of the European Orthopaedic Research Society. Izmir Sept. 16-19, 2020.**

90 Mell SP, Wimmer MA, Lundberg HJ: **Sensitivity of total knee replacement wear to variability in motion and load input: A parametric finite element analysis study. JOrthopaedic Research 2020;38(7):1538-49.**

91 Abdelgaied A, Fisher J, Jennings LM: **A comprehensive combined experimental and computational framework for pre-clinical wear simulation of total knee replacements. J Mechanical Behavior of Biomedical Materials 2018; 78:282-91.**

92 Fregly BJ, Marquez-Barrientos C, Banks SA, DesJardins JD: **Increased conformity offers diminishing returns for reducing total knee replacement wear. J Biomechanical Engineering 2010;132:2.**

93 Knight LA, Pal S, Coleman JC, Bronson F, Haider H, Levine DL, Taylor M, Rullkoetter PJ: **Comparison of long-term numerical and experimental total knee replacement wear during simulated gait loading. J of Biomechanics 2007;40(7):1550-8.**

94 Kebbach M, Darowski M, Krueger S, Schilling C, Grupp TM, Bader R, Geier A (2020) **Musculoskeletal multibody simulation analysis on the impact of patellar component design and positioning on joint dynamics after unconstrained total knee arthroplasty. Materials Manuscript ID: Materials 2020, 13, 1098; doi:10.3390/ma13102365.**

95 Amstutz HC, Campbell P, Kossovsky N, Clarke IC: **Mechanism and clinical significance of wear debris induced osteolysis. Clin. Orthop. Rel. 1992;276:7-18.**

96 Algan SM, Purdon M, Horowitz SM: **Role of tumor necrosis factor alpha in particulate-induced bone resorption. J Orthop Res 1996;14(1):30-5.**

97 Revell PA, Al-Saffar N, Kobayashi A: **Biological reaction to debris in relation to joint prostheses. Proc. Instn. Mech. Engrs. Part H J Engineering in Medicine 1997; 211:187-97.**

98 Catelas I, Huk OL, Petit A, Zukor DJ, Marchand R, Yahia L: **Flow cytometric analysis of macrophage response to ceramic and polyethylene particles: Effects of size, concentration and composition. J. Biomed. Mater. Res. 1998;41:600-7.**

99 Ingham E, Fisher J: **Biological reactions to wear debris in total joint replacement. Proc. Instn. Mech. Engrs. Part H J Engineering in Medicine 2000;214:21-37.**

100 Ingram JH, Stone M, Fisher J, Ingham E: **The influence of molecular weight, crosslinking and counterface roughness on TNF-alpha production by macrophages in response to ultra high molecular weight polyethylene particles. Biomaterials 2004;25:3511-22.**

101 Green TR, Fisher J, Stone M, Wroblewski BM, Ingham E: **Polyethylene particles of a `critical size` are necessary for the induction of cytokines by macrophages in-vitro. Biomaterials 1998;19:2297-302.**

102 Green TR, Fisher J, Matthews JB, Stone MH, Ingham E: **Effect of size and dose on bone resorption activity of macrophages by in vitro clinically relevant ultra high molecular weight polyethylene particles. J Biomedical Materials Research Appl. Biomater 2000;53:490-97.**

103 Matthews JB, Stone MH, Wroblewski BM, Fisher J, Ingham E: **Evaluation of the response of primary human peripheral blood mononuclear phagocytes challenged with in vitro generated clinically relevant particles of known size and dose. J Biomed Mater Res Appl Biomaterials 2000;44:296-307.**

104 Tipper JL, Galvin AL, Williams S, McEwen HMJ, Stone MH, Ingham E, Fisher J: **Isolation and characterization of UHMWPE wear particles down to ten nanometers in size from in vitro hip and knee joint simulators. J Biomed Mater Res Part A 2006;473-80.**

105 Affatato S, Fernandes B, Tucci A, Esposito L, Toni A: **Isolation and morphological characterisation of UHMWPE wear debris generated in vitro. Biomaterials 2001;22(17):2325-31.**

106 Affatato S, Emiliani D, Bersaglia G, Traina F, Toni A: **An easy technique to digest and isolate UHMWPE wear particles from a hip joint simulator. Int. J Artificial Organs 2004;27(5):424-32.**

107 Richards L, Brown C, Stone MH, Fisher J, Ingham E, Tipper JL: **Identification of nanometre-sized ultra-high molecular weight polyethylene wear particles in samples retrieved in vivo. J Bone Joint Surg [Br] 2008;90B:1106-13.**

108 Grupp TM, Utzschneider S, Schröder C, Schwiesau J, Fritz B, Maas A, Blömer W, Jansson V: **Biotribology of alternative bearing materials for unicompartmental knee arthroplasty. Acta Biomaterialia 2010;6:3601-10.**

109 Tipper JL, Ingham E, Hailey JL, Besong AA, Fisher J, Wroblewski BM, Stone MH. **Quantitative analysis of polyethylene wear debris, wear rate and head damage in retrieved Charnley hip prostheses. Materials Science, Materials in Medicine 2000;11:117-24.**

110 Tipper JL, Hatton A, Nevelos JE, Ingham E, Doyle C, Streicher R, Nevelos AB, Fisher J: **Alumina-alumina artificial hip joints. Part II: Characterisation of the wear debris from in vitro hip joint simulations. Biomaterials 2002;23(16):3441-8**.

111 Minoda Y, Kobayashi A, Iwaki H, Miyaguchi M, Kadoya Y, Ohashi H, Takaoka K: **Characteristics of polyethylene wear particles isolated from synovial fluid after mobile-bearing and posterior-stabilized total knee arthroplasties. J Biomed Mater Res Appl Biomater 2004;61:1-6.**

112 Minoda Y, Aihara M, Sakawa A, Fukuoka S, Hayakawa K, Tomita M, Umeda N, Ohzono K: **Comparison between highly cross-linked and conventional polyethylene in total knee arthroplasty. Knee. 2009;16:348-351.**

113 Orita K, Minoda Y, Sugama R, Ohta Y, Ueyama H, Takemura S, Nakamura H: **Vitamin E-infused highly cross-linked polyethylene did not reduce the number of in vivo wear particles in total knee arthroplasty. Bone Joint J 2020;102B:1527-34.**

114 Inacio MCS, Cafri G, Paxton EW, Kurtz SM, Namba RS: **Alternative bearings in total knee arthroplasty: risk of early revision compared to traditional bearings – An analysis of 62,177 primary cases. Acta Orthopaedica. 2013;84:145-152.**

115 Beyer F, Luetzner C, Kirschner S, Luetzner J: **Midterm Results After Coated and Uncoated TKA: A Randomized Controlled Study. Orthopedics 2016;39(3):S13-7.**

116 Reed M, Howard P, Brittain R, Lawrence S, Stonadge J, Wilkinson M, Wilton T: **National Joint Registry England, Wales, Northern Ireland & Isle of Man. 16th Annual Report 2019.**

117 Graves S, de Steiger R, Lewis P, Harris I: **Australian Orthopaedic Association National Joint Replacement Registry for Hip, Knee & Shoulder Arthroplasty. Annual Report 2019.**

118 Porter M, Howard P, Lawrence S, Reed M, Stonadge J, Wilkinson M: **National Joint Registry England, Wales, Northern Ireland & Isle of Man. Annual Report 2018:102–51.**

119 Grimberg A, Jansson V, Melsheimer O, Steinbrueck A: **German Arthroplasty Registy (EPRD) German Society of Orthopaedics & Orthopaedic Surgery (DGOOC) Annual Report 2019:21–9.**

120 Grimberg AW, Elliott J, Grupp TM, Melsheimer O, Jansson V, Steinbrueck A: **Ceramic-coating in cemented primary total knee arthroplasty is not associated with decreased risk of revision due to early prosthetic joint infection. Journal of Arthroplasty 2021;36:991–7.**

121 Boehler C, Kolbitsch P, Schuh R, Lass R, Kubista B, Giurea A.: **Midterm Results of a New Rotating Hinge Knee Implant: A 5-Year Follow-Up," Biomed Res Int; 2017:Article-ID7532745.**

122 Grupp TM, Kaddick C, Schwiesau J, Maas A, Stulberg SD: **Fixed and mobile bearing total knee arthroplasty – Influence on wear generation, corresponding wear areas, knee kinematics and particle composition. Clinical Biomechanics 2009;24:210–17.**

123 Ingham E, Fisher J: **The role of macrophages in osteolysis of total joint replacement. Biomaterials 2005;26(11):1271–86.**

124 Purdue PE, Koulouvaris P, Potter HG, Nestor BJ, Sculco TP: **The cellular and molecular biology of periprosthetic osteolysis. Clin Orthop Relat Res 2007;454:251–61.**

125 Utzschneider S, Paulus A, Datz JC, Schroeder C, Sievers B, Wegener B, Jansson V: **Influence of design and bearing material on polyethylene wear particle generation in total knee replacement. Acta Biomaterialia. 2009;5:2495–2502.**

126 Utzschneider S, Becker F, Grupp TM, Sievers B, Paulus A, Gottschalk O, Jansson V: **Inflammatory response against different carbon-fiber-reinforced PEEK wear particles compared to UHMWPE in vivo. Acta Biomaterialia 2010;6(11):4296–304.**

127 Zysk SP, Gebhard H, Plitz W, Buchhorn GH, Sprecher CM, Jansson V: **Influence of orthopaedic particulate biomaterials on inflammation and synovial microcirculation in the murine knee joint. J Biomed Mater Res B Appl Biomater 2004;71(1):108–15.**

128 Zysk SP, Gebhard HH, Kalteis T, Schmitt-Sody M, Jansson V, Messmer K, et al. **Particles of all sizes provoke inflammatory responses in vivo. Clin Orthop Relat Res 2005;433:258–64.**

129 Neuerburg C, Loer T, Mittlmeier L, Polan C, Farkas Z, Holdt LM, Utzschneider S, Schwiesau J, Grupp TM, Boecker W, Aszodi A, Wedemeyer C, Kammerlander C: **Impact of vitamin E-blended UHMWPE wear particles on the osseous microenvironment in particle-induced osteolysis. Int. Journal of Molecular Medicine 2016;38:1652–60.**

130 Popoola OO, Orozco Villasenor DA, Fryman JC, Mimnaugh K, Rufner A: **High cycle in vitro hip wear of and in vivo biological response to vitamin E blended highly crosslinked polyethylene. Biotribology 2018;16:10–6.**

RESEARCH TOPIC 9

Methodologies to Evaluate Linear Penetration and Volumetric Wear in Clinical Studies – Imaging Techniques for the Evaluation Of Articulations in THA, TKA to Assess Wear and Function of the Joint Throughout the Expected Implant Lifetime

Authors

Marcus Jäger[1,2], Marcel Dreischarf[3], Thomas M. Grupp[4,5], Claude Rieker[6]

[1]Department of Orthopaedics, Trauma & Reconstructive Surgery, St. Marien Hospital Mülheim an der Ruhr, Mülheim, Germany

[2]Orthopaedics and Trauma Surgery, University of Duisburg-Essen, Essen, Germany

[3]Berlin Institute of Health, Julius Wolff Institute for Biomechanics and Musculoskeletal Regeneration, Charité-Universitätsmedizin Berlin, Germany

[4]Aesculap AG Research & Development, Tuttlingen, Germany

[5]Ludwig Maximilians University Munich, Department of Orthopaedic and Trauma Surgery, Musculoskeletal University Center Munich (MUM), Campus Grosshadern, Munich, Germany

[6]Zimmer Biomet EMEA, Scientific Affairs, Zug, Switzerland

1. Abstract

Question: Biotribology

From a clinical point of view, what kind of proof can show that the planned articulations (THA, TKA) enable the function of the joint throughout the expected implant lifetime without producing a critical amount of wear? Which methodologies are suitable to evaluate linear penetration and volumetric wear in clinical studies (PMCF) and which (imaging) techniques are necessary for the evaluation?

Summary / Recommendation:

The abrasion of polyethylene particles from wear is seen as one main reason for aseptic loosening of prosthetic components in total hip arthroplasty. To non-invasively determine implant wear in vivo, femoral head penetration (FHP) is typically quantified as a measure to determine wear. However, traditional methods such as HAS Martell, EBRA, etc. require substantial manual interaction and thus still suffer from larger error variance and lower reproducibility.

Therefore, these methods require large cohorts for a meaningful evaluation and are less suitable to evaluate modern implants (i.e., highly crosslinked polyethylene, vitamin infusion or blended polyethylene, etc.) with wear rates of less than 30 micro-meters per year.

For the objective evaluation of THA wear in vivo of modern implants with low wear rates, it is recommended to employ accurate and precise methods such as RSA and fully automated CAD 2D-3D-registration.

For TKA, RSA also yields in high accuracy and precision with aforementioned limitations. Consistently to THA, first results show that the CAD-2D-3D registration approach also achieves comparable results to RSA in the wear measurement of TKA on routine C-rays.

2. Level of Evidence

Strong

3. Consensus Delegate Vote

93% - super majority, strong consensus (93% agree / 2% disagree / 5% abstain)

4. Graphical Abstract

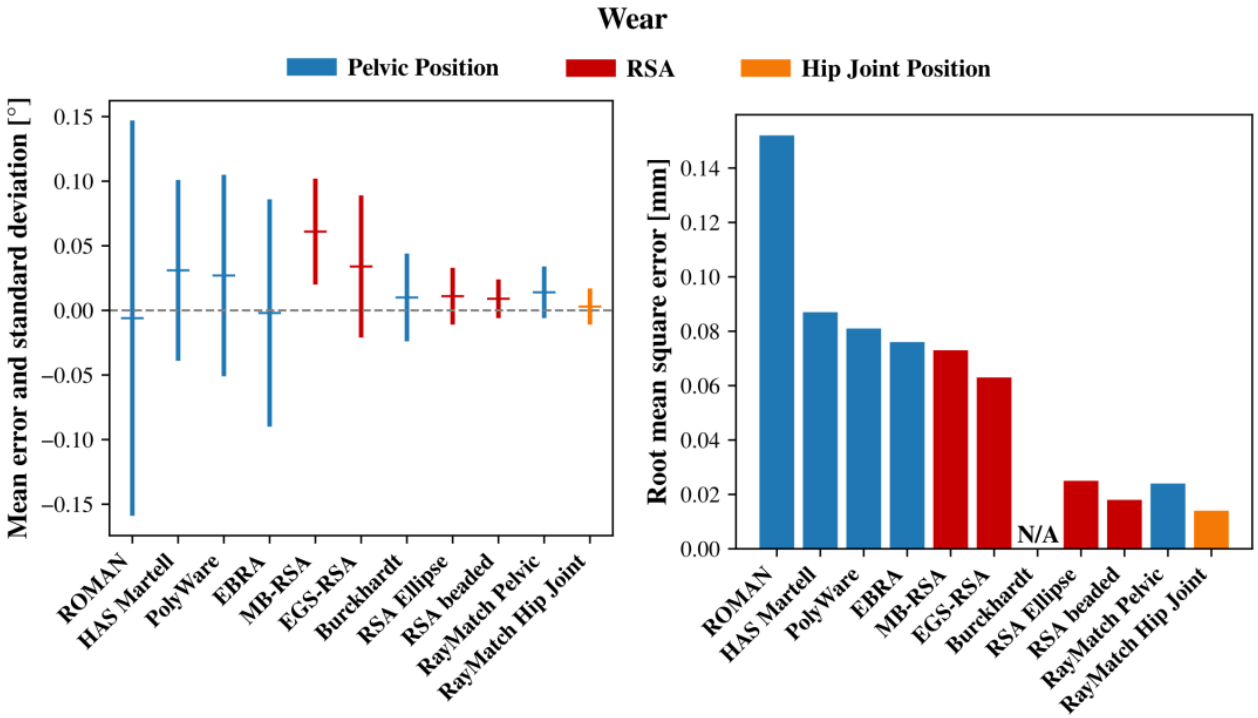


Figure. 1: Accuracy and precision of available wear measurement methods for THA.

5. Search Strategy

Manual search was performed focused on methodologies to evaluate linear penetration and volumetric wear in clinical studies (PMCF) and related (imaging) techniques.

By the expertise given in the group additional publications dealing on THA & TKA in vivo wear measurements and possible thresholds were selected.

Potentially eligible articles were screened in detail for applicability and 25 articles were included in the references.

6. Rationale

From a clinical point of view, what kind of proof can show that the planned articulations (THA, TKA) enable the function of the joint throughout the expected implant lifetime without producing a critical amount of wear?

Which methodologies are suitable to evaluate linear penetration and volumetric wear in clinical studies (PMCF) and which (imaging) techniques are necessary for the evaluation?

The abrasion of polyethylene particles from wear is seen as one main reason for aseptic loosening of prosthetic components in total hip arthroplasty (THA) [1].

To non-invasively determine implant wear in vivo, femoral head penetration (FHP) is typically quantified as a measure to determine wear within the frontal plane between two consecutive anterior-posterior (AP) X-ray images [2, 3]. Based on this idea, several measurement approaches had been developed in past decades. Early approaches started with manual, calliper-based methods proposed by Livermore [4] or Dorr and Wan [5], with limited accuracy and precision.

To overcome these subjective manual measurements, computer-assisted approaches developed by Martell [8] and Devane [7-9] and methods such as EBRA [10], ROMAN [11] and POLYWEAR [12] offer edge-detection-based wear measurements. The highest precision was obtained with Martell's method, guiding the way to a more objective evaluation of wear in vivo (see Figure 1). However, these methods require substantial manual interaction and thus still suffer from larger error variance and lower reproducibility. Therefore, these methods require large cohorts for a meaningful evaluation and are less suitable to evaluate modern implants (i.e., highly crosslinked polyethylene, vitamin infusion or blended polyethylene etc.) with wear rates of less than 30 micro-meters per year [13].

The current considered gold standard for wear measurement is the radio-stereometric analysis (RSA), which uses two X-ray images covering the field of view from two distinct directions and thus offers additional depth information that is unavailable in one plain AP X-ray image [14]. The most accurate RSA approach is based on the implantation of additional measuring beads, which can be easily detected in radiographs, but cannot be implanted in routine clinical practice. Although it provides accurate results with a mean error of 0.009 mm and a standard deviation of 0.015 mm [12], its invasive nature makes it unsuitable for large cohorts and real-world evaluations. Other model-based RSA-approaches (MB-RSA) utilize computer-aided design (CAD) models [15], computer-generated elementary geometric shape models (EGS-RSA) [16], or simply the visible ellipsoidal opening of the cup (RSA ellipse) [17]. Callary et al [12] and Stilling et al [18] evaluated these RSA approaches (MB-RSA, EGS-RSA, RSA Ellipse) without the implantation of additional measuring beads and confirmed their high accuracy and precision (see Figure 1). While RSA offers superior accuracy and precision in the area of two-dimensional wear determination compared to the aforementioned approaches, it comes with the disadvantages of a complex and expensive radiographic setup, not allowing an analysis in clinical routine for multi-center studies of large cohorts.

Another measurement approach is based on a CAD 2D-3D-registration and was developed by Burckhardt et al. [2], Haversath et al. [19] and Klebingat et al. (RayMatch, Raylytic) [20]. This measurement procedure uses implant CAD models and clinical routine AP X-rays. The use of the 3D-information through the CAD models in an iterative optimization process leads to improved accuracy and precision in the determination of implant wear. While Burckhardt et al. [2] still requires the user to manually select a region of interest, the CAD 2D-3D-registration according to the RayMatch method is fully automated and therefore does not rely on subjective user input resulting in full objectivity. Compared to other methods, it does not depend on additional implanted beads, complex setups or manual interaction. This approach resulted in the lowest mean error and standard deviation (0.003 mm $\pm$ 0.014 mm for hip joint images), as shown in Figure 1. Furthermore, it was shown that in vitro determined wear rates [21] were precisely confirmed in vivo using the CAD 2D-3D-registration method [22].

For the objective evaluation of THA wear in vivo of modern implants with low wear rates, it is recommended to employ accurate and precise methods such as RSA and fully automated CAD 2D-3D-registration.

Accurate measurement of wear is also required for other clinical uses cases such as total knee arthroplasty (TKA). Manual approaches such as the Metal-to-Metal or Metal-to-Middle method [23] are available to measure the minimum distances between the implant components and thus the wear on consecutive X-rays with very limited accuracy. For TKA, RSA also yields in high accuracy and precision with aforementioned limitations [24]. Consistently to THA, first results show that the CAD-2D-3D registration approach also achieves comparable results to RSA in the wear measurement of TKA on routine X-rays, but without the described disadvantages [25].

7. References

1. Jiang Y, Jia T, Wooley PH, Yang S-Y. 2013. **Current research in the pathogenesis of aseptic implant loosening associated with particulate wear debris. Acta Orthop Belg 79(1):1–9.**

2. Burckhardt K, Dora C, Gerber C, et al. 2006. **Measuring orthopedic implant wear on standard radiographs with a precision in the 10 microm-range. Med. Image Anal. 10(4):520–529.**

3. Martell JM, Berdia S. 1997. **Determination of polyethylene wear in total hip replacements with use of digital radiographs. J Bone Jt. Surg Am 79(11):1635–41.**

4. Livermore J, Ilstrup D, Morrey B. 1990. **Effect of femoral head size on wear of the polyethylene acetabular component. J. Bone Joint Surg. Am. 72(4):518–528.**

5. Dorr LD, Wan Z. 1995. **Ten Years of Experience With Porous Acetabular Components for Revision Surgery. Clin. Orthop. Relat. Res. 319:191–200.**

6. Martell JM, Berdia S. 1997. **Determination of polyethylene wear in total hip replacements with use of digital radiographs. J Bone Jt. Surg Am 79(11):1635–41.**

7. Devane PA, Bourne RB, Rorabeck CH, et al. 1995. **Measurement of polyethylene wear in metal-backed acetabular cups. I. Three-dimensional technique. Clin. Orthop. (319):303–316.**

8. Devane PA, Bourne RB, Rorabeck CH, et al. 1995. **Measurement of polyethylene wear in metal-backed acetabular cups. II. Clinical application. Clin. Orthop. (319):317–326.**

9. Devane PA, Horne JG. 1999. **Assessment of Polyethylene Wear in Total Hip Replacement: Clin. Orthop. 369:59–72.**

10. Krismer M, Bauer R, Tschupik J, Mayrhofer P. 1995. **EBRA: a method to measure migration of acetabular components. J. Biomech. 28(10):1225–1236.**

11. Geerdink CH, Grimm B, Vencken W, et al. 2008. **The determination of linear and angular penetration of the femoral head into the acetabular component as an assessment of wear in total hip replacement: a comparison of four computer-assisted methods. J. Bone Joint Surg. Br. 90(7):839–846.**

12. Callary SA, Solomon LB, Holubowycz OT, et al. 2017. **Accuracy of methods to measure femoral head penetration within metal-backed acetabular components: accuracy of current methods to measure wear. J. Orthop. Res. 35(5):988–996.**

13. Thoen PS, Nordsletten L, Pripp AH, Röhrl SM. 2020. **Results of a randomized controlled trial with five-year radiostereometric analysis results of vitamin E-infused highly crosslinked versus moderately crosslinked polyethylene in reverse total hip arthroplasty. Bone Jt. J. 102-B(12):1646–1653.**

14. Kärrholm J, Gill RHS, Valstar ER. 2006. **The History and Future of Radiostereometric Analysis: Clin. Orthop. 448:10–21.**

15. Kaptein BL, Valstar ER, Stoel BC, et al. 2003. **A new model-based RSA method validated using CAD models and models from reversed engineering. J. Biomech. 36(6):873–882.**

16. Kaptein BL, Valstar ER, Spoor CW, et al. 2006. **Model-based RSA of a Femoral Hip Stem Using Surface and Geometrical Shape Models: Clin. Orthop. 448:92–97.**

17. Börlin N, Röhrl SM, Bragdon CR. 2006. **RSA wear measurements with or without markers in total hip arthroplasty. J. Biomech. 39(9):1641–1650.**

18. Stilling M, Kold S, de Raedt S, et al. 2012. **Superior accuracy of model-based radiostereometric analysis for measurement of polyethylene wear. Bone Jt. Res. 1(8):180–191.**

19. Haversath M, Klebingat S, Jäger M. 2018. **Abriebanalyse mit virtuellen CAD-basierten Röntgenaufnahmen in der Endoprothetik. Orthop. 47(10):811–819.**

20. Klebingat S, Jäger M, Bien T, et al. 2021. **Highly Accurate Determination of Hip Implant Wear, Cup Anteversion and Inclination Through an AI Automated 2D-3D Registration: An In-Vitro Validation. J. Orthop. Res. (under review)**

21. Grupp TM, Holderied M, Mulliez MA, Streller R, Jäger M, Blömer W, Utzschneider S. **Biotribology of a vitamin E-stabilized polyethylene for hip arthroplasty – Influence of artificial ageing and third-body particles on wear. Acta Biomaterialia 10 (2014) 3068–3078**

22. Busch A, Jäger M, Klebingat S, Baghdadi J, Floerkemeier T, Hütter F, Grupp TM, Haversath M. **Vitamin E-blended highly cross-linked polyethylene liners in total hip arthroplasty: a randomized, multicenter trial using virtual CAD-based wear analysis at 5-year follow-up. Archives of Orthopaedic and Trauma Surgery volume 140, pages1859–1866 (2020)**

23. Collier M, Jewett B, Engh C. 2003. **A Clinical Assessment of Tibial Polyethylene Thickness. J. Arthroplasty 18(7):860–866.**

24. IJsseldijk EA, Valstar ER, Stoel BC, et al. 2011. **The robustness and accuracy of in vivo linear wear measurements for knee prostheses based on model-based RSA. J. Biomech. 44(15):2724–2727.**

25. Lutter C, Grupp T, Selig M, et al. 2021. **Accurate Wear Determination in Clinical Radiographs after Total Knee Arthroplasty using 2D-3D Registration and Artificial Intelligence: An In Vivo Validation. 10th Congress of the German Knee Society. (under review).**

RESEARCH TOPIC 10

Detection of Wear/Debris Complications in Total Hip and Knee Arthroplasty at an Early Follow-up with Regard to Complementary Tests for "Silent Bone Loss" Detection and Risky Modular Implants

Authors

Eduardo Garcia-Rey[1,2], José Cordero-Ampuero[3], George Babis[4], Francesco Benazzo[5,6], Michael Morlock[7]

[1]Hôpital Universitario La Paz-Idi Paz, Madrid, Spain
[2]Centro de Investigación Biomédica en Red de Bioingeniería, Biomateriales y Nanomedicina, CIBER-BBN, Madrid, Spain
[3]Hospital Universitario de La Princesa, Madrid, Spain
[4]2nd Department of Orthopaedics, National and Kapodistrian University of Athens, Medical School, Nea Ionia General Hospital Konstantopouleio, Athens, Greece
[5]Sezione di Chirurgia Protesica ad Indirizzo Robotico, Unità di Traumatologia dello Sport, U.O.C. Ortopedia e Traumatologia Fondazione Poliambulanza, Brescia, Italy
[6]IUSS, Istituto Universitario di Studi Superiori, Pavia, Italy
[7]Hamburg University of Technology, Hamburg, Germany

1. Abstract

Question: Biotribology (wear simulation, wear debris release and biological response)

How can we detect wear/debris complications at an early follow-up?

1. Which complementary tests can improve "silent bone loss" detection like blood samples or advanced imaging in risky patients?
2. How to detect wear/debris from non-articulating interfaces in risky modular implants?

Summary / Recommendation:

Currently, there are different tools and protocols that have demonstrated to be useful to detect "silent bone loss" secondary to wear/debris in patients undergoing total joint replacement. Previous international consensus regarding on the use of risky implants like metal-on-metal total hip replacements (THRs) on the European level with a multidisciplinary approach did provide important information to facilitate this difficult topic. Although most experts agreed that annual clinical and conventional radiological evaluation is recommended, additional imaging with ultrasound, computed tomography (CT)-scan and magnetic resonance imaging (MRI) with metal artifaction sequence (MARS) clearly show significant information. Despite ion analysis is more controversial, cobalt levels in the whole blood can be more practical particularly for risky implants. Metallic femoral heads with a diameter of 36 mm or more, previously recalled resurfacing types and modular-neck THRs can be considered as risky implants.

Further investigations have also reported that wear and debris have been clearly associated with blood metal ion levels and periarticular tissue damage in large-friction THR. More evidence is needed in TKR.

Although risky modular implants are usually easy to define, these imaging techniques and blood samples are very useful to assess pain or other complications like bone loss or mechanical loosening in patients with total joint replacement.

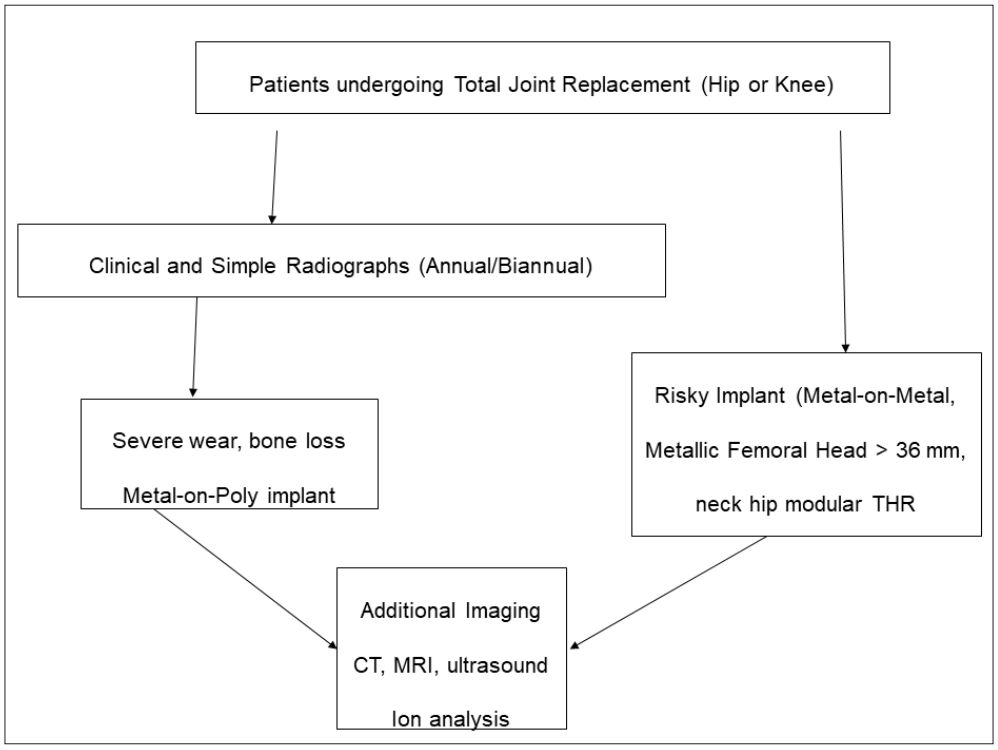
2. Level of Evidence

Strong

3. Consensus Delegate Vote

88% - super majority, strong consensus (88% agree / 10% disagree / 2% abstain)

4. Graphical Abstract



5. Search Strategy

Indexed publications from relevant Journals (Web of Knowledge) and Experts' opinion following EFORT recommended evaluation criteria.

6. Rationale

Concerns about wear in patients undergoing total joint replacement were raised early (1). Pre-clinical and clinical studies have expanded knowledge regarding this topic over the last decades. Wear and osteolysis around implants are considered as a "silent" disease (2). Moreover, further changing implants designs clearly affected the results of total hip replacement (THR), "the operation of the century", due to metallic corrosion from modular interfaces, not only due to wear and debris from bearing surfaces (3). Significant poor survival rates and radiographic bone changes around the implants were early detected. After considering these important clinical findings, it is critical to evaluate wear and debris at both early- and long-term follow-up and secondary bone damage. The interaction between the surgeon, the patient and the implants must be closely monitored.

1. "Risky patient". Young patients usually show high wear rates at long-term (4). Nevertheless, it is probably more important diagnosis and physical activity levels (5), Nowadays, findings obtained from blood samples and advanced imaging techniques can add significant information to detect early "silent bone loss" and wear in this population.

Ion analysis can show a high correlation between elevated blood metal ions (cobalt and chromium concentration) and damaged peri-prosthetic tissue (6), however, the complexity of this response and hypersensitivity influence on the validity of serum metal content as a good predictor (7). Other authors found that the process of ion release in total knee replacement (TKR) is different from metal-on-metal THR (8). The large surface area of the femoral component in TKR is probably a source of metal ion release by corrosion. Lastly, synovial fluid metal levels may show better correlation than blood levels in failed THR (9).

Usually, imaging is easier to do during patients' follow-up. Ultrasound is cheap and can suggest adverse local tissue reaction when compared to magnetic resonance imaging (MRI) (10), however, it is radiologist-dependent (11).

Computed tomography (CT) was considered valid and useful to detect osteolysis, determined bone defects and correlate with wear (12,13). Further investigations confirmed these results (14).

MRI with metal artifaction sequence (MARS) is another very useful tool (10,15,16). A potential prognostic test is dynamic contrast-enhanced MRI (DCE-MRI) as being reported that this novel technique as feasible for evaluation of tissue surrounding THR, however, further clinal studies are needed to confirm those preliminary results (17).

Interestingly, the combination of both blood samples and imaging CT can be useful to detect adverse events associated to metallic implants (18). Well-defined protocols are being reported by experts in international consensus providing valid ways to identify possible complications (19). These multidisciplinary approaches support the evidence for imaging but do not support with strong evidence findings obtained from blood samples.

2. "Risky implants". Again, although basic THR concepts were clearly defined, changing implants offering "newer" solutions did not reach expectations (3). Wear debris can be produced also from non-articulating interfaces (20). Moreover, large friction THR with metallic femoral head diameters larger with 36 mm or more, metal-on-metal THRs, hip resurfacing, and neck-modular hips and can be considered as "risky implants" (21-25). Nevertheless, above mentioned approaches are valid when evaluating these patients.

Although most metal-on-polyethylene THRs are not related to wear debris from non-articulating interfaces, many clinical presentations can present some years postoperatively, particularly with large femoral heads (26). Despite more evidence is suggested due to the lack of retrieval analysis, there is a risk of misdiagnosis of corrosion products from head- neck junction in patients with unexplained pain following metal-on-polyethylene THR and a femoral head diameter of 36 mm (27). The prevalence of symptomatic pseudotumor formation requiring revision surgery with 28 or 32 mm CrCo femoral heads is low and is not associated with high blood ion levels, suggesting other etiologies like individual response (28).

The reintroduction of metal-on-metal hip resurfacing produced new problems associated with the bearing surface, which were severe in the presence of a taper connection (29, 30). Another important issue is that metal-on-metal hip replacement can confound the diagnosis of periprosthetic infection (31).

Other risky implants are neck-modular THRs. The junction of different metallic alloys and the combination of large femoral heads can affect outcome due to the presence of pseudotumours and high serum cobalt-ion levels despite in the absence of pain (32). The revision surgery of these implants is usually more difficult including high dislocation rates (33). Recommendations with complete imaging (ultrasound, CT, MARS-MRI) and whole blood metal ions determination (Cobalt and chromium) are valid (34).

Significant important information has been reported during the last years regarding this difficult topic. It is critical to know risky implants, particularly in patients with a painful joint replacement in order to prevent further complications. The concept of risky patient is more controversial since age, diagnosis, and physical activity is changing with improved medical treatments for systemic diseases and quality of life in elderly patients. Furthermore, to detect early bone loss may be of importance to avoid large bone which may do revision surgery less complicated.

7. References

1. Charnley J. **The long-term results of low-friction arthroplasty of the hip performed as a primary intervention. J Bone Joint Surg Br. 1972 Feb;54(1):61-76. PMID: 5011747.**

2. Harris WH. **The problem is osteolysis. Clin Orthop Relat Res. 1995 Feb;(311):46-53. PMID: 7634590.**

3. Morlock MM. **The taper disaster--how could it happen? Hip Int. 2015 Jul-Aug;25(4):339-46. doi: 10.5301/hipint.5000269. Epub 2015 May 20. PMID: 26044535.**

4. García-Rey E, Carbonell-Escobar R, Cordero-Ampuero J, García-Cimbrelo E. **Outcome of a hemispherical porous-coated acetabular component with a proximally hydroxyapatite-coated anatomical femoral component: an update at 23 to 26 years' follow-up. Bone Joint J. 2019 Apr;101-B(4):378-385. doi: 10.1302/0301-620X.101B4.BJJ-2018-1223.R1. PMID: 30929484.**

5. Garcia-Rey E, Cruz-Pardos A, Garcia-Cimbrelo E. **Alumina-on-alumina total hip arthroplasty in young patients: diagnosis is more important than age. Clin Orthop Relat Res. 2009 Sep;467(9):2281-9. doi: 10.1007/s11999-009-0904-9. Epub 2009 Jun 4. PMID: 19495898; PMCID: PMC2866919.**

6. Lehtovirta L, Reito A, Parkkinen J, Hothi H, Henckel J, Hart A, Eskelinen A. **Analysis of bearing wear, whole blood and synovial fluid metal ion concentrations and histopathological findings in patients with failed ASR hip resurfacings. BMC Musculoskelet Disord. 2017 Dec 11;18(1):523. doi: 10.1186/s12891-017-1894-5. PMID: 29228956; PMCID: PMC5725985.**

7. Lohmann CH, Meyer H, Nuechtern JV, Singh G, Junk-Jantsch S, Schmotzer H, Morlock MM, Pflüger G. **Periprosthetic tissue metal content but not serum metal content predicts the type of tissue response in failed small-diameter metal-on-metal total hip arthroplasties. J Bone Joint Surg Am. 2013 Sep 4;95(17):1561-8. doi: 10.2106/JBJS.L.01273. PMID: 24005196**

8. Postler A, Beyer F, Lützner C, Tille E, Lützner J. **Similar outcome during short-term follow-up after coated and uncoated total knee arthroplasty: a randomized controlled study. Knee Surg Sports Traumatol Arthrosc. 2018 Nov;26(11):3459-3467. doi: 10.1007/s00167-018-4928-0. Epub 2018 Apr 3. PMID: 29616285.**

9. Houdek MT, Taunton MJ, Wyles CC, Jannetto PJ, Lewallen DG, Berry DJ. **Synovial Fluid Metal Ion Levels are Superior to Blood Metal Ion Levels in Predicting an Adverse Local Tissue Reaction in Failed Total Hip Arthroplasty. The Journal of Arthroplasty. 2021, in press**

10. Kwon YM, Dimitriou D, Liow MH, Tsai TY, Li G. **Is Ultrasound As Useful As Metal Artifact Reduction Sequence Magnetic Resonance Imaging in Longitudinal Surveillance of Metal-on-Metal Hip Arthroplasty Patients? J Arthroplasty. 2016 Aug;31(8):1821-7. doi: 10.1016/j.arth.2016.01.033. Epub 2016 Jan 29. PMID: 26895821.**

11. Robinson DJ, Lee S, Marks P, Schneider ME. **Ultrasound Screening for Adverse Local Tissue Reaction after Hip Arthroplasty. Ultrasound Med Biol. 2017 Jul;43(7):1549-1556. doi: 10.1016/j.ultrasmedbio.2017.02.001. Epub 2017 Apr 12. PMID: 28411966.**

12. Howie DW, Neale SD, Stamenkov R, McGee MA, Taylor DJ, Findlay DM. **Progression of acetabular periprosthetic osteolytic lesions measured with computed tomography. J Bone Joint Surg Am. 2007 Aug;89(8):1818-25. doi: 10.2106/JBJS.E.01305. PMID: 17671023.**

13. Garcia-Cimbrelo E, Tapia M, Martin-Hervas C. **Multislice computed tomography for evaluating acetabular defects in revision THA. Clin Orthop Relat Res. 2007 Oct;463:138-43. PMID: 17960676.**

14. Schierjott RA, Hettich G, Graichen H, Jansson V, Rudert M, Traina F, Weber P, Grupp TM. **Quantitative assessment of acetabular bone defects: A study of 50 computed tomography data sets. PLoS One. 2019 Oct 17;14(10):e0222511. doi: 10.1371/journal.pone.0222511. PMID: 31622343; PMCID: PMC6797127.**

15. Morozov PP, Sana M, McGrory BJ, Farraher SW, Abrahams TG. **Comparison of Pre-Revision Magnetic Resonance Imaging and Operative Findings in Mechanically Assisted Crevice Corrosion in Symptomatic Metal-on-Polyethylene Total Hip Arthroplasties. J Arthroplasty. 2017 Aug;32(8):2535-2545. doi: 10.1016/j.arth.2017.03.021. Epub 2017 Mar 18. PMID: 28479059.**

16. Koff MF, Esposito C, Shah P, Miranda M, Baral E, Fields K, Bauer T, Padgett DE, Wright T, Potter HG. **MRI of THA Correlates With Implant Wear and Tissue Reactions: A Cross-sectional Study. Clin Orthop Relat Res. 2019 Jan;477(1):159-174. doi: 10.1097/CORR.0000000000000535. PMID: 30499779; PMCID: PMC6345304.**

17. Anwander H, Cron GO, Rakhra K, Beaule PE. **Perfusion MRI in hips with metal-on-metal and metal-on-polyethylene total hip arthroplasty: A pilot study. Bone Joint Res. 2016 Mar;5(3):73-9. doi: 10.1302/2046-3758.53.2000572. PMID: 26935768; PMCID: PMC4852810.**

18. van der Veen HC, Reininga IH, Zijlstra WP, Boomsma MF, Bulstra SK, van Raay JJ. **Pseudotumour incidence, cobalt levels and clinical outcome after large head metal-on-metal and conventional metal-on-polyethylene total hip arthroplasty: mid-term results of a randomised controlled trial. Bone Joint J. 2015 Nov;97-B(11):1481-7. doi: 10.1302/0301-620X.97B11.34541. PMID: 26530649.**

19. Hannemann F, Hartmann A, Schmitt J, Lützner J, Seidler A, Campbell P, Delaunay CP, Drexler H, Ettema HB, García-Cimbrelo E, Huberti H, Knahr K, Kunze J, Langton DJ, Lauer W, Learmonth I, Lohmann CH, Morlock M, Wimmer MA, Zagra L, Günther KP. **European multidisciplinary consensus statement on the use and monitoring of metal-on-metal bearings for total hip replacement and hip resurfacing. Orthop Traumatol Surg Res. 2013 May;99(3):263-71. doi: 10.1016/j.otsr.2013.01.005. Epub 2013 Mar 16. PMID: 23507457.**

20. Huk OL, Bansal M, Betts F, Rimnac CM, Lieberman JR, Huo MH, Salvati EA. **Polyethylene and metal debris generated by non-articulating surfaces of modular acetabular components. J Bone Joint Surg Br. 1994 Jul;76(4):568-74. PMID: 8027142.**

21. Smith AJ, Dieppe P, Vernon K, Porter M, Blom AW; **National Joint Registry of England and Wales. Failure rates of stemmed metal-on-metal hip replacements: analysis of data from the National Joint Registry of England and Wales. Lancet. 2012 Mar 31;379(9822):1199-204. doi: 10.1016/S0140-6736(12)60353-5. Epub 2012 Mar 13. PMID: 22417410.**

22. Smith AJ, Dieppe P, Howard PW, Blom AW; **National Joint Registry for England and Wales. Failure rates of metal-on-metal hip resurfacings: analysis of data from the National Joint Registry for England and Wales. Lancet. 2012 Nov 17;380(9855):1759-66. doi: 10.1016/S0140-6736(12)60989-1. Epub 2012 Oct 2. PMID: 23036895.**

23. Mihalko WM, Wimmer MA, Pacione CA, Laurent MP, Murphy RF, Rider C. **How have alternative bearings and modularity affected revision rates in total hip arthroplasty? Clin Orthop Relat Res. 2014 Dec;472(12):3747-58. doi: 10.1007/s11999-014-3816-2. PMID: 25070918; PMCID: PMC4397767.**

24. Meneghini RM, Lovro LR, Wallace JM, Ziemba-Davis M. **Large Metal Heads and Vitamin E Polyethylene Increase Frictional Torque in Total Hip Arthroplasty. J Arthroplasty. 2016 Mar;31(3):710-4. doi: 10.1016/j.arth.2015.09.020. Epub 2015 Sep 30. PMID: 26482682.**

25. Tsikandylakis G, Kärrholm J, Hailer NP, Eskelinen A, Mäkelä KT, Hallan G, Furnes ON, Pedersen AB, Overgaard S, Mohaddes M. **No Increase in Survival for 36-mm versus 32-mm Femoral Heads in Metal-on-polyethylene THA: A Registry Study. Clin Orthop Relat Res. 2018 Dec;476(12):2367-2378. doi: 10.1097/CORR.0000000000000508. PMID: 30260863; PMCID: PMC6259897.**

26. Lachiewicz PF, O'Dell JA. **Trunnion corrosion in metal-on-polyethylene hip arthroplasty: a simple diagnosis and treatment? Bone Joint J. 2018;100-B:898-902. doi: 10.1302/0301-620X.100B7.BJJ-2017-1127.R2.**

27. Whitehouse MR, Endo M, Zachara S, Nielsen TO, Greidanus NV, Masri BA, Garbuz DS, Duncan CP. **Adverse local tissue reactions in metal-on-polyethylene total hip arthroplasty due to trunnion corrosion: the risk of misdiagnosis. Bone Joint J. 2015 Aug;97-B(8):1024-30. doi: 10.1302/0301-620X.97B8.34682**

28. Persson A, Eisler T, Bodén H, Krupic F, Sköldenberg O, Muren O. **Revision for Symptomatic Pseudotumor After Primary Metal-on-Polyethylene Total Hip Arthroplasty with a Standard Femoral Stem. J Bone Joint Surg Am. 2018 Jun 6;100(11):942-949. doi: 10.2106/JBJS.17.00616. PMID: 29870445.**

29. Renner L, Faschingbauer M, Schmidt-Braekling T, Boettner F. **Cobalt serum levels differ in well functioning Birmingham resurfacing and Birmingham modular THA. Arch Orthop Trauma Surg. 2016 May;136(5):715-21. doi: 10.1007/s00402-016-2439-3. Epub 2016 Mar 16. PMID: 26983720.**

30. Sidaginamale RP, Joyce TJ, Bowsher JG, Lord JK, Avery PJ, Natu S, Nargol AV, Langton DJ. **The clinical implications of metal debris release from the taper junctions and bearing surfaces of metal-on-metal hip arthroplasty: joint fluid and blood metal ion concentrations. Bone Joint J. 2016;98-B:925-33. doi: 10.1302/0301-620X.98B7.37029.**

31. Yi PH, Cross MB, Moric M, Levine BR, Sporer SM, Paprosky WG, Jacobs JJ, Della Valle CJ. **Do serologic and synovial tests help diagnose infection in revision hip arthroplasty with metal-on-metal bearings or corrosion? Clin Orthop Relat Res. 2015 Feb;473(2):498-505. doi: 10.1007/s11999-014-3902-5. PMID: 25171935; PMCID: PMC4294906.**

32. Kwon YM, Khormaee S, Liow MH, Tsai TY, Freiberg AA, Rubash HE. **Asymptomatic Pseudotumors in Patients with Taper Corrosion of a Dual-Taper Modular Femoral Stem: MARS-MRI and Metal Ion Study. J Bone Joint Surg Am. 2016 Oct 19;98(20):1735-1740. doi: 10.2106/JBJS.15.01325. PMID: 27869625.**

33. Inoue D, Restrepo C, Nourie B, Restrepo S, Hozack WJ. **Patients With Modular-Neck Total Hip Arthroplasty: A Brief Five-Year Follow-Up Study. J Arthroplasty. 2020 Jun;35(6S):S268-S272. doi: 10.1016/j.arth.2020.02.059. Epub 2020 Mar 3. PMID: 32199762.**

34. Grupp TM, Baxmann M, Jansson V, Windhagen H, Heller KD, Morlock MM, Knaebel HP. **How to proceed with asymptomatic modular dual taper hip stems in the case of acetabular revision. Materials (Basel). 2020 Mar 2;13(5):1098. doi: 10.3390/ma13051098. PMID: 32121627; PMCID: PMC7084958.**

RESEARCH TOPIC

SIZE RANGE AND ANATOMICAL DESIGN

RESEARCH TOPIC 11

Appropriateness of Implant Geometry, Sizing Range and Increments for Reconstruction of Anatomical Structures in Hip and Knee Arthroplasty – Review and Suggestion of Methods for Pre-clinical Evaluation

Authors

Dennis Janssen[1], Massimiliano Bernardoni[2], Ingrid Dupraz[3], Ronja A. Schierjott[3]

[1]Radboud University Medical Center, Orthopaedic Research Lab, Research Institute for Medical Innovation, Nijmegen, the Netherlands
[2]Medacta International SA, Castel San Pietro, Switzerland
[3]Aesculap AG Research & Development, Tuttlingen, Germany

1. Abstract

Question: Size range and anatomical design

How can the appropriateness of the implant geometry, sizing range and increments be assessed with respect to the reconstruction of anatomical structures?

Q1: Which methods exist to evaluate the overall implant fit?
Q2: Which methods exist to assess adverse interactions of implants with the surrounding soft and hard tissue?
Q3: How can different sources of anatomical variation be taken into account when assessing implant geometry and sizing?
Q4: Which anatomical variations need to be covered by X-Ray templates or trial implants?

Summary / Recommendation:

Q1: Considering the large variety in measures and relations reported in literature, standardization of morphometric analysis of hip and knee anatomy is recommended. Physical fitting experiments are valuable to confirm implant geometry and sizes after the design phase and to simulate the surgical procedure but should only be used for questions that cannot be answered by alternative methods. Virtual fitting is based on X-rays, CT, or MRI and can be performed in 2D or 3D. The geometric information of such datasets can be captured in statistical shape models (SSMs), which are useful for implant design and evaluation of fit.
Q2: Methods described under Q1 are also suitable to evaluate adverse interactions with the surrounding soft and hard tissues, which may limit range of motion or cause pain.
Q3: Sources of variability (patient age and gender, ethnicity, and disease status and pathology) need to be taken into account when assessing fit and function of implant systems. This can be achieved by matching the input data to the intended patient category. Comprehensive standardized databases summarizing information from large populations are valuable for assessment of fit.
Q4: Templates and virtual trial implants should replicate the final implant design and variation as closely as possible. Trial implants need to reflect fixation status, both for cemented and cementless implant systems.

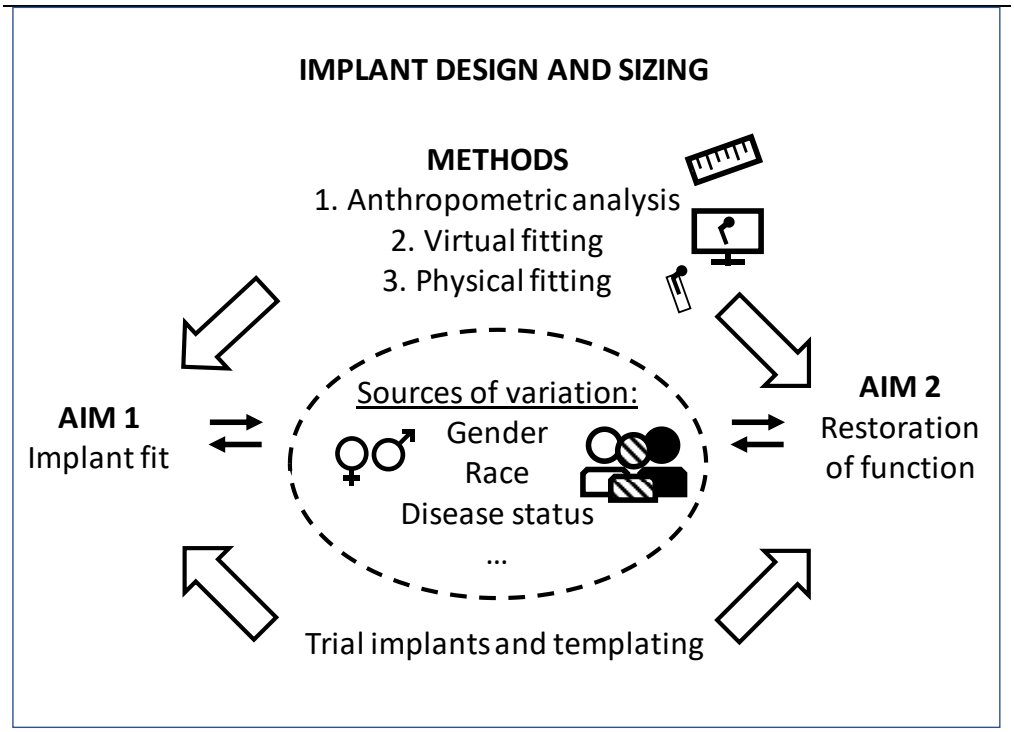
2. Level of Evidence

Moderate

3. Consensus Delegate Vote

98% - unanimous, strongest consensus (98% agree / 2% disagree / 0% abstain)

4. Graphical Abstract



5. Search Strategy

NA

6. Rationale

Total hip arthroplasty (THA) and total knee arthroplasty (TKA) are procedures that are performed worldwide. In order to serve each individual patient, anatomical variation in the patient population needs to be taken into account when designing new implant systems. THA and TKA implant systems should therefore be designed in various shapes and several size increments to accommodate to the individual shape of the patient. In addition, modular stem options are often provided to offer additional flexibility, especially for revision THA and TKA.

The objectives of the sizing range and shape adjustments in implant systems are: 1) to achieve an optimal fit of the implant in the body, and 2) to allow for an optimal functional reconstruction of the joint function.

This consensus paper aims at answering four questions related to implant shape and sizing, with a focus on geometrical fit in the anatomy, and on reconstruction of joint function.

Q1: The fit of an implant system in the patient anatomy is determined by the size range and implant shape. Three main methods are available to evaluate implant fit: 1) morphometric analysis, 2) virtual fitting experiments, and 3) physical fitting experiments.

Morphometric analysis is used to describe the geometrical relations between specific measures describing the native hip and knee joint. Numerous studies have investigated pelvic, femoral, and tibial morphology based on medical imaging data (CT or MRI) [1-5]. For tibial TKA implant systems typically parameters such as medio-lateral (ML) and anterior-posterior (AP) dimensions are measured [4,6,7] to evaluate coverage of the tibial plateau. For the distal femoral geometry, typical parameters include AP and ML width, distal condylar offset and posterior femoral offset [3,8,9]. For THA implant systems studies mainly focus on the width of the intramedullary canal to assess fit of the femoral stem [10]. Pelvic measurements are mainly aimed at measurements of the acetabular diameter to determine required implant sizing [1].

These studies provide a good first overview of the size distribution. A limitation is that currently no standard exists for the specific measures that are used for the description of joint and bone anatomy, and as a consequence a large variety in outcome measures has been reported in literature. In addition, the review and summary of the data is time consuming. Hence, a database summarizing this information from thousands of subjects in a standardized way would be very helpful.

In virtual fitting experiments the fit of implant systems is determined by digitally placing the implant geometry in medical imaging datasets of the joint anatomy. Virtual fitting can be performed in 2D (using planar radiographs) or in 3D models (based on CT or MRI data). Some medical imaging datasets that can be used for such experiments are publicly available, such as through the Osteoarthritis Initiative [11], but manufacturers may build up their own data base [12]. Datasets are also commercially available from selected providers. Data collection and structuring are time-consuming and imaging data is often limited to pathological cases. In addition, the possibility to share knowledge and information is limited.

A traditional analysis of fit can be performed by direct placement of implant geometries in the imaging dataset to assess if the planned implant sizes and increments can cover the patient population. This method can be time-consuming as measurements need to be performed on each individual item of the dataset.

Alternatively, imaging datasets can be used to create a statistical shape model (SSM) of the anatomy. A statistical shape model is a statistical description of the geometries of the training dataset. This training population consists of 3D bone models segmented from numerous medical imaging data sets (e.g. 80 patients) of the anatomy of interest [13,14]. The model is expressed as an average shape that can be manipulated to create new virtual geometries of variable size and morphology using principal component analysis [15]. Hence, it could be used to derive the smallest through largest size required by the patient population (the training dataset), as well as useful size increments [16].

An SSM is only as good as the underlying input dataset and should be validated to ensure that it well represents the patient population (for more information, please see guidelines for validation).

In addition to the three mentioned main methods, data on implant sizes from larger arthroplasty registries can be used in the beginning of the implant development phase as an orientation of common sizing increments, and also retrospectively to confirm the provided implant sizes and size increments.

In physical fitting experiments the fit of an implant system is assessed in cadaveric specimens. Fitting experiments are the closest approximation of the actual clinical situation (soft-tissue interaction, simulation of surgery). Besides the risks of working with cadaveric materials and the costs associated with specimens, limitations include their general limited availability. In order to cover the whole patient population, including very large and very small sizes, a large number of specimens would be required. Hence, the physical models could be useful to additionally confirm implant geometry and sizes after the design phase and to simulate the surgical procedure but should not be used primarily to define the sizes and size increments required by the patient population. They represent a valuable resource and should hence only be used for questions that cannot be answered by alternative methods, such as virtual fitting experiments.

Q2: Placement of an implant in the joint space may lead to adverse interactions with the surrounding soft and hard tissues. The soft tissue envelope therefore already needs to be taken into account in the pre-clinical phase.

Analysis of the range of motion (ROM) can be used to minimize the risk of dislocation of THA reconstructions. Such analysis is possible through computational models of the implant components, bony geometry, and even the surrounding soft tissue structures [17]. Such models simulate impingement of the implant components and bony structures, and their effect on ligament stresses. Computational models allow for the evaluation of the effect of design parameters such as head size, neck geometry, head-neck ratio, femoral offset, and implant positioning (e.g. combined anteversion of the acetabular and femoral component).

Undesired interactions with surrounding soft tissues in TKA can result from overhang of the prosthesis, which may lead to knee pain [18,19]. Oversizing or malpositioning of the tibial components may result in impingement of the medial collateral ligament, popliteus tendon, or the iliotibial band. Oversizing of the femoral component in the anterior-posterior direction can lead to overstuffing of the patellofemoral joint [20], although its effect on knee pain is subject to debate [21]. Moreover, reconstruction of patellar kinematics depends on the trochlear groove in the implant design, which may differ from the native trochlea [22] and is influenced by femoral component rotation [23].

Pre-clinical evaluation of soft tissue impingement can be performed using MRI imaging of reconstructions with non-metallic implant models [24]. Implant overhang can furthermore be evaluated geometrically in 3D models or SSMs as described under sub-question 1.

Q3: Anatomical variability may affect the fit of THA and TKA implant systems. Anatomical variability therefore needs to be taken into account when determining the appropriate implant geometry and sizing, to ensure the implant system matches the target patient population. Sources of variability reported in literature include patient age and gender [25,26], ethnicity [5,27,28], and disease status and pathology [29]. Differences between patient categories may influence specific anatomy and morphology, such as champagne flute-type compared to stovepipe-type femurs [30], anteversion of the femur and acetabulum [31], or (varus) alignment of the knee [32].

Evaluation of implant fit and interactions with surrounding soft and hard tissues requires categorization of the underlying anatomical dataset, matching the target patient population and category. Analyses as described under sub-questions 1 and 2 therefore are only appropriate if the underlying medical imaging dataset is representative of the target population.

Through benchmarking against a comprehensive database that covers a specific target population it is possible to evaluate if another database is representative of a specific population. Benchmarking entails comparison of quantitative parameters (e.g. AP or ML measurements) between the new database and the source database [4]. Benchmarking values can be derived from literature or large databases that summarize the information from multiple studies. Particularly valuable is the comparison of the smallest and largest sizes in input datasets with the benchmarking data to confirm that the dataset covers a realistic range of sizes. Without access to benchmarking values describing a specific target population, it is difficult to assess how many specific cases are required to form a new dataset that is representative of the population. Depending on the number of input datasets, it cannot be guaranteed that the population is fully covered, including the extremes (e.g. largest and smallest 2% of the population). In general, the larger the training set, the more likely these marginal groups are represented in the SSM.

Benchmarking can also be used to assess the quality of SSMs in describing a specific patient population [33]. Additional routine checks for SSMs include assessment of normality of the modes of variation, model convergence (i.e. testing how much the SSM changes if patients are added), and specificity (i.e. how well is SSM able to create virtual patients with a realistic anatomy). The 'leave one out test' is an informative test to confirm if sufficient patients have been included in the source dataset. One patient is removed from the training dataset, after which an SSM is created without this patient. Subsequently, it is assessed how well the SSM is able to predict the anatomy of the patient that was left out [34].

Q4: Pre-operative planning through templating is an important tool to choose the correct implant size and design for each individual patient. Planning can be performed in 2D, based on digital X-rays, and in 3D based on CT, MRI or 2D-3D reconstructions [35]. The applied X-ray templates (2D) and virtual trial implants (3D) should replicate the final implant design and variation as close as possible. Hence, they should cover the same variations as the implants they are mimicking.

Trial implants are used intra-operatively to verify the fit and function of the reconstruction prior to placing the final implant. Both in cemented and cementless reconstructions the dimensions of the trial implants differ from the final implant. To replicate the position and alignment of the final implant, for cemented reconstructions the trial implant needs to take into account the intended cement mantle thickness. For cementless reconstructions the trial implant has to replicate the final implant position without damaging the bone bed, as this would affect the press-fit fixation of the final implant.

7. References

1. Yang Z, Jian W, Li ZH, et al. **The geometry of the bone structure associated with total hip arthroplasty. PLoS One 2014; 9: e91058.**

2. Al-Otaibi ML, Alshumrani GA. **Western-designed total knee implants mismatch Saudi patient knees. Anatomical bony measurements documented by computed tomography. Saudi Med J 2020; 41: 1227-1233.**

3. Meier M, Zingde S, Steinert A, et al. **What Is the Possible Impact of High Variability of Distal Femoral Geometry on TKA? A CT Data Analysis of 24,042 Knees. Clin Orthop Relat Res 2019; 477: 561-570.**

4. Dai Y, Bischoff JE. **Comprehensive assessment of tibial plateau morphology in total knee arthroplasty: Influence of shape and size on anthropometric variability. J Orthop Res 2013; 31: 1643-52.**

5. Dai Y, Scuderi GR, Penninger C, et al. **Increased shape and size offerings of femoral components improve fit during total knee arthroplasty. Knee Surg Sports Traumatol Arthrosc 2014; 22: 2931-40.**

6. Servien E, Saffarini M, Lustig S, et al. **Lateral versus medial tibial plateau: morphometric analysis and adaptability with current tibial component design. Knee Surg Sports Traumatol Arthrosc 2008; 16: 1141-5.**

7. Zhang Y, Chen Y, Qiang M, et al. **Morphometry of the Tibial Plateau at the Surface and Resected Levels. J Arthroplasty 2017; 32: 2563-2567.**

8. Bonnin MP, Saffarini M, Bossard N, et al. **Morphometric analysis of the distal femur in total knee arthroplasty and native knees. Bone Joint J 2016;98-B: 49-57.**

9. Mahfouz M, Abdel Fatah EE, Bowers LS, et al. **Three-dimensional morphology of the knee reveals ethnic differences. Clin Orthop Relat Res 2012; 470: 172-85.**

10. Noble PC, Kamaric E, Sugano N, et al. **Three-dimensional shape of the dysplastic femur: implications for THR. Clin Orthop Relat Res 2003; 417: 27-40.**

11. https://nda.nih.gov/oai/

12. Banerjee S, Faizan A, Nevelos J, et al. **Innovations in hip arthroplasty three-dimensional modeling and analytical technology (SOMA). Surg Technol Int 2014; 24: 288-94.**

13. Heimann T, Meinzer HP. **Statistical shape models for 3D medical image segmentation: a review. Med Image Anal 2009; 13: 543-63.**

14. Vanden Berghe P, Demol J, Gelaude F, et al. **Virtual anatomical reconstruction of large acetabular bone defects using a statistical shape model. Comput Methods Biomech Biomed Engin 2017; 20: 577-586.**

15. Ambellan F, Lamecker H, von Tycowicz C, et al. **Statistical Shape Models: Understanding and Mastering Variation in Anatomy. In: Rea P. Biomedical Visualisation. Advances in Experimental Medicine and Biology, vol 1156. Springer, Cham.; 2019. p.67-84.**

16. Dupraz I, Bollinger A, Utz M, et al. **Using Statistical Shape Models to Define a Set of Femur Geometries That Well Represent the Population, Orthopaedic Proceedings 2020; 102-B Supp. 2: 80.**

17. Bunn A, Colwell CW Jr, D'Lima DD. **Effect of head diameter on passive and active dynamic hip dislocation. J Orthop Res 2014; 32: 1525-31.**

18. Mahoney OM, Kinsey T. **Overhang of the femoral component in total knee arthroplasty: risk factors and clinical consequences. J Bone Joint Surg Am 2010; 92: 1115-21.**

19. Bonnin MP, Schmidt A, Basiglini L, et al. **Mediolateral oversizing influences pain, function, and flexion after TKA. Knee Surg Sports Traumatol Arthrosc 2013; 21: 2314-24.**

20. Matz J, Lanting BA, Howard JL. **Understanding the patellofemoral joint in total knee arthroplasty. Can J Surg 2019; 62: 57-65.**

21. Kemp MA, Metcalfe AJ, Sayers A, et al. **Does overstuffing of the patellofemoral joint in total knee arthroplasty have a significant effect on postoperative outcomes? Knee 2018; 25: 874-881.**

22. Barink M, Van de Groes S, Verdonschot N, et al. **The difference in trochlear orientation between the natural knee and current prosthetic knee designs; towards a truly physiological prosthetic groove orientation. J Biomech 2006; 39: 1708-15.**

23. Anglin C, Brimacombe JM, Hodgson AJ, et al. **Determinants of patellar tracking in total knee arthroplasty. Clin Biomech (Bristol, Avon) 2008; 23: 900-10.**

24. Bonnin MP, Van Hoof T, De Kok A, et al. **Imaging the implant-soft tissue interactions in total knee arthroplasty. J Exp Orthop 2016; 3: 24.**

25. Carmona M, Tzioupis C, LiArno S, et al. **Upper Femur Anatomy Depends on Age and Gender: A Three-Dimensional Computed Tomography Comparative Bone Morphometric Analysis of 628 Healthy Patients' Hips. J Arthroplasty 2019; 34: 2487-2493.**

26. Koh YG, Nam JH, Chung HS, et al. **Morphometric study of gender difference in osteoarthritis posterior tibial slope using three-dimensional magnetic resonance imaging. Surg Radiol Anat 2020; 42:667-672.**

27. Sengodan VC, Sinmayanantham E, Kumar JS. **Anthropometric analysis of the hip joint in South Indian population using computed tomography. Indian J Orthop 2017; 51:155-161.**

28. Ma QL, Lipman JD, Cheng CK, et al. **A Comparison Between Chinese and Caucasian 3-Dimensional Bony Morphometry in Presimulated and Postsimulated Osteotomy for Total Knee Arthroplasty. J Arthroplasty 2017; 32: 2878-2886.**

29. Noble PC, Kamaric E, Sugano N, et al. **Three-dimensional shape of the dysplastic femur: implications for THR. Clin Orthop Relat Res 2003; 417: 27-40.**

30. Noble PC, Alexander JW, Lindahl LJ, et al. **The anatomic basis of femoral component design. Clin Orthop Relat Res 1988; 235: 148-65.**

31. Maheshwari AV, Zlowodzki MP, Siram G, et al. **Femoral neck anteversion, acetabular anteversion and combined anteversion in the normal Indian adult population: A computed tomographic study. Indian J Orthop 2010; 44: 277-82.**

32. Shetty GM, Mullaji A, Bhayde S, et al. **Factors contributing to inherent varus alignment of lower limb in normal Asian adults: role of tibial plateau inclination. Knee 2014; 21: 544-8.**

33. Ali AA, Clary CW, Smoger LM, et al. **Computational framework for population-based evaluation of TKR-implanted patellofemoral joint mechanics. Biomech Model Mechanobiol 2020; 19: 1309-1317**.

34. Deckx J, Jacobs M, Dupraz I, et al. **Ability of a statistical shape model of the Caucasian knee to predict the morphology of any Caucasian knee. Orthopaedic Proceedings 2020; 102-B Supp. 2: 84.**

35. Guenoun B, Zadegan F, Aim F, et al. **Reliability of a new method for lower-extremity measurements based on stereoradiographic three-dimensional reconstruction. Orthop Traumatol Surg Res 2012; 98: 506-13.**

RESEARCH TOPIC 12

Appropriateness of Implant Geometry, Sizing Range and Increments for Reconstruction of Anatomical Structures in Hip and Knee Arthroplasty – Review and Suggestion of Methods for Clinical Evaluation

Authors

Francesco Benazzo[1,2], Bernd Grimm[3], Claudia Mazzà[4,5], Fabio Mancino[6–8], Ronja A. Schierjott[9]

[1]Sezione di Chirurgia Protesica ad Indirizzo Robotico, Unità di Traumatologia dello Sport, U.O.C. Ortopedia e Traumatologia Fondazione Poliambulanza, Brescia, Italy

[2]IUSS,Istituto Universitario di Studi Superiori, Pavia, Italy

[3]Human Motion, Orthopaedics, Sports Medicine and Digital Methods Group, Luxembourg, Institute of Health, Transversal activities, Luxembourg, Luxembourg

[4]INSIGNEO Institute for in silico Medicine, The University of Sheffield, Sheffield, UK

[5]Department of Mechanical Engineering, The University of Sheffield, Sheffield, UK

[6]Division of Orthopaedics and Traumatology, Department of Aging, Neurological, Orthopaedic and Head-Neck studies, Fondazione Policlinico Universitario Agostino Gemelli IRCCS, Rome, Italy

[7]Catholic University of the Sacred Heart, Rome, Italy

[8]Adult Reconstruction and Joint Replacement Service, Division of Orthopaedics and Traumatology, Department of Aging, Neurological, Orthopaedic and Head-Neck studies, Fondazione Policlinico, Universitario Agostino Gemelli IRCCS, Roma, RM, Italy

[9]Aesculap AG Research & Development, Tuttlingen, Germany

1. Abstract

Question: Size range and anatomical design

How can it be clinically assessed that provided implant sizes can cover the majority of the patients' characteristics in terms of size increments and range and that the implant's geometry allows appropriate reconstruction of the anatomical structures?

Summary / Recommendation:

The coverage of the patient population should already be assessed pre-clinically while clinical studies and registry data should only be used for further confirmation. The implant fit for each individual patient should be assessed as part of pre-operative planning.

Large databases (radiographs, CT, MRI data) and statistical shape models (SSMs) nowadays provide the possibility to virtually implant the prosthesis in numerous patients prior to clinical application.* The possibility to develop and validate implant designs, sizes etc. based on sufficiently large virtual patient cohorts should be used before first clinical application of the implant to avoid potential risks for the patient due to improper component geometry or sizing. Clinical studies can be used to additionally confirm the implant fit, while registry data are needed to evaluate the overall performance. On the other hand, pre-operative planning based on state-of-the art 2D or 3D reconstructions is recommended for each individual patient in order to provide the best possible fit while optimizing implant stocking, ensuring availability for planned surgical cases, maximizing operating room efficiency, and minimizing the risks related to malpositioning and incorrect sizing of the prosthetic implants.

* Care should be taken regarding input data for such databases and SSMs in terms of:

- Sufficiently large number of data sets
- Gender
- Age
- Ethnicity
- Height, weight
- Osteoarthritis status

Please refer to Consensus Statement 11/Size range and anatomical design for more detailed information.

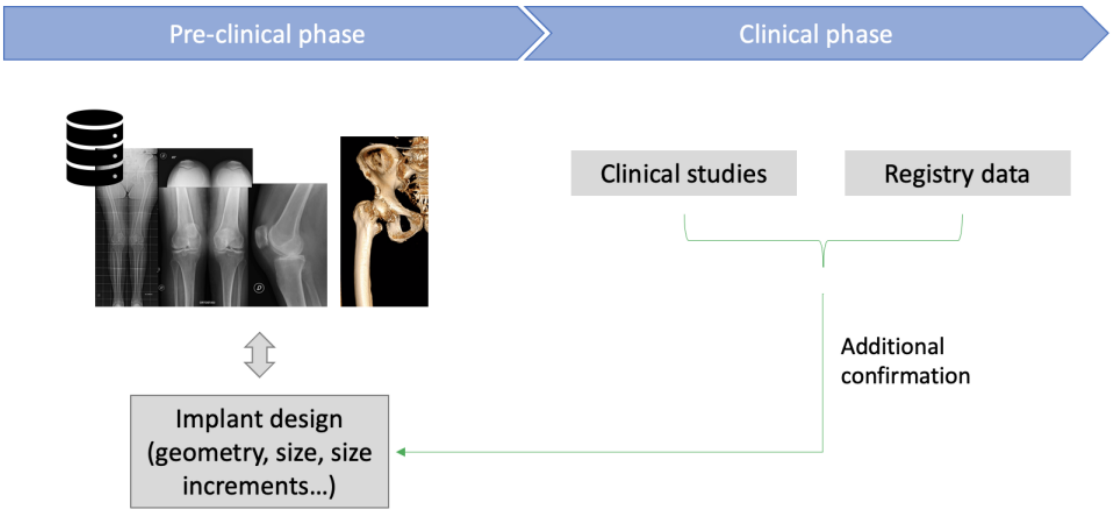
2. Level of Evidence

Very high

3. Consensus Delegate Vote

97% - unanimous, strongest consensus (97% agree / 0% disagree / 3% abstain)

4. Graphical Abstract



Adapted from Benazzo et al., 2019 (15); Chen et al., 2020 (55)

5. Search Strategy

Two literature searches (one for each sub question) were conducted in April 2021.

For the first sub question (How can different variants of an implant (sizes / combinations) be assessed clinically and how can be proven that they fit well and are not associated with any disadvantages for the patient?) the electronic databases EMBASE including PubMed were used applying the following search-terms combined with the Boolean operator "AND":

- (THA specific:)Total hip prosthesis OR total hip replacement OR total hip resurfacing arthroplasty / (TKA specific:) Total knee arthroplasty OR total knee prosthesis
- (THA specific:) Implant OR prosthesis OR stem / (TKA specific:) implant OR prosthesis OR tibial component OR tibial tray OR femoral component
- Size OR sizing OR mismatch OR overhang OR bone coverage OR shape OR geometry OR size increment OR size combination OR morphology OR design
- Functional assessment OR clinical assessment OR radiograph OR tomography OR outcome assessment

To obtain insight in the current state of the art, only articles between 2011 and 2021 were included. Full-text articles and conference articles were included. Article language was not restricted.

For the second sub question (Which anatomical variations / conditions need to be covered by X- Ray templates or trial implants and is the supply of such templates compulsory?) the US National Library of Medicine (PubMed/MEDLINE), EMBASE, and the Cochrane Database of Systematic Reviews were queried for publications utilizing various combinations of the search terms "total joint arthroplasty," "total knee arthroplasty," "TKA," "total hip arthroplasty," "THA," "implant size," "trial size," "implant fit," "preoperative planning," "morphometric," and "preoperative templating," in combination with the Boolean operators (AND, OR, *) since inception of database to April 2021. No limit was set with regard to the year of publication. Only articles written in English language were selected. Only abstracts that were considered relevant with the assigned question were reviewed. If the title and abstract of each study contained insufficient information, the full manuscript was reviewed.

Given the short timeframe, not all publications found to be eligible, but 54 representative articles were included in this statement.

6. Rationale

Appropriate fit of the prothesis of total hip arthroplasty / total knee arthroplasty (THA / TKA) is crucial to avoid the risk of overhanging (with consequent soft tissue impingement and irritation, and possible overstuffing) (1,2), underhanging (with consequent possible subsidence) (3), and intraoperative fractures (hammering/reaming) (4). However, this process is complicated by the fact that the bone morphology, especially in the proximal and distal part of the femur, as well as of the proximal part of the tibia varies among ethnicities (5–7), genders (5,8) and age groups (predominantly in women) (8,9). The challenge for implant development is hence to design implants that cover this broad range of aspects, i.e. to provide well-fitting implants for the whole patient population.

In order to provide / develop an implant system that covers the patient population in terms of appropriate implant geometry, sizes, size increments and size combinations, it is necessary to establish methods that can allow to assess and validate this coverage. However, it should be kept in mind that the quality of fixation is surgeon and technique dependent and as such not strictly related only to fitting.

A review of current literature has been conducted to answer the question of how these factors could be assessed clinically. Given the short timeframe, this review was primarily driven by previous knowledge and work in the field already carried out by the team members.

Clinical / functional outcome assessments after THA / TKA

In clinical practice, the outcome of a total hip (THA) or total knee arthroplasty (TKA) can be assessed using different outcome scores (e.g. Harris Hip Score or American Knee Society Score), Patient Reported Outcome Measurements (PROMs), as well as capacity and performance tests (e.g. timed 20-metre-walk or step test) (10). In addition, novel scores such as the Forgotten Joint Score have been developed in order to assess how natural the prosthesis feels after total joint arthroplasty (TJA) (11). Based on these measurements, one can see if an implant or a treatment does not work well for a specific patient. However, it is not possible to infer if this may be due to malalignment, improper implant geometry or sizing, or other reasons (soft tissue problems, disappointing surgery). In fact, patient dissatisfaction after TJA is multifactorial, even though persistent pain is considered the most frequent reason of dissatisfaction after either THA or TKA (12). Among the causes of persistent pain, implant malpositioning is commonly recognized as one of the more common after TJA (13, 14), likely associated to soft tissue unbalancing.

Image-based assessments

To assess implant sizing, positioning, bone coverage and potential component overhangs or component size mismatch post-operatively, medical imaging (e.g. radiographs, CT etc.) is required (15, 16).

In TKA, anterior-posterior (AP), lateral radiographs and full-length radiographs of both lower limbs can be used for the assessment of component alignment (11), as well as to assess tibial or femoral overhanging lines, and anterior and posterior coverage lines (15). In a study by Klasan et al., 2020 a full-length CT-scan in combination with registration of the implant 3D-model additionally enabled analyzing the percentage of bone coverage, percentage and position of overhang and rotation of the tibial component (17).

In THA, AP-radiographs (partially combined with lateral radiographs) can be used to assess stem alignment and acetabular cup positioning (18) and to confirm adequacy of the chosen component size (16). In addition, it is possible to assess the intramedullary fill

achieved by the femoral component (18), as well as the bony coverage of the acetabular component (18) and potential anterior or lateral overhangs of the acetabular component (19). CT-scans and superimposition of the 3D-model or contour of the implanted acetabular component additionally offers the possibility to assess bone coverage in 3D (20, 21).

Component size mismatch (CSM) in terms of incorrectly applied size combinations of implant components in THA can occur between the femoral head and acetabular cup, between femoral head and trunnion stem (22,23) and, in TKA, between femoral and tibial component (24). Potentially, medical imaging techniques could be used to detect CSM. However, CSM can be missed on plain radiographs (22). According to the NHS, CSM in THA represents an avoidable Never Event (22). Hence the risk of CSM should be eliminated by measures such as clear component labeling, well-organized hospital inventories and pre-operative planning, which become effective prior to the surgery (22).

The question if an implant system covers the majority of patients could hence potentially be assessed in a clinical study, where medical imaging can be the tool to enable this assessment. However, the meaningfulness of such a study would be very limited for the large and small implant sizes (margin sizes), as these are usually not sufficiently represented in the patient population. In order to account for this, a very large patient cohort would be required. In addition, this information is already needed at an earlier stage, i.e. during implant development and hence obtaining it retrospectively from a clinical study would represent a potential risk for the patients involved. In addition, this would entail design loops at an unnecessary late stage.

Besides cadaver specimens (physical models), large medical imaging databases and statistical shape models (SSMs) nowadays offer the possibility to design and validate implants based on large virtual patient cohorts (25,26). (Please refer to Consensus Statement 11/Size range and anatomical design for more detailed information.)

Hence, the question if an implant system covers the majority of patients should be primarily answered pre-clinically to avoid any potential issues related to the anatomical mismatch.

Applying this concept and addressing these questions already in the pre-clinical phase would help to ensure that any potential problems related to sizes / size combinations or implant geometry are detected before the first clinical application in the patient. Nonetheless, image-based assessments within clinical studies (pre-CE or PMCF) and registry data can be used to further confirm implant sizing, and to assess whether there are any potential complications related to the chosen size combinations (e.g. femorotibial size combinations).

Pre-operative planning and templating

For each patient individually pre-operative planning is vital to choose component size and size combination and position most suitable for the patient. This essentially corresponds to the same process as suggested for the pre-clinical and implant development phase - i.e. virtual implantation –but in this case planning / virtual implantation are performed for a specific patient and not for a whole patient cohort.

If the (X-Ray) templates are developed based on numerous clinical datasets alongside the corresponding implants, they should also allow to cover most of the anatomical variations in the clinical application (27). As a matter of fact, templates can be built on plain X-rays, or CT scans: this is the stage in which anatomical variations are acquired and the surgeon tries to fit the metal on them (28). At that point, they do not cover the target (joint), but only the bullets (prosthesis). Hitting the target is a two-step procedure: pre-op (planning), and intra-op (surgery). At the end of the road, pre-clinical anatomical data to design the prosthesis upon, is still the main point. Appropriate sizing of the prosthetic components is of crucial importance to accommodate the reconstruction of the anatomy and the function of the joint (29-31).

Preoperative planning is a crucial element for the success of modern joint arthroplasty, considering the number of sizes and designs commercially available (32,33). It is an important way to help optimize implant stocking, ensuring availability for planned surgical cases, and to maximize operating room efficiency (34). Correct prediction of component sizes can result in reduced surgical times and intraoperative complications including component malpositioning, leg length discrepancy and incorrect reestablishment of femoral offset and center of rotation (35). In TKA, oversizing is frequently encountered with commonly used implants, and overhang rates may be reduced with the usage of systems with narrow options (36,37). However, despite a common tendency to attribute poor fitting of the prosthetic component to patient anatomy and implant design, the surgeon's decision also plays a crucial role (38). The development of trials of multiple sizes based on anthropometric measurements allows a surgeon to identify the best fit for the patient's anatomy finding the right balance between bone coverage and soft tissue tension. Pre-operative templating, and avoidance of the above-described errors, is mandatory to avoid legal issues (liability). In addition, it compensates for the hospital's economic effort by reducing the implant inventory. However, careful evaluation of the available imaging during preoperative planning is necessary in case of severe deformities when the plan is performed on the contralateral side since side-to-side differences may affect the final results (39-45).

Use of technology in pre-operative planning and implant sizing

In the last two decades, progressive advances in technology have led to the development of a more personalized approach to TJA based on additional preoperative studies in order to better find the correct placement and fit of the implant components. The computer-assisted-surgery technique (CAS) was developed in TKA in order to improve intraoperatively the positioning and sizing of the prosthetic components and to provide better functional outcomes. CAS technique provided live on-screen information on patient anatomy and knee kinematics, sizing and positioning of the implant, after acquisition of the bony landmarks (bone morphology and mechanical axis). The bone mapping of the patient could have been obtained through preoperative CT scan (image-based navigation) or through intraoperative mapping of bony anatomical landmarks (non-image-based navigation). However, despite the improved alignment obtained, CAS systems have not become a panacea, due to increased capital costs, operative times, and because the associated learning curve limited its widespread acceptance. The objective of combining the advantages of CAS techniques, while eliminating its significant disadvantages, has led to

the development of patient-specific cutting guides in TKA. In fact, patient-specific instrumentation (PSI) (46), also known as "patient-matched instrumentation," "custom-fit instrumentation," or "custom-made instrumentation", has been developed to reduce the technical difficulties and invasiveness associated with standard TKA. A three-dimensional (3D) model of the patient's anatomy can be generated from preoperative CT or MRI in order to allow the exact templating of the proposed bony resections. In addition, 3D models of the femoral and tibial components are created to determine their optimal size, position, and alignment. Finally, rapid prototyping technology is used to fabricate disposable, custom-cutting guides that fit on the patient's native anatomy. In this way, the surgical planning process was moved into the early preoperative time period in order to allow a personalization of the surgical technique. However, multiple disadvantages were recognized including delayed surgery, reduced accuracy in case of severe deformities, and the additional costs necessary for preoperative studies and manufacturing to fabricate the actual cutting blocks. In addition, the application of this technology was expected to improve functional outcomes and decrease revision rates. However, TKA-PSI did not show superior clinical or implant positioning outcomes when compared with standard TKA groups (47). Further application of technology led to the development of robotic-assisted TJA, where computer software converts anatomical information into a virtual patient-specific three-dimensional (3D) reconstruction of the patient's joint. Robotic-assistance is used by the operating surgeon to calculate optimal bone resection, implant sizing, positioning, and ligaments compliance. This intraoperative robotic device helps to execute the preoperative patient-specific plan with a higher level of accuracy compared to conventional techniques (48).

The application of 3D technology has shown favorable results also in THA (49,50). In fact, 3D preoperative planning has been reported to improve accuracy in predicting femoral and acetabular component size compared to conventional 2D planning (51,52,53). Nonetheless, multiple concurrent factors have been addressed including radiation dose, technical difficulties, and increased costs. In addition, despite the promising results for the accuracy of 3D templating, there may still be potential for improvements regarding the prediction of stem and cup sizes (53). Still, preoperative planning relies on subjective decision of the examiner, which may lead to inaccuracy (54). In order to clarify the long-term impact of 3D planning in long-term outcomes and survivorship after TJA, further studies are needed.

Despite that, the additional imaging necessary for the development of PSI instrumentation and image-based robotic-assisted surgery allowed the formation of large data sets of morphotype, precious for the preclinical studies necessary for new implants development.

The role of 3D-printing technology in preoperative planning

3D-printed technology seems to represent the natural evolution of the application of technology advances in the setting of TJA and it has become a fast-growing innovation in the medical field. An increase in research and publications involving 3D-printing applications in orthopedic surgery and related fields was observed, especially in recent ten years. Today, 3D-printing technology is used to create simulation models or medical implants, thus significantly aiding doctors and medical companies by optimizing the way a surgeon plans and executes a procedure (41). The use of 3D-printed technology can be used to develop anatomical models based on patient imaging necessary for better understanding the patho-anatomy of the patient in case of complex procedures, enabling surgeons to simulate the surgery and potentially improve its execution. In addition, it allows the synthesis of patient-specific instruments (PSI) that may increase the accuracy of a surgery. Moreover, it may be used to produce arthroplasty implants in order to better fit the bony anatomy and increasing its primary stability and adaptation to the patient's characteristics. Finally, it may be used for the development of custom implants. In fact, unlike standard sized implants, a custom implant created using patient-specific medical images can be a perfect match for the patient's unique anatomy. Despite the fascinating application of 3D-printing technology, there remain multiple limitations including accuracy and resolution to be further proven, and increased costs. Recognizing that 3D-printing technology represents the future of TJA, we cannot state that it represents the present. In fact, in order to actively apply 3D-printing technology to the hip and knee surgery fields multiple advancement are still necessary including the advances in biomimetic material, the development of integrated all-in-one computer platform that allows easy planning, better clarification of countries' regulations and regulatory requirements, and implement validation and quality assurance steps when using customized tools and implants.

7. References

1. Barrow JA, Divecha HM, Panchani S, Boden R, Porter ML, Board TN. **Does oversizing an uncemented cup increase post-operative pain in primary total hip arthroplasty? Eur J Orthop Surg Traumatol 2019; 29(1):97–102.**

2. Garceau SP, Enns PA, Teo GM, Weinblatt AI, Aggarwal VK, Long WJ. **Lack of small tibial component size availability for females in a highly utilized total knee arthroplasty system. Knee Surg Sports Traumatol Arthrosc 2020.**

3. Lee YS, Yun JY, Lee BK. **Tibial component coverage based on bone mineral density of the cut tibial surface during unicompartmental knee arthroplasty: clinical relevance of the prevention of tibial component subsidence. Arch Orthop Trauma Surg 2014; 134(1):85–9.**

4. Bonnin MP, Neto CC, Aitsiselmi T, Murphy CG, Bossard N, Roche S. **Increased incidence of femoral fractures in small femurs and women undergoing uncemented total hip arthroplasty - why? Bone Joint J 2015; 97-B(6):741–8.**

5. Yang Z, Jian W, Li Z, Jun X, Liang Z, Ge Y et al. **The geometry of the bone structure associated with total hip arthroplasty. PLoS One 2014; 9(3):e91058.**

6. Al-Otaibi ML, Alshumrani GA. **Western-designed total knee implants mismatch Saudi patient knees. Anatomical bony measurements documented by computed tomography. Saudi Med J 2020; 41(11):1227–33.**

7. Bansal V, Mishra A, Verma T, Maini D, Karkhur Y, Maini L. **Anthropometric assessment of tibial resection surface morphology in total knee arthroplasty for tibial component design in Indian population. Journal of Arthroscopy and Joint Surgery 2018; 5(1):24–8.**

8. Dundon JM, Felberbaum DL, Long WJ. **Femoral stem size mismatch in women undergoing total hip arthroplasty. J Orthop 2018; 15(2):293–6.**

9. Boymans TAEJ, Heyligers IC, Grimm B. **The Morphology of the Proximal Femoral Canal Continues to Change in the Very Elderly: Implications for Total Hip Arthroplasty. J Arthroplasty 2015; 30(12):2328–32.**

10. Blom AW, Artz N, Beswick AD, Burston A, Dieppe P, Elvers KT et al. **Improving patients' experience and outcome of total joint replacement: the RESTORE programme. Southampton (UK); 2016.**

11. Nakajima A, Yamada M, Sonobe M, Akatsu Y, Saito M, Yamamoto K et al. **Three-year clinical and radiological results of a cruciate-retaining type of the knee prosthesis with anatomical geometry developed in Japan. BMC Musculoskelet Disord 2021; 22(1):241.**

12. Halawi MJ, Jongbloed W, Baron S, Savoy L, Williams VJ, Cote MP. **Patient Dissatisfaction After Primary Total Joint Arthroplasty: The Patient Perspective. J Arthroplasty 2019; 34(6):1093–6.**

13. Li CY, Ng Cheong Chung KJ, Ali OME, Chung NDH, Li CH. **Literature review of the causes of pain following total knee replacement surgery: prosthesis, inflammation and arthrofibrosis. EFORT Open Rev 2020; 5(9):534–43.**

14. Forster-Horvath C, Egloff C, Valderrabano V, Nowakowski AM. **The painful primary hip replacement - review of the literature. Swiss Med Wkly 2014; 144:w13974.**

15. Benazzo F, Ghiara M, Rossi SMP, Pruneri E, Tiwari V, Perelli S. **Clinical and radiological analysis of a personalized total knee arthroplasty system design. Int Orthop 2019; 43(5):1113–21.**

16. McConnell JS, Syed FA, Saunders P, Kattimani R, Ugwuoke A, Magra M et al. **Definition and validation of a system for classifying the undersized Corail femoral stem in total hip arthroplasty. Hip Int 2021:1120700021996001.**

17. Klasan A, Twiggs JG, Fritsch BA, Miles BP, Heyse TJ, Solomon M et al. **Correlation of tibial component size and rotation with outcomes after total knee arthroplasty. Arch Orthop Trauma Surg 2020; 140(11):1819–24**.

18. Kutzner KP, Freitag T, Bieger R. **Defining 'undersizing' in short-stem total hip arthroplasty: the importance of sufficient contact with the lateral femoral cortex. Hip Int 2020:1120700020940276.**

19. Goyal P, Lau A, Naudie DD, Teeter MG, Lanting BA, Howard JL. **Effect of Acetabular Component Positioning on Functional Outcomes in Primary Total Hip Arthroplasty. J Arthroplasty 2017; 32(3):843–8.**

20. Ueno T, Kabata T, Kajino Y, Ohmori T, Yoshitani J, Tsuchiya H. **Three-Dimensional Host Bone Coverage Required in Total Hip Arthroplasty for Developmental Dysplasia of the Hip and Its Relationship With 2-Dimensional Coverage. J Arthroplasty 2019; 34(1):93–101.**

21. Xu J, Qu X, Li H, Mao Y, Yu D, Zhu Z. **Three-Dimensional Host Bone Coverage in Total Hip Arthroplasty for Crowe Types II and III Developmental Dysplasia of the Hip. J Arthroplasty 2017; 32(4):1374–80.**

22. Whittaker RK, Hexter A, Hothi HS, Panagiotidou A, Bills PJ, Skinner JA et al. **Component size mismatch of metal on metal hip arthroplasty: an avoidable never event. J Arthroplasty 2014; 29(8):1629–34.**

23. Ljung P, Lidgren L, Rydholm U. **Hip socket wear due to component mismatch. A case report. Acta Orthop Scand 1989; 60(2):223–4.**

24. Young SW, Clarke HD, Graves SE, Liu Y-L, Steiger RN de. **Higher Rate of Revision in PFC Sigma Primary Total Knee Arthroplasty With Mismatch of Femoro-Tibial Component Sizes. J Arthroplasty 2015; 30(5):813–7.**

25. Dupraz I, Thorwaechter C, Grupp TM, Woiczinski M, Jansson V, Mueller PE et al. **Impact of tibiofemoral size combinations on knee kinematics after total knee arthroplasty. In: Orthopeadic Proceedings [cited 2021 May 7]. Available from: URL: https://online.boneandjoint.org.uk/doi/abs/10.1302/1358-992X.2020.2.088.**

26. Dai Y, Scuderi GR, Penninger C, Bischoff JE, Rosenberg A. **Increased shape and size offerings of femoral components improve fit during total knee arthroplasty. Knee Surg Sports Traumatol Arthrosc 2014; 22(12):2931–40.**

27. Derias M, Khan MA, Buchanan J. **Digital templating without a calibration marker is accurate at predicting implant size for hip hemiarthroplasty. Eur J Trauma Emerg Surg. 2020 Feb;46(1):115-120. doi: 10.1007/s00068-018-0994-8. Epub 2018 Sep 12. PMID: 30209525.**

28. Derias M, Khan MA, Buchanan J. **Digital templating without a calibration marker is accurate at predicting implant size for hip hemiarthroplasty. Eur J Trauma Emerg Surg. 2020 Feb;46(1):115-120. doi: 10.1007/s00068-018-0994-8. Epub 2018 Sep 12. PMID: 30209525**.

29. Sas A, Pellikaan P, Kolk S, Marty P, Scheerlinck T, van Lenthe GH. **Effect of anatomical variability on stress-shielding induced by short calcar-guided stems: Automated finite element analysis of 90 femora. J Orthop Res. 2019 Mar;37(3):681-688. doi: 10.1002/jor.24240. Epub 2019 Feb 28. PMID: 30720205.**

30. Cristofolini L, Erani P, Bialoblocka-Juszczyk E, Ohashi H, Iida S, Minato I, Viceconti M. **Effect of undersizing on the long-term stability of the Exeter hip stem: A comparative in vitro study. Clin Biomech (Bristol, Avon). 2010 Nov;25(9):899-908. doi: 10.1016/j.clinbiomech.2010.07.003. Epub 2010 Jul 24. PMID: 20659780.**

31. Piriou P, Bugyan H, Casalonga D, Lizée E, Trojani C, Versier G. **Can hip anatomy be reconstructed with femoral components having only one neck morphology? A study on 466 hips. J Arthroplasty. 2013 Aug;28(7):1185-91. doi: 10.1016/j.arth.2012.08.023. Epub 2012 Dec 13. PMID: 23245588**

32. Levine B, Fabi D, Deirmengian C. **Digital templating in primary total hip and knee arthroplasty. Orthopedics. 2010 Nov 2;33(11):797. doi: 10.3928/01477447-20100924-04. PMID: 21053892.**

33. Woo SH, Sung MJ, Park KS, Yoon TR. **Three-dimensional-printing Technology in Hip and Pelvic Surgery: Current Landscape. Hip Pelvis. 2020 Mar;32(1):1-10. doi: 10.5371/hp.2020.32.1.1. Epub 2020 Feb 26. PMID: 32158723; PMCID: PMC7054078.**

34. De Clerico M, Bordini B, Stea S, Viceconti M, Toni A. **Relationship between biometric characteristics and stem size of uncemented hip prostheses. Artif Organs. 2007 Jun;31(6):480-3. doi: 10.1111/j.1525-1594.2007.00410.x. PMID: 17537062.**

35. Blevins JL, Rao V, Chiu YF, Lyman S, Westrich GH. **Predicting implant size in total knee arthroplasty using demographic variables. Bone Joint J. 2020 Jun;102-B(6_Supple_A):85-90. doi: 10.1302/0301-620X.102B6.BJJ-2019-1620.R1. PMID: 32475285.**

36. Halawi MJ, Jongbloed W, Baron S, Savoy L, Williams VJ, Cote MP. **Patient Dissatisfaction After Primary Total Joint Arthroplasty: The Patient Perspective. J Arthroplasty. 2019 Jun;34(6):1093-1096. doi: 10.1016/j.arth.2019.01.075. Epub 2019 Feb 5. PMID: 30799270.**

37. Li CY, Ng Cheong Chung KJ, Ali OME, Chung NDH, Li CH. **Literature review of the causes of pain following total knee replacement surgery: prosthesis, inflammation and arthrofibrosis. EFORT Open Reviews. 2020 Sep;5(9):534-543. DOI: 10.1302/2058-5241.5.200031**.

38. Fong TS, Kim SC, Kim JE, Lee ES, Kim TW, Lee YS. **Morphometric Assessment of Resected Femoral Cut Surface in Korean Knees and Its Fitting With Western-Designed Femoral Components. J Arthroplasty. 2019 Sep;34(9):1929-1937. doi: 10.1016/j.arth.2019.04.061. Epub 2019 May 6. PMID: 31138501.**

39. Della Valle, Alejandro González MD; Padgett, Douglas E. MD; Salvati, Eduardo A. **MD Preoperative Planning for Primary Total Hip Arthroplasty, Journal of the American Academy of Orthopaedic Surgeons: November 2005 - Volume 13 - Issue 7 - p 455-462**

40. Crooijmans HJ, Laumen AM, van Pul C, van Mourik JB. **A new digital preoperative planning method for total hip arthroplasties. Clin Orthop Relat Res. 2009;467(4):909-916. doi:10.1007/s11999-008-0486-y**

41. Zeng Y, Lai OJ, Shen B, et al. **Three-dimensional computerized preoperative planning of total hip arthroplasty with high-riding dislocation developmental dysplasia of the hip. Orthop Surg. 2014;6(2):95-102. doi:10.1111/os.12099**

42. Shimmin AJ, Walter WL, Esposito C. **The influence of the size of the component on the outcome of resurfacing arthroplasty of the hip: a review of the literature. J Bone Joint Surg Br. 2010 Apr;92(4):469-76. doi: 10.1302/0301-620X.92B4.22967. PMID: 20357319.**

43. Schiffner E, Latz D, Jungbluth P, Grassmann JP, Tanner S, Karbowski A, Windolf J, Schneppendahl J. **Is computerised 3D templating more accurate than 2D templating to predict size of components in primary total hip arthroplasty? Hip Int. 2019 May;29(3):270-275. doi: 10.1177/1120700018776311. Epub 2018 May 20. PMID: 29781288.**

44. Crosswell S, Akehurst H, Ramiah R, Navadgi B. **Preoperative sizing of hip hemiarthroplasties to accurately estimate head size from non- standardised pelvic radiographs: Can it be done? Injury. 2019 Nov;50(11):2030-2033. doi: 10.1016/j.injury.2019.08.025. Epub 2019 Aug 17. PMID: 31447209.**

45. Hasler J, Hoch A, Fürnstahl P, Ackermann J, Zingg PO, Vlachopoulos L. **Is the contralateral lesser trochanter a reliable reference for planning of total hip arthroplasty - a 3-dimensional analysis. BMC Musculoskelet Disord. 2021 Mar 11;22(1):268. doi: 10.1186/s12891-021-04131-w. PMID: 33706727; PMCID: PMC7953689.**

46. Nam D, McArthur BA, Cross MB, Pearle AD, Mayman DJ, Haas SB. **Patient-specific instrumentation in total knee arthroplasty: a review. J Knee Surg. 2012 Jul;25(3):213-9. doi: 10.1055/s-0032-1319785. PMID: 23057140.**

47. Kizaki, K., Shanmugaraj, A., Yamashita, F. et al. **Total knee arthroplasty using patient-specific instrumentation for osteoarthritis of the knee: a meta-analysis. BMC Musculoskelet Disord 20, 561 (2019). https://doi.org/10.1186/s12891-019-2940-2**

48. Kayani B, Konan S, Ayuob A, Onochie E, Al-Jabri T, Haddad FS. **Robotic technology in total knee arthroplasty: a systematic review. EFORT Open Rev. 2019 Oct 1;4(10):611-617. doi: 10.1302/2058-5241.4.190022. PMID: 31754467; PMCID: PMC6836078**.

49. Wegrzyn J, Roux JP, Loriau C, Bonin N, Pibarot V. **The tridimensional geometry of the proximal femur should determine the design of cementless femoral stem in total hip arthroplasty. Int Orthop. 2018 Oct;42(10):2329-2334. doi: 10.1007/s00264-018-3843-9. Epub 2018 Feb 22. PMID: 29468265.**

50. Heinert G, Hendricks J, Loeffler MD. **Digital templating in hip replacement with and without radiological markers. J Bone Joint Surg Br. 2009 Apr;91(4):459-62. doi: 10.1302/0301-620X.91B4.21730. PMID: 19336804.**

51. Sariali E, Mauprivez R, Khiami F, Pascal-Mousselard H, Catonné Y. **Accuracy of the preoperative planning for cementless total hip arthroplasty. A randomised comparison between three-dimensional computerised planning and conventional templating. Orthop Traumatol Surg Res. 2012 Apr;98(2):151-8. doi: 10.1016/j.otsr.2011.09.023. Epub 2012 Feb 29. PMID: 22381566.**

52. Mainard D, Barbier O, Knafo Y, Belleville R, Mainard-Simard L, Gross JB. **Accuracy and reproducibility of preoperative three-dimensional planning for total hip arthroplasty using biplanar low-dose radiographs : A pilot study. Orthop Traumatol Surg Res. 2017 Jun;103(4):531-536. doi: 10.1016/j.otsr.2017.03.001. Epub 2017 Mar 18. PMID: 28323248.**

53. Moralidou M, Di Laura A, Henckel J, Hothi H, Hart AJ. **Three-dimensional pre-operative planning of primary hip arthroplasty: a systematic literature review. EFORT Open Rev. 2020 Dec 4;5(12):845-855. doi: 10.1302/2058-5241.5.200046. PMID: 33425373; PMCID: PMC7784139.**

54. Knafo Y, Houfani F, Zaharia B, Egrise F, Clerc-Urmès I, Mainard D. **Value of 3D Preoperative Planning for Primary Total Hip Arthroplasty Based on Biplanar Weightbearing Radiographs. Biomed Res Int. 2019 Mar 10;2019:1932191. doi: 10.1155/2019/1932191. PMID: 30984776; PMCID: PMC6431504.**

RESEARCH TOPIC

MODULARITIES/ INTERFACES

RESEARCH TOPIC 13

Pre-clinical Assessment of the in vivo Behaviour of Non-articulating Interfaces Between Implant Components Concerning the Consequences of Micro Motion or Corrosion Processes – Test Methodology, and Requirements for Testing and Characterization of Implant Modularities

Authors

Jan Philippe Kretzer[1], Thomas M. Grupp[2,3], Christian Kaddick[4], Richard T. Mayer[5], Michael Morlock[6]

[1]Laboratory of Biomechanics & Implant Research, Department of Orthopaedic Surgery, Heidelberg University Hospital, Heidelberg, Germany

[2]Aesculap AG Research & Development, Tuttlingen, Germany

[3]Ludwig Maximilians University Munich, Department of Orthopaedic and Trauma Surgery, Musculoskeletal University Center Munich (MUM), Campus Grosshadern, Munich, Germany

[4]EndoLab Mechanical Engineering GmbH, Riedering, Germany

[5]Johnson & Johnson, Office of the Chief Medical Officer New Brunswick NJ, Zuchwil, Switzerland

[6]Hamburg University of Technology, Hamburg, Germany

1. Abstract

Question: Modularities / interfaces

How can the in vivo behaviour of interfaces between implant components (e.g. head-conus-connection, not articulation partners) be assessed pre-clinically, for example concerning the consequences of micro motion or corrosion processes?

1. Which test methods concerning modularities / corrosion / performance in vivo/in vitro do exist?
2. Which aspects cannot be covered by the current test methods?
3. Which are the requirements for the testing / the characterization of implant modularities and what needs to be considered?

Summary / Recommendation:

Test methods concerning modular interfaces are available (e.g. ASTM F1875) and they are frequently used in short- or long-term applications. By mechanically loading the interfaces, several outcome variables like electrochemical parameters or the amount of lost material are assessed. Furthermore, some studies focus on the analysis of interface micro motions. Nevertheless, in the past, clinically relevant failure modes regarding modular connections were not completely predictable by these testing methods. Current research has pointed out the importance of an appropriate test fluid and the addition of pH lowering additives or corrosion enhancing supplements such as H_2O_2, $FeCl_3$ or HCl are of interest and should be considered in pre-clinical testing. To evaluate the fretting and fatigue performance of modular connections the addition of bovine calf serum to the test fluid should be considered.

Furthermore it appears to be advisable to perform long-term tests and to gain quantitative information about the amount and composition of the degraded material.

2. Level of Evidence

Moderate

3. Consensus Delegate Vote

92% - super majority, strong consensus (92% agree / 3% disagree / 5% abstain)

4. Graphical Abstract

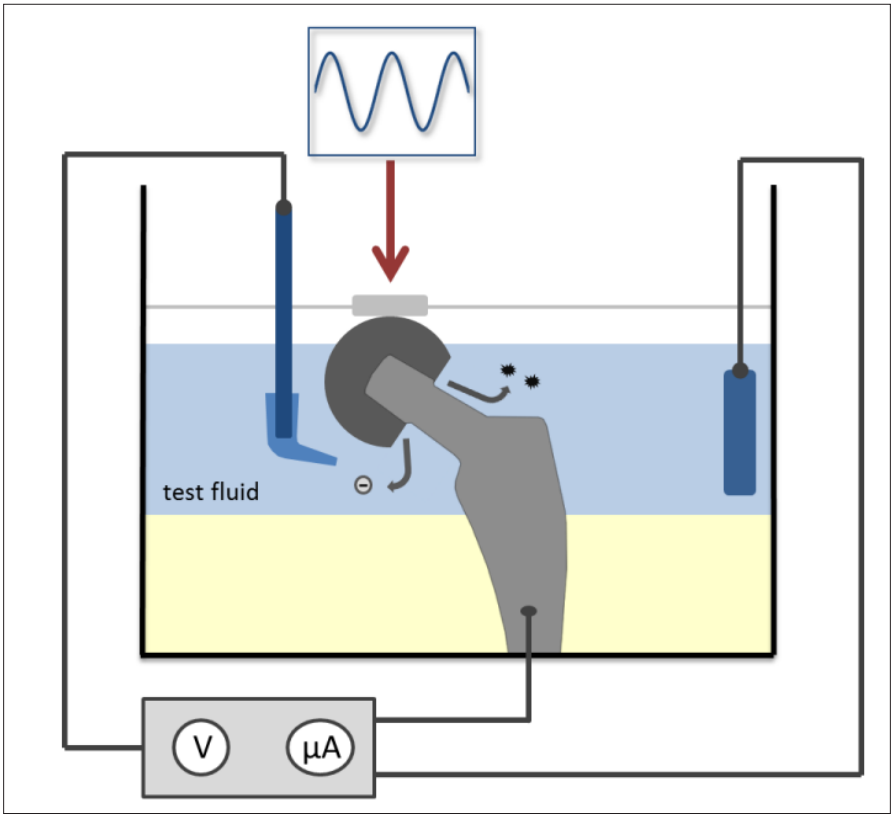


Figure.1: Example of a testing principle to pre-clinically study modular interfaces.

5. Search Strategy

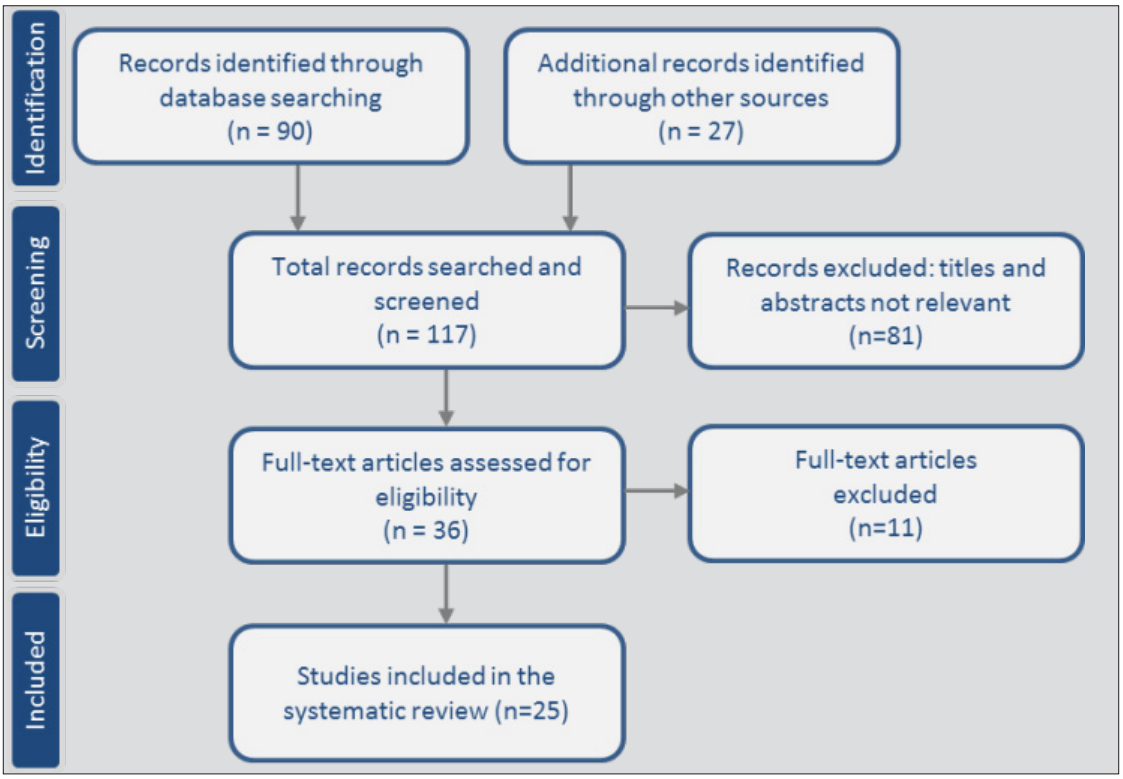
Search platform used: Web of Science

Search terms:

TS=(pre-clinical OR preclinical OR in vitro OR experimental) AND TS=(modular OR taper) AND TS=(corrosion OR fretting OR micro motion) AND TS=(hip OR knee OR arthroplasty OR joint replacement) NOT TS=(finite element OR FEM)

Timeframe: ALL (1920-2021)

Flow chart literature selection process:



6. Rationale

Wear and corrosion at modular connections in arthroplasty have been recognized to cause potential concerns, with multiple clinical reports on adverse local tissue reactions and subsequent early failure of joint replacements (1). Fretting and corrosion at the taper junction are initiated by small oscillating movements (micro motions) at the interfaces leading to distortion of the protective passive layer of the alloy. Starting from mechanically initiated cracks in the passive layer crevice, corrosion and galvanic corrosion may take place and accelerate the corrosive attack on the alloys (2-4).

Modular interfaces may be assessed pre-clinically by dynamic fretting corrosion testing. Herby, the modular connection is immersed in a corrosion promoting fluid and the components are mechanically loaded (Figure 1). Corrosive degradation is typically assessed by determining electrochemical parameters or by measuring the amount of lost material. Furthermore, some studies focus on the amount of interface micro motions, as these are considered to promote or initiate the corrosive attack. Table 1 summarizes the findings of the systematic review on studies that are describing testing methods of modular interfaces in that context. Hereby the testing conditions in terms of applied forces, test frequency and test duration as well as the used test fluid are summarized. Moreover, the outcome parameters are given, that were used to describe or measure the corrosive induced degradation. Mainly head-stem taper connections were studied but also the interface of acetabular cups (5) or modular connections of revision implants (6). The components were loaded with maximum forces ranging between 2.1 and 5.3 kN and most studies applied low frequencies (1-5 Hz). The studies can be divided into short-term (~ a couple of thousand load cycles) and long-term (~ a couple of million load cycles) studies. Regarding the test fluid, most studies used some kind of pH-neutral saline solution (PBS, Ringer, etc.) and some studies applied serum. Only few studies reported the use of fluids with a low pH-value (7-9).

Regarding the assessment of corrosion, the short-term studies mainly focused on electrochemical parameters like the open corrosion potential or the fretting current, whereas long-term studies mainly evaluated the amount of released degradation products by determining the gravimetric or volumetric material loss at the components or the metal ions that were released into the test fluid.

			Test conditions				Outcome parameters							
#	Author(s)	Year	max. Force in N	Frequency in Hz	Duration in cycles C MC = 10^6 cycles	Test fluid	Surface characterization (visual, SEM, etc.)	Subjective scoring	Corrosion current	Corrosion potential	Ion release	Mass or volume material loss	Micromotion	Force to disconnect the connection
1	Viceconti	1996	2100	2	1 MC	a) Ringer b) $FeCl_3$ c) air	X					X		
2	Bhambri	1997	5340	5	10 MC	low pH Ringer	X	X			X			
3	Lambert	1997	± 2.5 Nm (torque)	3 to 5	10MC	dry	X						X	
4	Goldberg	1997	2000	5	1.3 MC	PBS	X		X	X	X			
5	Jani	1997	3303	10	1 MC	Lactated Ringer	X				X			X
6	McLean	1997	4900	10	10 MC	Ringer	X				X	X		X
7	Viceconti	1997	3300	10	5.5 MC / 20MC	Ringer	X					X		
8	Schramm	2000	4000	2	12 MC	NaCl	X							
9	Goldberg	2003	3200	3	1 MC	PBS	X		X	X	X			
10	Hallab	2004	2060	2	1 MC	Serum	X		X	X	X			
11	Mroczkowski	2006	5340	3	18900 C	Ringer			X	X				
12	Gilbert	2009	3300	3	1 MC	PBS	X		X	X			X	
13	Kretzer	2009	2300	3 and 15	10 MC	Serum	X				X			
14	Jauch	2011	2300	1	10000 C	dry (clean vs. contaminated)	X						X	
15	Panagiotidou	2013	3100	4	10 MC	PBS	X		X	X				
16	Mali	2015	4000	3	25 incrementals of 540 cycles: 13500 C	PBS			X				X	X
17	Nambu	2015	3300	10	5 MC / 10 MC	Acidified solution: 0.9 % saline plus HCl: pH 1.5	X					X		X
18	Panagiotidou	2015	2300	3	1000 C	10% bovine calf serum diluted with PBS			X					
19	Swaminathan	2015	3200	3	1 MC	PBS			X	X			X	
20	Panagiotidou	2017	2300	3	1000 C / 5 MC	10% bovine calf serum diluted with PBS	X		X			X		
21	Baxmann	2017	3800	15	10 MC	a) Ringer b) Bovine serum c) $FeCl_3$	X				X			
22	Haschke	2019	5300	1	n/a	n/a							X	X
23	Ouelette	2019	4000	3	25 incrementals of 540 cycles: 13500 C	PBS			X				X	
24	Falkenberg	2020	5000	quasistatic		n/a							X	
25	Woon	2020	3400	3	22 incrementals of 540 cycles: 11880 C	PBS			X					

Table 1: Systematically analyzed studies on corrosion testing of modular interfaces (5-29).

In 1998, the ASTM released the ASTM F1875 standard that serves as a standard practice for fretting corrosion testing of modular implant interfaces. Two methods are part of this standard, a quantitative analysis (method I) and a qualitative assessment (method IIa/b).

According to method I, the test samples should be loaded with a maximum force of 3.3 kN for at least $10x10^6$ load cycles at a frequency of 5 Hz and a 0.9 % NaCl solution containing proteins (10% calf serum) is used as lubricant. Material degradation is evaluated by analyzing the released particles or ions.

According to method II the test samples are loaded with a maximum force of 2.0 kN for 3600 load cycles at a frequency of 1 Hz and a 0.9 % NaCl solution is used. Corrosion is characterized by electrochemical evaluation (measurement of corrosion potential IIa or corrosion peak current IIb).

Until approx. 2008-2010, implant modularity was only weakly associated with real clinical problems and material degradation was mainly reported as phenomenon or some single case reports addressed clinical issues (1, 30). With the introduction of modular necks and large metal heads in THA, this situation has dramatically changed. Several studies identified the modular connection as a source for corrosive degradation and the release of degradation products (e.g. metal ions, particles, metalorganics, etc.). A number of studies quantified the amount of degraded material at the modular connection of clinically failed implants (Table 2).

#	Author(s)	Year	Number of analysed retrievals	Type of Implant	Amount of degradation / wear in mm^3/year mean or (median)
1	Langton	2012	63	ASR	0,44
2	Langton	2012	43	Articuleze	0,13
3	Nassif	2013	8	Taper Type 1	0,5*
4	Nassif	2013	6	Taper 11/13	1,8*
5	Nassif	2013	26	Taper 12/14	0,7*
6	Matthies	2013	110	Diverse LH-MoM	0,85
7	Bishop	2013	2	Pinnacle	0,97
8	Bishop	2013	3	Durom LDH	2,8
9	Brock	2015	72	Corail	0,13* (median)
10	Brock	2015	32	SROM	0,08* (median)
11	Hothi	2015	10	Corail	0,24 (median)
12	Hothi	2015	10	SROM	0,13 (median)
13	Whittaker	2017	50	Corail	0,27 (median)

*Table 2: Comparison of various studies on annual taper wear (3, 31-36). *- values taken from diagram.*

If the studies with larger numbers of cases (n>10) are taken into account, the mean annual taper degradation volumes vary between 0.08 mm3 and 0.85 mm3. These values appear to be low if they are compared to the articular wear of metal-on-metal bearings, but corrosive degradation products have been considered to be more aggressive compared to the articular wear products (3, 37, 38).

In addition to the studies listed in Table 2, there are two studies in which taper wear was quantified for metal-polyethylene bearing without reporting taper wear associated clinical problems (39, 40). These studies reported annual taper wear rates ranging from 0.03 mm^3/year to 0.05 mm^3/year. If the study results on taper wear summarized here (Table 2) are evaluated, it is difficult to differentiate between

uncritical taper wear and problematic taper wear. However, it can be stated that numerous studies describe mean annual taper wear rates in the range of approximately 0.08 mm^3 to 0.85 mm^3, where metal-associated reasons for revision are increasingly reported. In contrast, metal-associated problems seem to play a minor role when taper wear is lower.

Retrospectively, it can clearly be stated that the existing methods to pre-clinically evaluate modular interfaces were not able to predict the aforementioned severe clinical problems related to modular interfaces.

Different reasons may be considered regarding this aspect. The mechanical loading conditions applied in the test methods used (Table 1 and ASTM F1875) are mostly close to the in-vivo loading conditions in terms of magnitude and frequency (41) and the long-term tests also represent several years of service in the patient (42). Nevertheless, time depending degradation processes are hard to be simulated pre-clinically.

Moreover, recent research has pointed out the importance of the test fluid to mimic the liquid in the joint (7, 9, 43-46). The joint fluid in the human body might be more aggressive in comparison to the mostly used saline based solutions in recent testing methods (Table 1, ASTM F1875). Consequently, the test fluid in pre-clinical studies might be altered to a corrosion promoting fluid. For example in retrieved implants it has been noticed that the pH value in the crevice may turn to be low and highly acidic (9). Such a reduced pH is considered to be related to crevice corrosion processes, which generate hydrogen ions that dissolve in the fluid within the crevice. Further retrieval studies have highlighted a potential inflammatory cell-induced corrosion due to the formation of reactive oxygen species, specifically hydrogen peroxide (H_2O_2) (46). Regarding the fretting and fatigue performance of modular titanium-titanium connections a solution that contains bovine calf serum has been proposed as a promising test fluid (7).

Consequently, current research is focusing on the development of appropriate test fluids and the addition of pH lowering additives or corrosion enhancing supplements. Additives like H_2O_2, $FeCl_3$ or HCl are of interest and should be considered in pre-clinical testing of modular interfaces. It should be noted that the ASTM 1875 is currently being revised. Therefore, regular monitoring of the standards and scientific data is recommended.

Micro motion analysis, as conducted in some studies (Table 1), might be an interesting parameter, to directly compare the fixation properties of modular interfaces. Nevertheless, the applied analytical methods vary and the magnitude of measured micro motions highly depends on the analytical method. Moreover, the corrosive induced material degradation is influenced by several further aspects like the crevice shape and volume or the surrounding fluid.

Regarding the aforementioned considerations, pre-clinical testing of modular interfaces appears to be essential. Short time tests may help to directly compare different designs, surface treatments etc. but they do not provide sufficient information regarding the amount of degraded material and its potential clinical relevance. To achieve such information it seems to be advisable to perform long term tests and to gain quantitative information about the amount and composition of the degraded material. It needs to be noticed that the currently applied testing methods appear to be improvable in particular regarding the application of appropriate testing fluids.

7. References

1. Esposito CI, Wright TM, Goodman SB, Berry DJ. **What is the trouble with trunnions? Clin Orthop Relat Res. 2014;472(12):3652-8.**

2. Goldberg JR, Gilbert JL, Jacobs JJ, Bauer TW, Paprosky W, Leurgans S. **A multicenter retrieval study of the taper interfaces of modular hip prostheses. Clin Orthop Relat Res. 2002(401):149-61.**

3. Matthies AK, Racasan R, Bills P, Blunt L, Cro S, Panagiotidou A, et al. **Material loss at the taper junction of retrieved large head metal-on-metal total hip replacements. Journal of orthopaedic research : official publication of the Orthopaedic Research Society. 2013;31(11):1677-85.**

4. Jacobs JJ, Gilbert JL, Urban RM. **Corrosion of metal orthopaedic implants. J Bone Joint Surg Am. 1998;80(2):268-82.**

5. Lambert RD, McLean TW. **Test Method comparing Torsional Fatigue of Modular Acetabular Components. Modularity of Orthopedic Implants - STP1301: ASTM International; 1997.**

6. Schramm M, Wirtz DC, Holzwarth U, Pitto RP. **The morse taper junction in modular revision hip replacement - A biomechanical and retrieval analysis. Biomedizinische Technik. 2000;45(4):105-9.**

7. Baxmann M, Pfaff AM, Schilling C, Grupp TM, Morlock MM. **Biomechanical Evaluation of the Fatigue Performance, the Taper Corrosion and the Metal Ion Release of a Dual Taper Hip Prosthesis under Physiological Environmental Conditions. Biotribology. 2017;12:1-7.**

8. Bhambri SK, Gilbertson L. **Fretting corrosion fatigue study of modular joints in total hip replacements by accelerated in vitro testing. Modularity of orthopedic implants - STP1301: ASTM International; 1997.**

9. Nambu S, Obert R, Roark M, Delvecchio E, Linton D, Bible S, et al. **Accelerated fretting corrosion testing modular necks for total hip arthroplasty. In: Greenwald SA, Kurtz SM, Lemons JE, Mihalko WM, editors. Modularity and Tapers in Total Joint Replacement Devices: ASTM International; 2015. p. 22.**

10. Viceconti M, Ruggeri O, Toni A, Giunti A. **Design-related fretting wear in modular neck hip prosthesis. J Biomed Mater Res. 1996;30(2):181-6.**

11. Goldberg JR, Buckley CA, Jacobs JJ, Gilbert JL. **Corrosion testing of modular hip implants. Modularity of Orthopedic Implants - STP1301: ASTM International; 1997. p. 157-76.**

12. Jani SC, Sauer WL, McLean TW, Lambert RD, Kovacs P. **Fretting corrosion mechanisms at modular implant interfaces. Modularity of Orthopedic Implants - STP1301: ASTM International; 1997.**

13. McLean TW, Lambert RD. **Test method for the assessment of frettimg at the femoral head and neck taper interface Modularity of Orthopedic Implants - STP1301: ASTM International; 1997. p. 104-13.**

14. Viceconti M, Baleani M, Squarzoni S, Toni A. **Fretting wear in a modular neck hip prosthesis. Journal of Biomedical Materials Research. 1997;35(2):207-16.**

15. Goldberg JR, Gilbert JL. **In vitro corrosion testing of modular hip tapers. Journal of Biomedical Materials Research Part B-Applied Biomaterials. 2003;64B(2):78-93.**

16. Hallab NJ, Messina C, Skipor A, Jacobs JJ. **Differences in the fretting corrosion of metal-metal and ceramic-metal modular junctions of total hip replacements. Journal of Orthopaedic Research. 2004;22(2):250-9.**

17. Mroczkowski ML, Hertzler JS, Humphrey SM, Johnson T, Blanchard CR. **Effect of impact assembly on the fretting corrosion of modular hip tapers. Journal of orthopaedic research : official publication of the Orthopaedic Research Society. 2006;24(2):271-9.**

18. Gilbert JL, Mehta M, Pinder B. **Fretting Crevice Corrosion of Stainless Steel Stem-CoCr Femoral Head Connections: Comparisons of Materials, Initial Moisture, and Offset Length. Journal of Biomedical Materials Research Part B-Applied Biomaterials. 2009;88B(1):162-73.**

19. Kretzer JP, Jakubowitz E, Krachler M, Thomsen M, Heisel C. **Metal release and corrosion effects of modular neck total hip arthroplasty. International Orthopaedics. 2009;33(6):1531-6.**

20. Jauch SY, Huber G, Hoenig E, Baxmann M, Grupp TM, Morlock MM. **Influence of material coupling and assembly condition on the magnitude of micromotion at the stem-neck interface of a modular hip endoprosthesis. Journal of biomechanics. 2011;44(9):1747-51.**

21. Panagiotidou A, Meswania J, Hua J, Muirhead-Allwood S, Hart A, Blunn G. **Enhanced Wear and Corrosion in Modular Tapers in Total Hip Replacement Is Associated With the Contact Area and Surface Topography. Journal of Orthopaedic Research. 2013;31(12):2032-9.**

22. Mali SA, Gilbert JL. **Correlating fretting corrosion and micromotions in modular tapers: test method development and assessment. Modularity and tapers in total joint replacement devices -STP 1591, ASTM International 2015;STP 11591:259-82.**

23. Panagiotidou A, Meswania J, Osman K, Bolland B, Latham J, Skinner J, et al. **The effect of frictional torque and bending moment on corrosion at the taper interface : an in vitro study. The bone & joint journal. 2015;97-B(4):463-72.**

24. Swaminathan V, Scholl L, Lee R, Faizan A, Thakore M, TenHuisen K, et al. **Simultaneous Hip Head-Stem Taper Junction Measurements of Electrochemical Corrosion and Micromotion: A Comparison of Taper Geometry and Stem Material. In: Greenwald SA, Kurtz SM, Lemons JE, Mihalko WM, editors. Modularity and Tapers in Total Joint Replacement Devices: ASTM International; 2015. p. 15.**

25. Panagiotidou A, Cobb T, Meswania J, Skinner J, Hart A, Haddad F, et al. **Effect of impact assembly on the interface deformation and fretting corrosion of modular hip tapers: An in vitro study. Journal of orthopaedic research : official publication of the Orthopaedic Research Society. 2017.**

26. Haschke H, Konow T, Huber G, Morlock MM. **Influence of flexural rigidity on micromotion at the head-stem taper interface of modular hip prostheses. Med Eng Phys. 2019;68:1-10.**

27. Ouellette ES, Mali SA, Kim J, Grostefon J, Gilbert JL. **Design, Material, and Seating Load Effects on In Vitro Fretting Corrosion Performance of Modular Head-Neck Tapers. The Journal of Arthroplasty. 2019;34(5):991-1002.**

28. Falkenberg A, Biller S, Morlock MM, Huber G. **Micromotion at the head-stem taper junction of total hip prostheses is influenced by prosthesis design-, patient- and surgeon-related factors. Journal of biomechanics. 2020;98:109424.**

29. Woon CYL, Wach A, Wright TM, Padgett DE. **Taper Design, Head Material, and Manufacturer Affect the Onset of Fretting Under Simulated Corrosion Conditions. The Journal of Arthroplasty. 2020;35(4):1117-22.**

30. Morlock MM. **The taper disaster--how could it happen? Hip international : the journal of clinical and experimental research on hip pathology and therapy. 2015;25(4):339-46.**

31. Langton DJ, Sidaginamale R, Lord JK, Nargol AVF, Joyce TJ. **Taper junction failure in large-diameter metal-on-metal bearings. Bone and Joint Research. 2012;1(4):56-63.**

32. Nassif NA, Nawabi DH, Stoner K, Elpers M, Wright T, Padgett DE. **Taper Design Affects Failure of Large-head Metal-on-metal Total Hip Replacements. Clinical orthopaedics and related research. 2013.**

33. Bishop N, Witt F, Pourzal R, Fischer A, Rutschi M, Michel M, et al. **Wear patterns of taper connections in retrieved large diameter metal-on-metal bearings. Journal of orthopaedic research : official publication of the Orthopaedic Research Society. 2013;31(7):1116-22.**

34. Brock TM, Sidaginamale R, Rushton S, Nargol AV, Bowsher JG, Savisaar C, et al. **Shorter, rough trunnion surfaces are associated with higher taper wear rates than longer, smooth trunnion surfaces in a contemporary large head metal-on-metal total hip arthroplasty system. Journal of orthopaedic research : official publication of the Orthopaedic Research Society. 2015;33(12):1868-74.**

35. Hothi HS, Whittaker RK, Meswania JM, Blunn GW, Skinner JA, Hart AJ. **Influence of stem type on material loss at the metal-on-metal pinnacle taper junction. Proceedings of the Institution of Mechanical Engineers Part H, Journal of engineering in medicine. 2015;229(1):91-7.**

36. Whittaker RK, Hothi HS, Eskelinen A, Blunn GW, Skinner JA, Hart AJ. **Variation in taper surface roughness for a single design effects the wear rate in total hip arthroplasty. Journal of orthopaedic research : official publication of the Orthopaedic Research Society. 2017;35(8):1784-92.**

37. Esposito CI, Wright TM, Goodman SB, Berry DJ. **What is the Trouble With Trunnions? Clin Orthop Relat Res. 2014.**

38. Langton D, Sidaginamale R, Lord J, Joyce T, Natu S, Nargol A. **Metal debris release from taper junctions appears to have a greater clinical impact than debris released from metal on metal bearing surfaces. Bone & Joint Journal Orthopaedic Proceedings Supplement. 2013;95-B(SUPP 1):28-.**

39. Langton DJ, Wells SR, Joyce TJ, Bowsher JG, Deehan D, Green S, et al. **Material loss at the femoral head taper. The bone & joint journal. 2018;100-b(10):1310-9.**

40. Hothi HS, Eskelinen AP, Henckel J, Kwon YM, Blunn GW, Skinner JA, et al. **Effect of Bearing Type on Taper Material Loss in Hips From 1 Manufacturer. J Arthroplasty. 2018;33(5):1588-93.**

41. Bergmann G, Graichen F, Rohlmann A, Bender A, Heinlein B, Duda GN, et al. **Realistic loads for testing hip implants. Biomed Mater Eng. 2010;20(2):65-75.**

42. Kinkel S, Wollmerstedt N, Kleinhans JA, Hendrich C, Heisel C. **Patient activity after total hip arthroplasty declines with advancing age. Clin Orthop Relat Res. 2009;467(8):2053-8.**

43. Ocran EK, Guenther LE, Brandt JM, Wyss U, Ojo OA. **Corrosion and Fretting Corrosion Studies of Medical Grade CoCrMo Alloy in a Clinically Relevant Simulated Body Fluid Environment. Metallurgical and Materials Transactions a-Physical Metallurgy and Materials Science. 2015;46A(6):2696-709.**

44. Royhman D, Patel M, Runa MJ, Jacobs JJ, Hallab NJ, Wimmer MA, et al. **Fretting-corrosion in Hip Implant Modular Junctions: New Experimental Set-up and Initial Outcome. Tribol Int. 2015;91:235-45.**

45. Royhman D, Pourzal R, Hall D, Lundberg HJ, Wimmer MA, Jacobs J, et al. **Fretting-corrosion in hip taper modular junctions: The influence of topography and pH levels – An in-vitro study. Journal of the Mechanical Behavior of Biomedical Materials. 2021;118:104443.**

46. Gilbert JL, Sivan S, Liu Y, Kocagöz SB, Arnholt CM, Kurtz SM. **Direct in vivo inflammatory cell-induced corrosion of CoCrMo alloy orthopedic implant surfaces. Journal of biomedical materials research Part A. 2015;103(1):211-23.**

RESEARCH TOPIC 14

Clinical Assessment of the in vivo Behaviour of Interfaces Between Non-Articulating Implant Components Concerning the Consequences of Micro Motion or Corrosion Processes

Authors

Francesco Traina[1,2], Michael Morlock[3], Richard T. Mayer[4], Alister Hart[5]

[1]Ortopedia-Traumatologia e Chirurgia protesica e dei reimpianti d'anca e di ginocchio, IRCCS Istituto Ortopedico Rizzoli, Bologna, Italy

[2]DIBINEM, Università di Bologna, Bologna, Italy

[3]Hamburg University of Technology, Hamburg, Germany

[4]Johnson & Johnson, Office of the Chief Medical Officer New Brunswick NJ, Zuchwil, Switzerland

[5]London Implant Retrieval Centre, Royal National Orthopaedic Hospital, Stanmore, UK

1. Abstract

Question: Modularities / Interfaces

How can the in vivo behaviour of interfaces between implant components (e.g. head-conus-connection, not articulation partners) be assessed clinically, for example concerning the consequences of micro motion or corrosion processes?

Summary / Recommendation:

The in vivo assessment of the behaviour of interfaces, like the clinical consequences of micro-motion or corrosion processes, includes an appropriate patient/implant stratification, plain X-rays, serum metal ion levels and cross-sectional imaging (MARS MRI). A patient-implant stratification is mandatory to predict the cases at higher risk of complications, when close surveillance is appropriate.

In vivo clinical manifestations of interface damages are adverse local tissue reactions (Cr-Co alloy modularity in single or dual taper implants) or catastrophic failures (Ti alloys in dual taper implants). The cases at high risk encompass: active overweight male patients, high offset, 36 mm heads or larger. In these cases, a simple X-ray should be performed and infection ruled out (inflammation markers and synovial fluid aspiration). Second level investigations include serum metal ion levels in Cr Co alloys (single taper: cobalt level >1 ng/ml (1 ppb) and the Co/Cr ratio >2/ dual taper: thresholds can be as high as 2.8 and 3.8). Ti blood ion assessment in dual taper Ti implants is less standardized probably not useful. Cross sectional imaging (MARS MRI, or CT and US as second choices) may identify adverse local tissue reactions, usually >3mm thick walled cystic/solid masses.

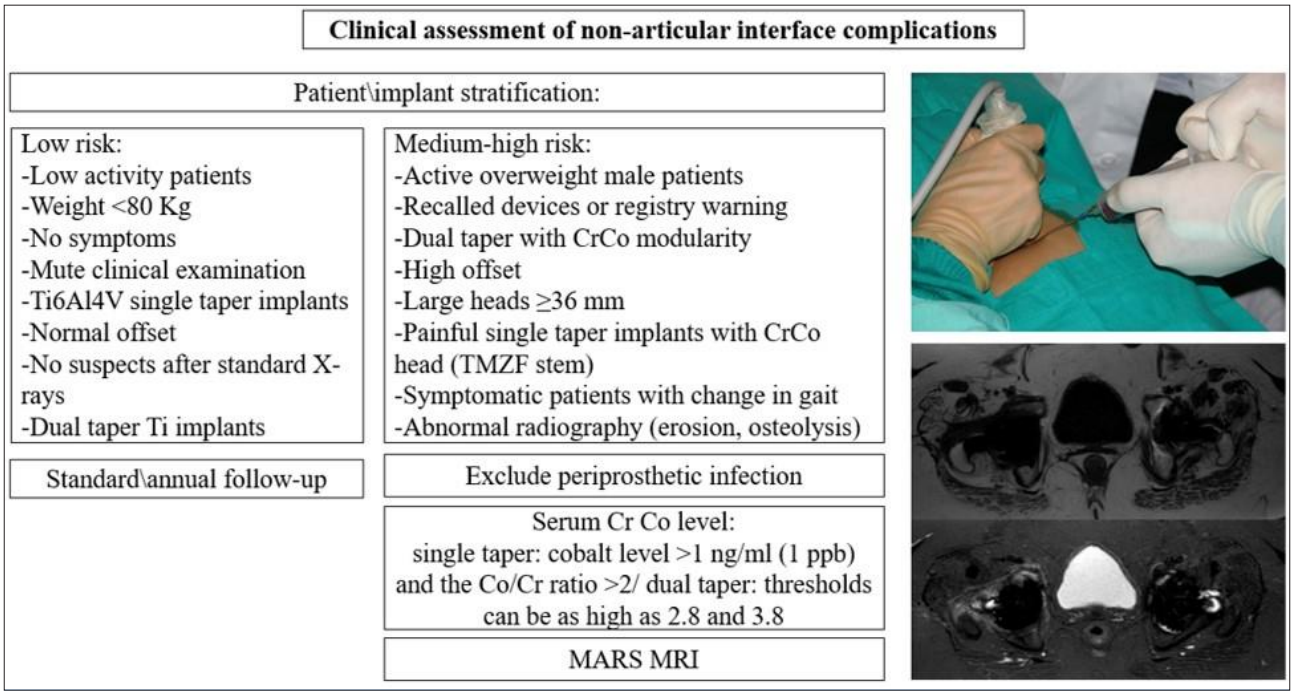
2. Level of Evidence

Moderate/Poor

3. Consensus Delegate Vote

84% - super majority, strong consensus (84% agree / 5% disagree / 11% abstain)

4. Graphical Abstract



5. Search Strategy

Not applicable

6. Rationale

Non articular interfaces in arthroplasty, while providing more choice and more modularity, promoting more "customizable" implants, are inevitably prone to some adjunctive failure modalities, which can be a relevant reason for revision and a cause of severe periprosthetic tissue damage (1-3). To date, two main failure modalities have been identified: mechanically assisted crevice corrosion and cantilever bending motion (3). These two failure modalities are strictly dependent on the involved metal alloys. CrCo alloys may cause corrosion ending in adverse local tissue reactions (ALTR), with a spectrum of MOM-like damages: these events may take place in single or dual taper implants (1,3,4). Ti alloys have a micromotion that in some cases may ultimately lead to catastrophic failures in dual taper implants (1,3,4).

Clinical assessment of non-articular interface complications can be important to anticipate the local damage, making the revision surgery potentially simpler and more effective (3). Considering that many failures are reported in short/mid-terms, appropriate evaluations should be made soon after the implantation (1-4).

The first step to take in the clinical assessment of non-articular interface complications is patient/implant stratification, with the aim to identify the cases at higher risk for non articular interface complications (4,5).

Very low risk patients/implants are:

- Low activity patients
- Weight under 80 Kg
- No symptoms
- Mute clinical examination
- Ti6Al4V single taper implants with 32 mm ceramic head or smaller
- Normal offset
- No suspects after standard radiographic follow-up

Ti6Al4V single taper implants with 36 mm ceramic head or larger appear low risk devices (5).

Dual taper Ti implants are considered low risk implants (4). However, dual taper Ti implants in active overweight male patients with high lever arms are implants at risk for catastrophic failure (6,7). In these cases, some Authors discourage the use of such implants (7). While it is difficult to provide a percentage of risk, as the catastrophic failures are also dependent on implant design, it is likely that less than 1% of the dual taper Ti implants may fail due to interface complications (junction breakage) (1). To date there is no ascertained way to predict the failure of dual taper Ti implants. In these cases, it seems that to date, nothing more than annual follow-up can be recommended (4). The only recommendation is to avoid implantation of dual taper Ti implants in patients at higher risk (7). Unfortunately, the blood and serum Ti levels cannot provide any help to clinicians, as normal ranges are not determined and there is a substantial lack of methodological standardization for every kind of hip arthroplasty (8-11).

Patient-implant stratification for medium/high risk cases (4,5,12):

- Active overweight male patients
- Dual taper implants with recalled devices (or registry warnings)
- Dual taper with CrCo modularity
- High offset
- Large heads >=36 mm
- Painful single taper implants with CrCo head (in particular V40 taper in TMZF stem)
- Symptomatic patients with change in gait
- Abnormal radiography (erosion, osteolysis)

In these cases, more strict monitoring should be considered (at least 6 months) (4). In case of symptomatic patients with abnormal radiographic findings, periprosthetic infection work-up is the first step to take: ESR, CRP ± synovial fluid analysis (4). The results may be misleading: CPR false positives are up to 30% in ALTR, synovial fluid analysis is not fully reliable in ALTR (13,14).

After ruling out infection, second level investigations are recommended. It should be clearly stated that, due to the complexity of these cases and the substantial lack of a very consistent literature, overreliance on any investigative tool should be discouraged (5,15-17).

First, the serum metal ion levels should be assessed (16,18,19). The proposed thresholds are: cobalt level >1 ng/ml (1 ppb) and Co/Cr ratio >2 (Sensibility-Specificity: 95%-94%; 83%-72% for single taper corrosion) (5,18). For dual taper Cr Co implants, the thresholds can be as high as 2.8 and 3.8, respectively (achieving lower sensibility and specificity, though) (15-17). Higher serum metal ion levels usually correlate with pseudotumors, but not with symptoms (15). As the Cr Co thresholds are low, environmental exposition may cause false positives and should be investigated and considered. Plus, using such low thresholds for Cr Co devices may be questionable: in the SCENIHR report about MOM implants, a serum Co threshold lower than 7 µg/L improved the sensibility, but deeply affected the specificity (20). Moreover, the substantial lack of standardization in metal ion assessment may furtherly discredit the validity of low thresholds.

Cross sectional imaging is the other second level investigation which should be performed in case of elevated serum ion levels (5,15,16,19). The ion-level based approach may be cost-effective, but may miss some cases with ALTR/pseudotumor, at least, as the natural history of ALTR is still largely unknown (17). Another logical approach would be a universal monitoring with a cross sectional imaging at two years after implantation, in all medium/high risk cases: the best surveillance has not been determined (17). MARS MRI is the best cross sectional imaging (lower quality alternatives are US and CT) (5,15,19). 3D multispectral imaging techniques as SEMAC MAVRIC advanced WARP are other promising MRI tools, but to date the use is limited and not standardized and no recommendations can be made (5). Pseudotumors can occur in up to 36% of dual taper CrCo implants at short terms (15). The most common findings in ALTR due to taper corrosion are >3mm thick walled cystic masses (around half of the cases); secondly, another finding can be solid masses (15,19). These findings are different from the MOM implants, which usually have cystic masses with <3 mm thickness, in more locations (18). However, to date, there is no validated scoring system for symptomatic ALTR and some variability still exists in MARS MRI protocols: thus, although MARS MRI is an highly sensitive tool, there are some notable limits (19).

The role of synovial fluid aspiration, with Cr Co analysis, has been used in MOM implants, but to date there is no study about non articular interfaces and no thresholds can be determined (14). Pre-operative periprosthetic biopsies to detect ALTR may be a solution in doubtful cases, but no study has been conducted so far: moreover, there is no correlation between local tissue response to metal debris and clinical problems possibly leading to revision (20). Differently from MOM bearings, dual taper implants released Cr_2O_3 and amorphous TiO_2 in laboratory setting, identifying two possible markers for metal ion analysis in periprosthetic tissues (21). However, there is no clinical validation to date.

To date, the rise of dual mobility implants introduces the issue of non-articular metal ion release from cup-liner modularity, possibly leading to ALTR. This possible source of metal ion was identified even in MOM implants with cup-liner modularity (22). The Cr Co release seems of minor concern in dual mobility implants with cup liner modularity. A mild elevation of blood Co level was detected in less than 20% of the patients, when ceramic head was implanted, no metal ion elevation was observed (23,24). While no recommendations can be made about modular dual mobility cups, to date the behaviour of cup-liner interfaces seems not worrisome.

In summary, the clinical assessment of the behaviour of non-articular interfaces is still poorly described, many of the investigations and their thresholds are based on very few qualitative studies.

While clinical assessment of non-articular interfaces is less defined, a few recommendations can be made about the use and the choice of single-taper and dual taper implants. The retrieval studies confirmed that Cr Co devices may cause metal ion release and ALTR, similarly to MOM implants (3,4,25). Mixing metal alloys may trigger mechanically assisted crevice corrosion and should be avoided: the Australian registry showed a higher revision rate for Cr Co exchangeable necks when compared to Ti necks (1,26). On the other side, while dual taper Ti junctions seem scarcely involved in metallosis and ALTR, these devices may break in active overweight patients with long lever arms (6,7). Two large registry studies advised against the routinary use of dual taper implants in osteoarthritis (1,2).

Other few recommendations can be made about the surgical technique. To date, it seems that no clinical study can support any technical tips/technique to assemble tapers. Moreover, to our knowledge there is no validated device/tool to assess the quality of the intraoperative assembly technique. The only way to assess the correct assembly is the surgeon visual inspection (thus inherently subjective). Most of the evidence about the correct surgical technique came from biomechanical studies and manufacturer's brochures (27-29). A correct matching is paramount: taper and bore should come from the same manufacturer (27,28). Cleaning and drying the taper (and the bore in particular) is universally recognized as a fundamental passage in order to provide a correct assembly (5-7,27-29). Moreover, the impaction force has been advocated as another important factor affecting the quality of the assembly: a minimum force of 4000 N should be exerted, using a 500 gr hammer on a specific impactor aligned with the axis of the taper (28,30). Ceramtec provided also some practical guidelines for ceramic head implantation summarizing the main advises (taper protection, clean and dry, impaction with moderate force, single or multiple hammer blows) (31).

7. References

1. Graves SE, de Steiger R, Davidson D, Donnelly W, Rainbird S, Lorimer MF, Cashman KS, Vial RJ. **The use of femoral stems with exchangeable necks in primary total hip arthroplasty increases the rate of revision. Bone Joint J. 2017 Jun;99-B(6):766-773. doi: 10.1302/0301-620X.99B6.38020.**

2. Colas S, Allalou A, Poichotte A, Piriou P, Dray-Spira R, Zureik M. **Exchangeable Femoral Neck (Dual-Modular) THA Prostheses Have Poorer Survivorship Than Other Designs: A Nationwide Cohort of 324,108 Patients. Clin Orthop Relat Res. 2017 Aug;475(8):2046-2059. doi: 10.1007/s11999-017-5260-6. Epub 2017 Feb 13.**

3. Srinivasan A, Jung E, Levine BR. **Modularity of the femoral component in total hip arthroplasty. J Am Acad Orthop Surg. 2012 Apr;20(4):214-22. doi: 10.5435/JAAOS-20-04-214.**

4. Kwon YM, Fehring TK, Lombardi AV, Barnes CL, Cabanela ME, Jacobs JJ. **Risk stratification algorithm for management of patients with dual modular taper total hip arthroplasty: consensus statement of the American Association of Hip and Knee Surgeons, the American Academy of Orthopaedic Surgeons and the Hip Society. J Arthroplasty. 2014 Nov;29(11):2060-4. doi: 10.1016/j.arth.2014.07.029.**

5. Fitz D, Klemt C, Chen W, Xiong L, Yeo I, Kwon YM. **Head-Neck Taper Corrosion in Metal-on-Polyethylene Total Hip Arthroplasty: Risk Factors, Clinical Evaluation, and Treatment of Adverse Local Tissue Reactions. J Am Acad Orthop Surg. 2020 Nov 15;28(22):907-913. doi: 10.5435/JAAOS-D-20-00475.**

6. Fokter SK, Molčnik A, Kavalar R, Pelicon P, Rudolf R, Gubeljak N. **Why do some titanium-alloy total hip arthroplasty modular necks fail? J Mech Behav Biomed Mater. 2017 May;69:107-114. doi: 10.1016/j.jmbbm.2016.12.012.**

7. Grupp TM, Weik T, Bloemer W, Knaebel HP. **Modular titanium alloy neck adapter failures in hip replacement--failure mode analysis and influence of implant material. BMC Musculoskelet Disord. 2010 Jan 4;11:3. doi: 10.1186/1471-2474-11-3.**

8 . Swiatkowska I, Martin NG, Henckel J, Apthorp H, Hamshere J, Hart AJ. **Blood and plasma titanium levels associated with well-functioning hip implants. J Trace Elem Med Biol. 2020 Jan;57:9-17. doi: 10.1016/j.jtemb.2019.09.005.**

9. Yao JJ, Lewallen EA, Thaler R, Dudakovic A, Wermers M, Day P, Eckdahl S, Jannetto P, Bornhorst JA, Larson AN, Abdel MP, Lewallen DG, van Wijnen AJ. **Challenges in the Measurement and Interpretation of Serum Titanium Concentrations. Biol Trace Elem Res. 2020 Jul;196(1):20-26. doi: 10.1007/s12011-019-01891-4.**

10. Omlor GW, Kretzer JP, Reinders J, Streit MR, Bruckner T, Gotterbarm T, Aldinger PR, Merle C. **In vivo serum titanium ion levels following modular neck total hip arthroplasty--10 year results in 67 patients. Acta Biomater. 2013 Apr;9(4):6278-82. doi: 10.1016/j.actbio.2012.12.001.**

11. Fokter SK, Zajc J, Merc M. **Interchangeable neck failures of bi-modular femoral stems in primary total hip arthroplasty cannot be predicted from serum trace element analysis. Int Orthop. 2021 Apr;45(4):877-881. doi: 10.1007/s00264-020-04812-6.**

12. Kwon YM, An S, Yeo I, Tirumala V, Chen W, Klemt C. **Radiographic Risk Factors Associated With Adverse Local Tissue Reaction in Head-Neck Taper Corrosion of Primary Metal-on-Polyethylene Total Hip Arthroplasty. J Am Acad Orthop Surg. 2021 Apr 15;29(8):353-360. doi: 10.5435/JAAOS-D-20-00473.**

13. Yi PH, Cross MB, Moric M, Levine BR, Sporer SM, Paprosky WG, Jacobs JJ, Della Valle CJ. **Do serologic and synovial tests help diagnose infection in revision hip arthroplasty with metal-on-metal bearings or corrosion? Clin Orthop Relat Res. 2015 Feb;473(2):498-505. doi: 10.1007/s11999-014-3902-5.**

14. Reito A, Parkkinen J, Puolakka T, Pajamäki J, Eskelinen A. **Diagnostic utility of joint fluid metal ion measurement for histopathological findings in metal-on-metal hip replacements. BMC Musculoskelet Disord. 2015 Dec 22;16:393. doi: 10.1186/s12891-015-0851-4.**

15. Kwon YM, Khormaee S, Liow MH, Tsai TY, Freiberg AA, Rubash HE. **Asymptomatic Pseudotumors in Patients with Taper Corrosion of a Dual-Taper Modular Femoral Stem: MARS-MRI and Metal Ion Study. J Bone Joint Surg Am. 2016 Oct 19;98(20):1735-1740. doi: 10.2106/JBJS.15.01325.**

16. Kwon YM, Tsai TY, Leone WA, Liow MHL. **Sensitivity and Specificity of Metal Ion Levels in Predicting "Pseudotumors" due to Taper Corrosion in Patients With Dual Taper Modular Total Hip Arthroplasty. J Arthroplasty. 2017 Mar;32(3):996-1000. doi: 10.1016/j.arth.2016.08.040.**

17. Incavo SJ. **Goals and Challenges of Identifying Asymptomatic Pseudotumors in Patients Treated with Total Hip Arthroplasty: Commentary on an article by Young-Min Kwon, MD, PhD, et al.: "Asymptomatic Pseudotumors in Patients with Taper Corrosion of a Dual-Taper Modular Femoral Stem: MARS-MRI and Metal Ion Study". J Bone Joint Surg Am. 2016 Oct 19;98(20):e93. doi: 10.2106/JBJS.16.00688.**

18. Hothi HS, Berber R, Whittaker RK, Blunn GW, Skinner JA, Hart AJ. **The Relationship Between Cobalt/Chromium Ratios and the High Prevalence of Head-Stem Junction Corrosion in Metal-on-Metal Total Hip Arthroplasty. J Arthroplasty. 2016 May;31(5):1123-7.**

19. Mahajan J, Bonner B, Oganesyan R, Yeo I, Klemt C, Kwon YM. **MARS MRI Characteristics of Adverse Local Tissue Reactions in Taper Corrosion of Metal-On-Polyethylene THA Differ From Metal-On-Metal THA. J Arthroplasty. 2020 Nov;35(11):3338-3342. doi: 10.1016/j.arth.2020.06.010.**

20. Scientific Committee on Emerging and Newly Identified Health Risks SCENIHR. The safety of Metal-on-Metal joint replacements with a particular focus on hip implants. https://www.siot.it/wp-content/uploads/pagine/analisi-documenti/doc/metal-metal.pdf. Last access: 13/05/2021

21. Di Laura A, Quinn PD, Panagiotopoulou VC, Hothi HS, Henckel J, Powell JJ, Berisha F, Amary F, Mosselmans JFW, Skinner JA, Hart AJ. **The Chemical Form of Metal Species Released from Corroded Taper Junctions of Hip Implants: Synchrotron Analysis of Patient Tissue. Sci Rep. 2017 Sep 8;7(1):10952. doi: 10.1038/s41598-017-11225-w.**

22. Hothi HS, Ilo K, Whittaker RK, Eskelinen A, Skinner JA, Hart AJ. **Corrosion of Metal Modular Cup Liners. J Arthroplasty. 2015 Sep;30(9):1652-6. doi: 10.1016/j.arth.2015.03.022.**

23. Markel DC, Bou-Akl T, Rossi MD, Pizzimenti N, Wu B, Ren W. **Blood metal levels, leucocyte profiles, and cytokine profiles in patients with a modular dual-mobility hip prosthesis: early results from a prospective cohort study. Bone Joint J. 2019 Sep;101-B(9):1035-1041. doi: 10.1302/0301-620X.101B9.BJJ-2019-0377.R2.**

24. Chalmers BP, Mangold DG, Hanssen AD, Pagnano MW, Trousdale RT, Abdel MP. **Uniformly low serum cobalt levels after modular dual-mobility total hip arthroplasties with ceramic heads: a prospective study in high-risk patients. Bone Joint J. 2019 Jun;101-B(6_Supple_B):57-61. doi: 10.1302/0301-620X.101B6.BJJ-2018-1403.R1.**

25. Di Laura A, Hothi HS, Henckel J, Kwon YM, Skinner JA, Hart AJ. **Retrieval Findings of Recalled Dual-Taper Hips. J Bone Joint Surg Am. 2018 Oct 3;100(19):1661-1672. doi: 10.2106/JBJS.17.00790.**

26. Mueller U, Panzram B, Braun S, Sonntag R, Kretzer JP. **Mixing of Head-Stem Components in Total Hip Arthroplasty. J Arthroplasty. 2018 Mar;33(3):945-951. doi: 10.1016/j.arth.2017.10.047.**

27. Wight CM, Lanting B, Schemitsch EH. **Evidence based recommendations for reducing head-neck taper connection fretting corrosion in hip replacement prostheses. Hip Int. 2017 Nov 21;27(6):523-531. doi: 10.5301/hipint.5000545.**

28. Rieker CB, Wahl P. **What the Surgeon Can Do to Reduce the Risk of Trunnionosis in Hip Arthroplasty: Recommendations from the Literature. Materials (Basel). 2020 Apr 21;13(8):1950. doi: 10.3390/ma13081950.**

29. Zimmer, Kinectiv technology. https://www.zimmerbiomet.com/content/dam/zimmer-biomet/medical-professionals/000-surgical-techniques/hip/zimmer-ml-taper-hip-prosthesis-with-kinectiv-technology-surgical-technique.pdf

30. Panagiotidou A, Cobb T, Meswania J, Skinner J, Hart A, Haddad F, Blunn G. **Effect of impact assembly on the interface deformation and fretting corrosion of modular hip tapers: An in vitro study. J Orthop Res. 2018 Jan;36(1):405-416. doi: 10.1002/jor.23601.**

31. Ceramtec head positioning. https://www.ceramtec.com/files/mt_clinical_guide_clinical_management_of_hip_arthroplasty.pdf

RESEARCH TOPIC IMPLANT FIXATION

RESEARCH TOPIC 15

Standard Test Methods for the Pre-Clinical Assessment of Primary and Secondary Stability in Cemented Total Joint Arthroplasty Taking into Consideration Physiological Force Transmission and Long-Term Fixation with Regard to the Clinical Application

Authors

Luca Cristofolini[1], Thomas M. Grupp[2,3], Volkmar Jansson[3], Richard T. Mayer[4]

[1]Department of Industrial Engineering, School of Engineering and Architecture, Alma Mater Studiorum – Università di Bologna, Bologna, Italy

[2]Aesculap AG Research & Development, Tuttlingen, Germany

[3]Ludwig Maximilians University Munich, Department of Orthopaedic and Trauma Surgery, Musculoskeletal University Center Munich (MUM), Campus Grosshadern, Munich, Germany

[4]Johnson & Johnson, Office of the Chief Medical Officer New Brunswick NJ, Zuchwil, Switzerland

1. Abstract

Question: Implant fixation

How can be assessed pre-clinically if a reasonable primary and secondary stability, as well as a physiological application of force / force transmission into the underlying bone can be achieved when using a cemented implant?

1. Which (standard) test methods can be used in case of a cemented implant fixation for the pre-clinical assessment of the primary stability and the cement mantle?
2. Are there test methods which enable the assessment of long-term fixation using a cemented implant (secondary stability)?
3. How meaningful are they for the actual clinical application?

Summary / Recommendation:

Loosening remains the main non-septic reason for revision both for knee and hip.

No standard is available for testing the primary or long-term stability of cemented implants, nor to assess the load transfer to the host bone. The only existing standards refer to the properties of the acrylic cement in itself.

Several testing methods to assess the primary and long-term performance of cemented hip stems have been published. Some of them have been validated by testing devices with clinically-known performance. Similarly, numerical studies have been published, based on the finite element method (FEM). Very little can be found for the stability of cemented acetabular cups (mainly under simplified conditions). The main limitation of experimental testing is that long-term simulations can only be performed on synthetic bone models. FEM allows simulating real bone and its adaptation over time. The main concern about FEM study is when the model is not validated.

Methods have been developed for testing the short-term and long-term stability of the femoral component of TKA, based on synthetic models. Also, the stability of the tibial component has been tested. FEM models have been developed also for the implant stability of TKA. Similar strengths and limitations apply to investigating the knee.

As reliable test protocols have been presented, these should be applied in the pre-clinical phase. Numerical models (FEM) shall be used during the design process to address those aspects that cannot be assessed experimentally. Only validated FEM should be accepted.

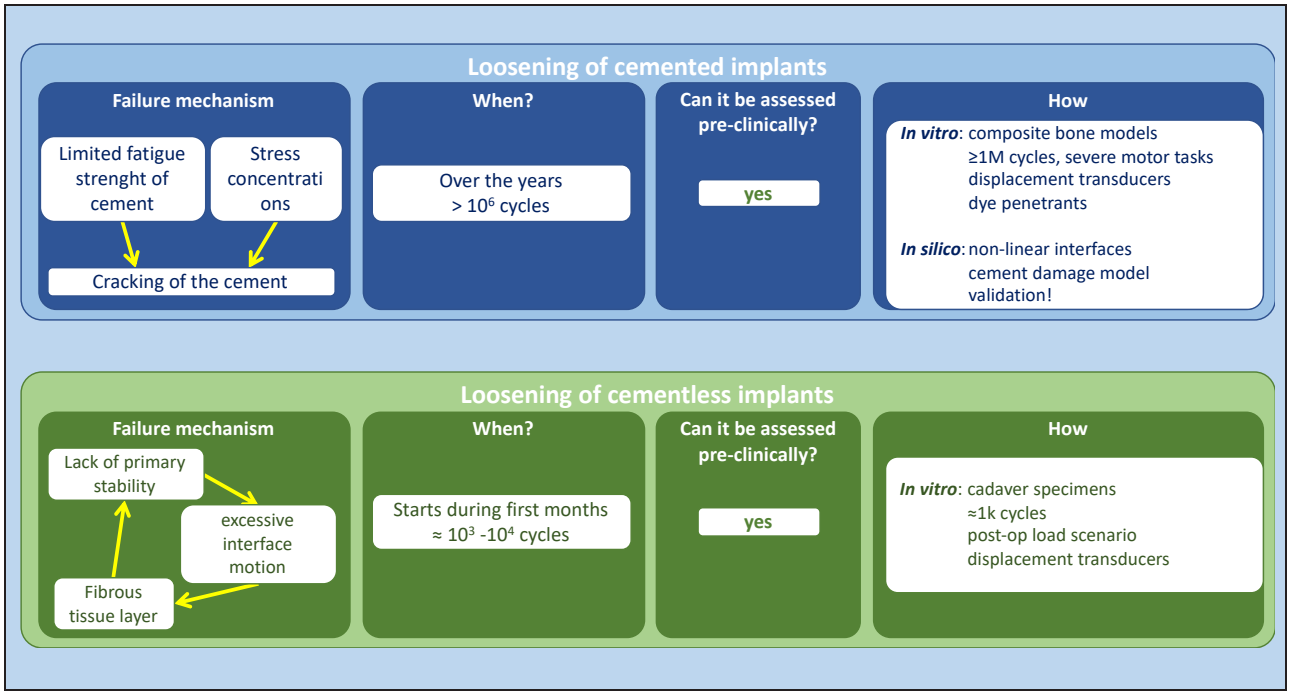
2. Level of Evidence

Moderate

3. Consensus Delegate Vote

93% - super majority, strong consensus (93% agree / 7% disagree / 0% abstain)

4. Graphical Abstract



5. Search Strategy

The search was carried out in three phases:
1) In the preliminary phase, the incidence of the different loosening failure scenarios was investigated. The Hip and Knee registries of arthroplasty were interrogated. Only registries meeting the following criteria were considered:

- Covering a population of at least 4 million citizens
- Including at least 1000 revision cases
- Follow up of at least 4 years

2) The ISO and ASTM database were interrogated for relevant test methods.
3) To investigate the published test methods, journal papers were searched, using PubMed (National Library of Medicine)

Search string:

(pre-clinical OR preclinical OR in vitro OR biomechanical) AND ("hip" OR "knee") AND ("implant" OR "prosthesis" OR "stem" OR "acetabulum" OR "tibial" OR "femoral" OR "patellar")

AND ("implant stability" OR "long term stability" OR "loosening") AND ("cemented" OR "cement")

Timeframe: ALL (1920-2020)

While the entire database, was searched, here only a selection of the relevant ones is presented.

6. Rationale

Aseptic loosening is the main cause of non-septic revision both for the hip (between 15.6 and 30.9% of the revisions, Table 1) and the knee (between 17.0 and 38.4%, Table 2). Most loosening events both for cemented hips and knee occur in the mid- and long-term and are associated with fatigue failure of the cement under cyclic loading (1–4). An active patient is likely to walk 1M steps every year or more.

While standards exist for testing the short-term (e.g., ISO 5833) and long-term performance (e.g., ISO 16402, ASTM F2118) of acrylic bone cement as a material, no standard exists for assessing the short- and long-term stability of cemented implants, nor the transfer of load from the prosthesis to the host bone.

Considering the relatively low failure rates of both THR and TKR, pre-clinical validation (both with in vitro experiments and with numerical models) must focus on the most critical scenarios (the tails of the distribution, associated with high BMI, active patients, poor bone quality etc) as these are the cases where failure might occur (5–7). Indeed, addressing the "average" patient would not capture the real risk of failure, which is an event occurring in non-average patients.

Stability of cemented hip stems

In-vitro test methods have been published to test the long-term stability of cemented hip stems (8–11). In order to test relevant fatigue scenarios 10 or more years of patient activities should be simulated. Considering the viscoelasticity of bone cements, test frequency should not exceed 4Hz. Therefore, an accelerated protocol is required, focusing on the most critical motor tasks (e.g., negotiating stairs, entering a car, stumbling) rather than low-magnitude ones (e.g., walking). Different motor tasks have been simulated in-vitro, highlighting the importance of a realistic load history (including occasional overloads) in eliciting failure of sub-optimal implants (12–14). Cadaveric specimens are not suitable for such long tests because of preservation issues, and because dead one tissue would undergo non-relevant fatigue fractures. Similar methods have also been implemented for cemented resurfacing prostheses (2). These protocols rely on displacement transducers (LVDT) and/or the use of dye penetrants to identify fatigue cracks in the cement mantle. The reliability of such protocols has been proven by applying the same protocol to commercial devices with known clinical follow-up (15,16). Also, several FE models have been published: these allow simulating different bone anatomies, bone quality and bone adaptation (17–20). Only models that have undergone a quantitative validation should be considered for pre-clinical purposes (21,22).

Stability of cemented hip cups

Very little can be found in the literature for the stability of cemented acetabula, probably also in relation to the limited use of cemented cups. Most published experiments relied on simplified test conditions, mostly blocks of synthetic bone models (2)(23,24)(25)(26,27). This approach offers two advantages: (i) the testing conditions are reproducible; (ii) long-term simulations can be performed without the problems of using cadaver tissue. The main limitation of this approach is the difference of anatomy and distribution of mechanical properties from the human pelvis. In vitro experiments with cadaveric pelvises focused on the short-term performance (14,28,29). Pilot studies have been published where synthetic (composite) pelvises have been used as a test bench. Digital image correlation (DIC) can be used to track implant motions and measure periprosthetic deformations at the same time. In principle, this approach could be extended to test the long-term performance of cemented cups in more realistic conditions.

While in principle it is possible to validate FE models, no detailed validated model to test acetabular stability has been presented.

Stability of cemented femoral knee components

Loosening of the femoral component of TKR is probably due to the "rolling force", which is loading the implant in a direction that can vary by nearly 90°. To simulate a relevant loading profile, this must be taken into account in the experiment. For this purpose, it is convenient to adopt the same simulators, same load profile and the same standard as for knee wear testing (ISO 14243). Also in this case, only synthetic bone models provide a suitable test bench to withstand cyclic loading for 1M cycles or more (30–32). The long-term stability of the femoral component of TKR has been investigated with similar tools (LVDTs and dye penetrants) to the cemented hip stem (33,34).

Numerical models enabled investigating in detail interface stresses and the load transfer mechanism (35–37).

	Range	RIPO Emilia Romagna	RIAP Italy*	Germany	Netherlands *	Norway **	Sweden		Uk	New Zealand	Australia
	Reference population	Emilia Romagna	(Lombardy, Tuscany, Marche, Apulia, Basilicata, Calabria, Sicily, and Campania, Bolzano, Trento and two hospital structures "Policlinico Città di Alessandria" and "Santa Maria della Misericordia" of Udine	Germany	Netherlands	Norway	Sweden		Uk	New Zealand	Australia
Diagnosis in revision	%	*(% of number of revisions)*	*(% of number of revisions)*	*(% of number of revisions)*	*(% of number of revisions)*	*(% of number of revisions)*	*dual mobility cup cemented (% of number of revisions)*	*dual mobility cup uncemented (% of number of revisions)*	*(% of number of revisions)*	*(% of number of revisions)*	*(% of number of revisions)*
Loosening of acetabulum component	15.6-30.9	30.9	18.1	15.6	22.8	30.8				21.7	
Loosening of femur component	8.1-31.1	13.3	10.8	10.9	18.9	31.1	8.1	29.2		17.2	
Loosening total aseptic	3.3-18.3	18.3	7.0	3.3							24.0
Loosening septic	1.1-60.4	1.1	8.8	15.2	19.1	6.5	60.4	41.7	12.3	12.2	
Loosening aseptic	26.2								26.2		
Prosthesis breakage*	1.8-5.2	5.2	3.5	1.8					2.5		
modular necks (% of Prosthesis breakage)		*34.4*									
liners (% of Prosthesis breakage)		*22.2*									
heads (% of Prosthesis breakage)		*14.5*									
stems (% of Prosthesis breakage)		*12.3*									
cups (% of Prosthesis breakage)		*12.5*									
liner and head (% of Prosthesis breakage)		*1.7*									
not specified (% of Prosthesis breakage)		*2.3*									
Primary instability	0.7	0.7									21.6
Prosthesis dislocation	9.2-19.0	9.2	13.9	11.7	19.0	8.3	16.2	16.7	12.0	18.6	
Inlay wear	4.2-18.0	4.2			18.0						
Wear	3.1-9.2		8.2	8.1		3.1			9.2		
Metallosis	0.8	0.8									
Femur fracture	3.7-13.2		12.3	10.9	13.2	3.7	10.7	12.5	10.4	12.5	
Acetabulum fracture	0.2	0.2									21.2
Bone fracture	7.1	7.1									
Pain	1.7-15.2	1.7	5.6			6.4			3.6	15.2	
Trauma	0.2	0.2									
Osteolysis acetabular (no loosening)	1.6					1.6					
Osteolysis femur	1.9					1.9					
Lysis	0.9-8.9		2.7	0.9					8.9		2.1
Condition after removal	9.2			9.2							
Girdlestone situation	2.1-5.5				5.5	2.1					
Two steps prosthesis removal	5.0	5.0									
Symptomatic mom bearing	3.7				3.7						
Previous prosthesis removal	1.6		1.6								
Malalignment	1.6-3.2			1.6					3.2		
Adverse reaction to particulate debris	7.8								7.8		
Peri-articular ossification	0.5-1.8	0.5			1.8						
Progression of arthrosis	0.3			0.3							
Disease progression	0.1		0.1								
Head-socket size mismatch	0.3								0.3		
Missing information	0.4					0.4					
Other	0.0-11.7	1.5	7.4	10.5	11.3	4.2	4.6	0.0	3.8		11.7

*= one patient may have more than one reason for revisionor re-surgery. As such the total proportion is over 100%
**= Diseases are not mutually exclusive

Table 1: Incidence of the different failure modes of THR according to the different implant registries.

Stability of cemented tibial knee components

Aseptic cemented tibial component loosening remains a major cause of TKA failure (38–42), beneath infection, instability and sub optimal alignment. In a cohort of 17,782 primary cemented TKA's reported to the Norwegian arthroplasty register during the years from 1994 to 2009, Gothesen et al. (38) evaluated the aseptic implant revisions, and found a ratio of 2.8 between tibial and femoral component loosening. Based on the overview on the incidence of different failure modes for TKA (Table 2) the ratio's in the registries are 1.8 in RIAP Italy, 2.1 in EPRD Germany, 2.0 in the Netherlands, 2.0 & 2.3 in Norway and 3.0 in New Zealand. Dyrhovden et al. (39) selected two cohorts of primary TKA's implanted during 1994 to 2004 (Period 1; n = 17,404) and during 2005 to 2015 (Period 2; n = 43,219) in Norway to investigate if the causes of revision for TKA have changed during the past two decades, and found that the relative risk for revision (RR) has been unchanged for the tibia (RR 1.0), but substantially decreased for the femur (RR 0.3). This underlines the need for stable primary and secondary fixation of cemented tibial trays. Several studies have been undertaken to analyze the primary and secondary stability of tibial trays in vitro, in vivo, in silico and ex vivo (40–45). To measure the primary implant-cement-bone interface fixation strength of cemented tibial components a standardised push-out testing in comparison to a long-term clinically established reference device has been

	Range	RIPO Emilia Romagna	RIAP Italy*	Germany	Netherlands *	Norway **			Sweden ***			Uk	New Zealand	Australia
	Reference population	Emilia Romagna	(Lombardy, Tuscany, Marche, Apulia, Basilicata, Calabria, Sicily, and Campania, Bolzano, Trento and two hospital structures "Policlinico Città di Alessandria" and "Santa Maria della Misericordia" of Udine	Germany	Netherlands	Norway			Sweden			Uk	New Zealand	Australia
Diagnosis in revision	%	-	-	-	-	total knee prosthesis with patella	total knee prosthesis without patella	Unicondylar knee prosthesis	TKA-O	TKA-RA	UKA-OA	-	-	-
		(% of number of revisions)	*(% of number of revisions)*	*(% of number of revisions)*	*(% of number of revisions)*	*(% of number of operations)*			*(% of number of revisions)*			*(% of number of revisions)*	*(% of number of revisions)*	*(% of number of revisions)*
Aseptic loosening	17.0-38.4	38.4							17.0	29.0	28.0	29.6		
Loosening of femoral component	2.6-9.0	2.6	5.5	4.4	9.0	0.9	0.4	2.3					8.0	
Loosening of patellar component	0.2-20		0.2	0.6	1.7	0.5	-	-					20.0	32.2
Loosening of several components	10.7-26.0		26.0	10.7										
Loosening of tibial component	9.3-24.0	9.6	9.9	9.3	21.0	2.1	0.9	2.4					24.0	
Septic loosening	2.6-24.0	2.6	15.5	14.7	18.0	1.5	0.8	0.3	31.0	27.0	3.0	24.0	23.0	
Patellar reasons	2.0-20.9								19.0	4.0	2.0			20.9
Arthrofibrosis	4.2-4.8			4.2	4.8									4.7
Progression of arthities	8.6-16.5				8.6							16.5		
Breakage of prosthesis	0.5-2.0	0.5	1.4	2.0										
Condition after removal	10.9			10.9										
Dislocation prosthesis	2.2-3.2	2.2	2.4									3.2		
Dislocation of patella	2.4				2.4	0.2	0.1	-						
Dislocation no patella	-					0.1	0.1	-						
Fractured spacer	0.1		0.1											
Periprosthetic fracture	1.2-4.5	1.7	1.2	3.0	1.9	0.3	0.2	0.5				4.5		
Implant wear	2.0-10.8	3.7	2.4	5.7	7.3				2.0	4.0	10.0	10.8		
Instability	1.8-26.3	1.8	3.5		26.3	1.3	0.6	0.8	14.0	9.0	10.0	14.7		11.2
Ligament instability	8.9			8.9										
Malalignment / rotation revision	1.8-12.6			1.8	12.6	0.7	0.3	0.9				5.5		
Osteolysis with fixed component	1.1			1.1										
Femoral components				*0.3*										
Tibial tray				*0.3*										
Patellar components				*0.1*										
Several components				*0.4*										
Pain	9.9-25.0	9.9	16.7		21.0	1.7	1.8	5.7				10.1	25.0	10.7
Trauma	0.5	0.5												
Progression of desease	0.9-32	0.9	2.3						13.0	20.0	32.0			
Restricted mobility	4.0			4.0										
Stiffness	1.0-4.9	1.0	1.6									4.9		
Defect polyethylene	-					1.2	0.2	0.7						
Two steps prosthesis removal	18.4	18.4												
Revision of knee removal	5.6				5.6									
Missing	-					0.1	0.1	0.2						
Other	3.0-20.4	6.3	11.4	18.7	8.0	0.7	0.4	1.1	3.0	4.0	14.0	8.8		20.4

*= one patient may have more than one reason for revisionor re-surgery. As such the total proportion is over 100%
**= Revision causes are not mutually exclusive
***= numbers taken from a histagram chart

TKA O = Total Knee Arthroplasty in Osteoarthritis
TKA RA = Total Knee Arthroplasty in Rheumatoid Arthritis
UKA OA = Unicompartmental Knee Arthroplasty

Table 2: Incidence of the different failure modes of TKR according to the different implant registries.

described, using a synthetic polyurethane foam model according to ASTM F1839-08 (46). Unless push-out test conditions do not reflect the physiologic TKA loading modes in vivo, where the tibial tray is predominantly subjected to combined compression and shear forces in a cyclic profile (40,47–49)of a full cemented keel and of an additional tibial stem on the primary stability of a posterior stabilised tibial plateau (VEGAÒ System Aesculap Tuttlingen, Germany, it is a suitable method to evaluate different influencing factors such as tibial tray material & surface textures, cement application & curing conditions (i.e. apposition time etc.) (46), surface contaminations (50) and lavage with dilute betadine to reduce the risk of acute peri-prosthetic joint infection (51).

To simulate the endurance behaviour of the implant-cement-bone interface fixation under clinically relevant severe anterior shear and internal-external (IE) torsional shear testing conditions a pre-clinical test methodology has been established, to distinguish in a direct comparison between tibial tray implant designs (52). To examine the influence of design, material and surface texture parameters on the long-term secondary fixation of cemented tibial trays, a demanding anterior shear and IE torsional shear fatigue testing simulating more than 10 years in vivo, shall be performed in comparison to a clinically established implant system. It is of importance to use a test setup and a bone model suitable for high cycle fatigue, allowing for appropriate cement penetration (40,52) and representing clinically relevant failure modes (44,53).

Additional investigations for specific failure modes and claims

Although stress shielding is generally not considered a critical issue, experimental methods to investigate load transfer have been published, either based on surface strain gauges (54,55) (56), or on digital image correlation (57–59), or on strain gauges and piezoelectric transducers embedded in the cement mantle or at the interface (60) (61). In this field, FE models are advantageous in estimating parameters that are difficult to assess experimentally, such as interface stress, simulating bone adaptation etc. (62–64). Again, only validated FE models should be considered.

Effect of patient- related and surgeon-related aspects

Experimental methods and numerical models have been successfully deployed also to assess the effects of sub-optimal surgery, e.g., excessively thin cement layer, implant mispositioning etc. (65–67). To assess the possible risks associated with patient factors (e.g., high BMI, poor bone quality) a sensitivity analysis should be performed whenever concerns exist about a specific design. Similarly, failures caused by sub-optimal implantation (e.g., implant mispositioning, improper use of acrylic cement, over/under-reaming) are unlikely to be detected in a Phase I or Phase II clinical trial, as highly expert and trained surgeons would be recruited: such failures occur when an implant undergoes widespread use, in a large number of centers. To detect such risks pre-clinically, the effects of all foreseeable surgical errors should be evaluated. The most flexible tool in this case are in silico simulations, which allow cost-effective exploration of multiple factors.

7. References

1. Lewis G. **Properties of acrylic bone cement: state of the art review. J Appl Biomater. 1997;38(2):155–82.**

2. Zant NP, Heaton-Adegbile P, Hussell JG, Tong J. **In Vitro Fatigue Failure of Cemented Acetabular Replacements: A Hip Simulator Study. J Biomech Eng. 2008 Apr 1;130(2):021019.**

3. Lennon AB, McCormack BA, Prendergast PJ. **The relationship between cement fatigue damage and implant surface finish in proximal femoral prostheses. Med Eng Phys. 2003;25(10):833–41.**

4. Erani P, Cristofolini L, Baleani M, Bignozzi MC, Cotifava M, Ihaddadene N, et al. **Quantitative crack surface morphology of bone cements in relation to propagation rate. Fatigue Fract Eng Mater Struct. 2007;30(9):783–95.**

5. Gunst, S., & Fessy, M. H.. **The effect of obesity on mechanical failure after total knee arthroplasty. Annals of translational medicine. 2015;3(20):310.**

6. Kenney C, Dick S, Lea J, Liu J, Ebraheim NA. **A systematic review of the causes of failure of Revision Total Hip Arthroplasty. J Orthop. 2019 Sep;16(5):393–5.**

7. Rosso F, Cottino U, Dettoni F, Bruzzone M, Bonasia DE, Rossi R. **Revision total knee arthroplasty (TKA): mid-term outcomes and bone loss/quality evaluation and treatment. J Orthop Surg. 2019 Dec;14(1):280.**

8. Maher SA, Prendergast PJ. **Discriminating the loosening behaviour of cemented hip prosthesis using measurements of migration and inducible displacement. J Biomech. 2002;35:257–65.**

9. Cristofolini L, Saponara Teutonico A, Monti L, Cappello A, Toni A. **Comparative in vitro study on the long term performance of cemented hip stems: validation of a protocol to discriminate between "good" and "bad" designs. J Biomech. 2003 Nov;36(11):1603–15.**

10. Cristofolini L, Duda G, Prendergast PJ. **Inter-laboratory consensus about stability testing of cemented and cementless hip stems. In: European Orthopaedic Research Society Transactions Vol 16, Baldini N, Savarino L eds, CLUEB Publ, Bologna: O66. 2006.**

11. Britton JR, Lyons CG, Prendergast PJ. **Measurement of the relative motion between an implant and bone under cyclic loading. Strain. 2004;40(4):193–202.**

12. Bergman G, Graichen, F, Rohlmann, A, Bender A, Heinlein B, Duda GN, et al. **Realistic loads for testing hip implants. Biomed Mater Eng. 2009;20(2):65–75.**

13. Cristofolini L, Juszczyk M, Martelli S, Taddei F, Viceconti M. **In vitro replication of spontaneous fractures of the proximal human femur. J Biomech. 2007 Jan;40(13):2837–45.**

14. Morosato F, Traina F, Cristofolini L. **Effect of different motor tasks on hip cup primary stability and on the strains in the periacetabular bone: An in vitro study. Clin Biomech. 2019 Dec;70:137–45.**

15. Cristofolini L, Saponara Teutonico A, Savigni P, Erani P, Viceconti M. **Preclinical assessment of the long-term endurance of cemented hip stems. Part 1: effect of daily activities--a comparison of two load histories. Proc Inst Mech Eng H. 2007 Aug;221(6):569–84.**

16. Cristofolini L, Erani P, Savigni P, Bordini B, Viceconti M. **Preclinical assessment of the long-term endurance of cemented hip stems. Part 2: in-vitro and ex-vivo fatigue damage of the cement mantle. Proc Inst Mech Eng H. 2007 Aug;221(6):585–99.**

17. Bagge M. A model of bone adaptation as an optimization process. **J Biomech. 2000 Nov;33(11):1349–57.**

18. Orwoll ES, Marshall LM, Nielson CM, Cummings SR, Lapidus J, Cauley JA, et al. **Finite Element Analysis of the Proximal Femur and Hip Fracture Risk in Older Men. J Bone Miner Res. 2009 Mar;24(3):475–83.**

19. Pahr DH, Zysset PK. **Finite Element-Based Mechanical Assessment of Bone Quality on the Basis of In Vivo Images. Curr Osteoporos Rep. 2016 Dec;14(6):374–85.**

20. Stops A, Wilcox R, Jin Z. **Computational modelling of the natural hip: a review of finite element and multibody simulations. Comput Methods Biomech Biomed Engin. 2011;963–79.**

21. Anderson AE, Ellis BJ, Maas SA, Peters CL, Weiss JA. **Validation of Finite Element Predictions of Cartilage Contact Pressure in the Human Hip Joint. J Biomech Eng. 2008 Oct 1;130(5):051008.**

22. Henninger HB, Reese SP, Anderson AE, Weiss JA. **Validation of computational models in biomechanics. Proc Inst Mech Eng [H]. 2010 Jul 1;224(7):801–12.**

23. Szivek JA, Thompson JD, Benjamin JB. **Chracterization of three formulations of a synthetic foam as models for a range of human cancellous bone types. J Appl Biomater. 1995;6:125–8.**

24. Schierjott RA, Hettich G, Baxmann M, Morosato F, Cristofolini L, Grupp TM. **Primary stability of a press-fit cup in combination with impaction grafting in an acetabular defect model. J Orthop Res Off Publ Orthop Res Soc. 2021;39(5):929–40.**

25. Palisserey V, Taylor M, Browne M. **Fatigue characterization of a polymer foam to use as a cancellous bone analog material in the assessment of orthopaedic devices. J Mater Sci Mater Med. 2004;15:61–7.**

26. Markel DC, Hora N, Grimm M. **Press-fit stability of uncemented hemispheric acetabular components: a comparison of three porous coating systems. Int Orthop. 2002;26(2):72–5.**

27. Crosnier EA, Keogh PS, Miles AW. **A novel method to assess primary stability of press-fit acetabular cups. Proc Inst Mech Eng [H]. 2014 Nov;228(11):1126–34.**

28. Morosato F, Cristofolini L, Castagnini F, Traina F. **Effect of cup medialization on primary stability of press-fit acetabular cups. Clin Biomech. 2020 Dec 1;80(paper 105172):1–6.**

29. Morosato F, Traina F, Schierjott RA, Hettich G, Grupp TM, Cristofolini L. **Primary Stability of Revision Acetabular Reconstructions Using an Innovative Bone Graft Substitute: A Comparative Biomechanical Study on Cadaveric Pelvises. Materials. 2020;13(19):1–11.**

30. Cristofolini L, Viceconti M, Cappello A, Toni A. **Mechanical validation of whole bone composite femur models. J Biomech. 1996 Apr;29(4):525–35.**

31. Cristofolini L, Viceconti M. **Mechanical validation of whole bone composite tibia models. J Biomech. 2000 Mar;33(3):279–88.**

32. Heiner AD. **Structural properties of fourth-generation composite femurs and tibias. J Biomech. 2008;41(15):3282–4.**

33. Cristofolini L, Affatato S, Erani P, Tigani D, Viceconti M. **Implant fixation in knee replacement: preliminary in vitro comparison of ceramic and metal cemented femoral components. Knee. 2009;16:101–8.**

34. Cristofolini L, Affatato S, Erani P, Leardini W, Tigani D, Viceconti M. **Long-term implant-bone fixation of the femoral component in total knee replacement. Proc Inst Mech Eng H. 2008 Apr;222(3):319–31.**

35. Afsar E, Taspinar F, Calik BB, Ozkan Y, Gok K. **Use of the finite element analysis to determine stresses in the knee joints of osteoarthritis patients with different Q angles. J Braz Soc Mech Sci Eng. 2017 Apr;39(4):1061–7.**

36. Huiskes R, Strens PHGE, van Heck J, Slooff TJJH. **Interface stresses in the resurfaced hip: Finite element analysis of load transmission in the femoral head. Acta Orthop Scand. 1985 Jan;56(6):474–8.**

37. Lenthe GHV. **Stress shielding after total knee replacement may cause bone resorption in the distal femur. J Bone Jt Surg. 1997;79(1):6.**

38. Gøthesen Ø, Espehaug B, Havelin L, Petursson G, Lygre S, Ellison P, et al. **Survival rates and causes of revision in cemented primary total knee replacement: A report from the Norwegian Arthroplasty Register 1994–2009. Bone Jt J. 2013 May;95-B(5):636–42.**

39. Dyrhovden GS, Lygre SHL, Badawy M, Gøthesen Ø, Furnes O. **Have the Causes of Revision for Total and Unicompartmental Knee Arthroplasties Changed During the Past Two Decades? Clin Orthop. 2017 Jul;475(7):1874–86.**

40. Grupp TM, Saleh KJ, Holderied M, Pfaff AM, Schilling C, Schroeder C, et al. **Primary stability of tibial plateaus under dynamic compression-shear loading in human tibiae – Influence of keel length, cementation area and tibial stem. J Biomech. 2017 Jul;59:9–22.**

41. Cawley DT, Kelly N, McGarry JP, Shannon FJ. **Cementing techniques for the tibial component in primary total knee replacement. Bone Jt J. 2013 Mar;95-B(3):295–300.**

42. Schlegel UJ, Bruckner T, Schneider M, Parsch D, Geiger F, Breusch SJ. **Surface or full cementation of the tibial component in total knee arthroplasty: a matched-pair analysis of mid- to long-term results. Arch Orthop Trauma Surg. 2015 May;135(5):703–8.**

43. Schlegel UJ, Siewe J, Delank KS, Eysel P, Püschel K, Morlock MM, et al. **Pulsed lavage improves fixation strength of cemented tibial components. Int Orthop. 2011 Aug;35(8):1165–9.**

44. Gebert de Uhlenbrock A, Püschel V, Püschel K, Morlock MM, Bishop NE. **Influence of time in-situ and implant type on fixation strength of cemented tibial trays — A post mortem retrieval analysis. Clin Biomech. 2012 Nov;27(9):929–35.**

45. Staats K, Wannmacher T, Weihs V, Koller U, Kubista B, Windhager R. **Modern cemented total knee arthroplasty design shows a higher incidence of radiolucent lines compared to its predecessor. Knee Surg Sports Traumatol Arthrosc. 2019 Apr;27(4):1148–55.**

46. Grupp TM, Schilling C, Schwiesau J, Pfaff A, Altermann B, Mihalko WM. **Tibial Implant Fixation Behavior in Total Knee Arthroplasty: A Study With Five Different Bone Cements. J Arthroplasty. 2020 Feb;35(2):579–87.**

47. Zhao D, Banks SA, D'Lima DD, Colwell CW, Fregly BJ. **In vivo medial and lateral tibial loads during dynamic and high flexion activities. J Orthop Res. 2007 May;25(5):593–602.**

48. Kutzner I, Heinlein B, Graichen F, Bender A, Rohlmann A, Halder A, et al. **Loading of the knee joint during activities of daily living measured in vivo in five subjects. J Biomech. 2010 Aug;43(11):2164–73.**

49. Bergmann G, Bender A, Graichen F, Dymke J, Rohlmann A, Trepczynski A, et al. **Standardized Loads Acting in Knee Implants. Williams BO, editor. PLoS ONE. 2014 Jan 23;9(1):e86035.**

50. Billi F, Kavanaugh A, Schmalzried H, Schmalzried TP. **Techniques for improving the initial strength of the tibial tray-cement interface bond. Bone Jt J. 2019 Jan;101-B(1_Supple_A):53–8.**

51. Grupp TM, Puente Reyna A, Schilling C, Masi B, Fritz B, Mihalko WM, et al. **Is there an influence of dilute povidone-iodine (betadine) lavage on the interface fixation of tibial baseplates under clinical relevant conditions? Eur Soc Biomech. 2021 Jul 12;(Abstract 26th).**

52. Grupp TM, Schilling C, fritz B, Puente Reyna A, Pfaff A, Taunt C, et al. **New method to evaluate endurance of cemented tibial tray fixation under anterior and internal-external torsional shear. Eur Orthop Res Soc. 2020 Sep 17;Abstract 28.**

53. Cerquiglini A, Henckel J, Hothi H, Allen P, Lewis J, Eskelinen A, et al. **Analysis of the Attune tibial tray backside: A comparative retrieval study. Bone Jt Res. 2019 Mar;8(3):136–45.**

54. Huiskes R, Janssen JD, Sloof TJ. **A detailed comparison of experimental and theoretical stress-analyses of a human femur. In: Cowin SC, editor. Mechanical properties of bone, AMD 45, ASME Publ, New York. 1981. p. 211–34.**

55. Cristofolini L. **A critical analysis of stress shielding evaluation of hip prostheses. Crit Rev Biomed Eng. 1997;25(4–5):409–83.**

56. Grassi L, Isaksson H. **Extracting accurate strain measurements in bone mechanics: A critical review of current methods. J Mech Behav Biomed Mater. 2015 Oct;50:43–54.**

57. Palanca M, Tozzi G, Cristofolini L. **The use of digital image correlation in the biomechanical area: a review. Int Biomech. 2016;3(1):1–21.**

58. Morosato F, Traina F, Cristofolini L. **A reliable in vitro approach to assess the stability of acetabular implants using digital image correlation. Strain. 2019;55(4 (e12318)):1–12.**

59. Grassi L, Isaksson H. **Extracting accurate strain measurements in bone mechanics: A critical review of current methods. J Mech Behav Biomed Mater. 2015;50:43–54.**

60. Cristofolini L, Marchetti A, Cappello A, Viceconti M. **A novel transducer for the measurement of cement-prosthesis interface forces in cemented orthopaedic devices. Med Eng Phys. 2000 Sep;22(7):493–501.**

61. Cristofolini L, Viceconti M. **Development and validation of a technique for strain measurement inside polymethylmethacrylate. J Strain Anal Eng Des. 2000;35(1):21–33.**

62. Chennimalai Kumar N, Dantzig JA, Jasiuk IM, Robling AG, Turner CH. **Numerical Modeling of Long Bone Adaptation due to Mechanical Loading: Correlation with Experiments. Ann Biomed Eng. 2010 Mar;38(3):594–604.**

63. Gefen A. **Computational simulations of stress shielding and bone resorption around existing and computer-designed orthopaedic screws. Med Biol Eng Comput. 2002 May;40(3):311–22.**

64. ten Broeke RHM, Tarala M, Arts JJ, Janssen DW, Verdonschot N, Geesink RGT. **Improving peri-prosthetic bone adaptation around cementless hip stems: A clinical and finite element study. Med Eng Phys. 2014 Mar;36(3):345–53.**

65. Liau J-J, Cheng C-K, Huang C-H, Lo W-H. **The effect of malalignment on stresses in polyethylene component of total knee prostheses – a finite element analysis. Clin Biomech. 2002 Feb;17(2):140–6.**

66. Shi J, Browne M, Strickland M, Flivik G, Taylor M. **Sensitivity analysis of a cemented hip stem to implant position and cement mantle thickness. Comput Methods Biomech Biomed Engin. 2014 Nov 18;17(15):1671–84.**

67. Viceconti M, Monti L, Muccini R, Bernakiewicz M, Toni A. **Even a thin layer of soft tissue may compromise the primary stability of cementless hip stems. Clin Biomech. 2001 Nov;16(9):765–75.**

RESEARCH TOPIC 16

Test Methods for the Pre-Clinical Assessment of Primary and Secondary Stability in Cementless Total Joint Arthroplasty Taking into Consideration Physiological Force Transmission, Stress Shielding and Long-Term Fixation with Regard to the Clinical Application

Authors

Dennis Janssen[1], Christoph Schilling[2], Jan Philippe Kretzer[3], Richard T. Mayer[4]

[1]Radboud University Medical Center, Orthopaedic Research Lab, Research Institute for Medical Innovation, Nijmegen, the Netherlands

[2]Aesculap AG Research & Development, Tuttlingen, Germany

[3]Laboratory of Biomechanics & Implant Research, Department of Orthopaedic Surgery, Heidelberg University Hospital, Heidelberg, Germany

[4]Johnson & Johnson, Office of the Chief Medical Officer New Brunswick NJ, Zuchwil, Switzerland

1. Abstract

Question: Implant fixation

How can the primary and secondary stability and physiological load transfer to the peri-prosthetic bone of cementless implants be assessed in a pre-clinical stage?

Q1: Which methods can be used to assess the primary stability of cementless implants?
Q2: Which methods can be used to assess the physiological load transfer and stress shielding of cementless implants?

Summary / Recommendation:

Q1: Primary stability of cementless hip and knee replacement implants can be assessed through analysis of implant-bone interface micromotions. Interface micromotions can be measured through physical testing of reconstructions, in synthetic or cadaveric bones, using various measurement techniques. Computational modelling of interface micromotions based on finite element analysis requires mechanical validation against physical testing and is suitable for obtaining a global overview of primary stability. Analysis of a population of models increases robustness and allows for analysis of outliers. Due to the lack of a validated clinically relevant threshold for micromotions, reference testing against predicate devices is necessary.

Q2: Physiological load transfer and stress-shielding of cementless devices can be assessed by comparing bone strains before and after reconstruction using experimental and computational techniques. Various measurements techniques are available for evaluation of bone strains, in reconstructions in synthetic and cadaveric bones. Computational models are suitable to assess the change in bone strains but can also be used to simulate long-term bone density changes through adaptive remodeling simulations. Clinical relevance needs to be demonstrated against clinical DEXA data. Therefore, reference testing against predicate devices is required for new devices.

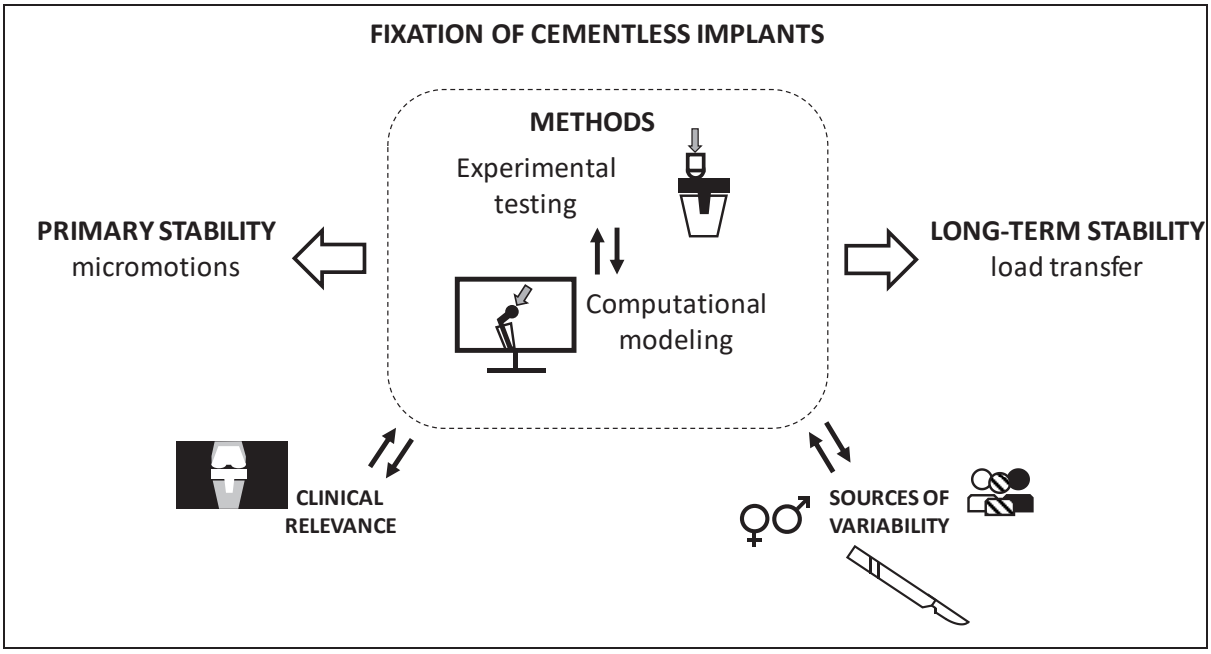
2. Level of Evidence

Moderate

3. Consensus Delegate Vote

97% - unanimous, strongest consensus (97% agree / 0% disagree / 3% abstain)

4. Graphical Abstract



5. Search Strategy

NA

6. Rationale

Q1:

Primary stability is evaluated by measuring micromotions at the implant-bone interface under physiological loading. Animal studies have demonstrated that osseointegration of the implant depends on the micromotions at the interface, with micromotions exceeding 150 µm leading to the formation of a soft tissue layer that may obstruct secondary fixation [1]. Assessment of interface micromotions can be divided in physical testing, and computational modelling.

In physical testing of primary stability reconstructions are subjected to loads representing activities of daily living. Loading configurations can be based on standards (see also Consensus Statements mechanical component testing) or databases on in vivo measurements of joint loads [2,3]. Loading configurations may either represent high-frequency, low impact activities (e.g. walking or stair climbing) or low-frequency, high-impact events (e.g. stumbling or a deep knee bend). Loading can be applied in experimental set-ups in a quasi-static manner by applying the peak load of a dynamic activity [4], while some set-ups enable application of full dynamic loading cycles [5].

Implant-bone micromotions can be measured using displacement transducers, in which the relative displacement is measured by placing transducers at specific locations across the interface [6]. Digital image correlation (DIC) is an optical method to measure micromotions by tracking relative displacements of applied dots or speckle patterns across the interface using a camera system [4,5]. Roentgen stereophotogrammetric analysis (RSA) uses bi-planar radiographs to measure relative displacements of rigid bodies constructed from markers attached to the implant and bone, and therefore provide information on migration and inducible displacements on a larger scale.

Testing can be performed either in cadaveric bones [4,6] or synthetic bone constructs [6,7]. Testing in cadaveric bones provides the closest approximation of the clinical situation, including the anatomical variability that is present in the patient population. Synthetic bone constructs provide consistency between specimens, thereby minimizing inter-specimen variability and increasing repeatability.

Computational analysis of implant-bone micromotions is generally based on finite element analysis. An advantage of finite element analysis over physical testing is that it allows for evaluation of micromotions over the full implant surface, including locations that are not visible in experimental testing [5,8]. A model of a reconstruction is created by combining the 3D design file of the implant with a model of the bone it is implanted in. CT scans are used to extract bone geometry and material properties. Implant positioning and bone preparation (bone cuts and broaching geometry) are replicated based on surgical instructions and 3D design files of instrumentation. Mechanical validity of computational models can be demonstrated through comparison against experimental testing. Important input model parameters for primary fixation of cementless implants are the coefficient of friction, the amount of press fit (or interference fit), and the (non-linear) material properties of the implant and bone.

The coefficient of friction of the implant-bone interface depends on the surface treatment of the implant and can be assessed in friction experiments [9]. Testing of surface specimens against cadaveric bone provides the closest approximation of the clinical situation, including the effect of fat and marrow on the frictional resistance. Testing against synthetic bone specimens reduces variability in bone material properties, and therefore facilitates inter-specimen comparison, which is mainly suitable in the early design phase.

The interference fit simulated in computational models can assume an idealized surface, replicating a uniform press-fit over the entire contact interface and representing the optimal fixation condition. A more clinically relevant situation can be simulated by including errors in the preparation of the bone bed, and implant positioning errors or malalignment [10].

Material properties of the implants and coating can be based on literature or based on specifications supplied by the manufacturer. Bone material properties are generally based on calibrated greyscale values derived from CT scans, providing a heterogeneous distribution of bone properties over the model. Several sources are available in literature providing equations relating greyscale values to elastic modulus, Poisson ratio, and yield strength. In computational studies on primary fixation linear elastic bone properties can be applied [11], or the post-yield behavior can be included in the simulations by using material models that include plasticity formulations [12]. Computational time is reduced by using elastic material properties, while the inclusion of plasticity provides a closer representation of the crushing of bone during implant insertion and impaction.

As implant-bone interface micromotions cannot be verified in patients, clinical validity of micromotions as measured in physical tests or as simulated in computational models cannot be demonstrated in a direct manner. The currently accepted approach for evaluation of new devices therefore is to compare results of testing or simulations against those of implants that have a good clinical track record as demonstrated in clinical studies or implant registries (predicate devices). Although no standards exist there are established methods for analysis femoral TKA stability in which micromotions are measured under gait and deep knee bend loading, and fixation is tested in deep flexion push-off tests [13,14]. Micromotion testing in tibial TKA components in synthetic bone has been widely published and has been described along with validation of an associated computational model by Navacchia et al. [15].

Simulating a population of models makes computational analysis of primary stability more robust, as it includes sources of variation in parameters related to the surgery (e.g. implant alignment, surgical technique) and patient (e.g. anatomy, sizing, BMI, bone quality). This allows for statistical analysis of the critical parameters that influence primary stability, but also the identification of outliers that may lead to clinical failure [16,17].

Q2:

By placing orthopaedic implants, the load transfer to the periprosthetic bone is changed relative to the native situation. In some locations bone stresses may be elevated, resulting in densification of the bone, while at other locations stress shielding of the bone may occur, which may result in local bone resorption. Methods for testing the physiological load transfer therefore aim at measuring the altered bone strains following a total hip or knee reconstruction. Assessment of bone strains can be performed through physical testing and computational modelling.

Similar to primary fixation testing, reconstructions are subjected to loads representing activities of daily living. Loading configurations can be based on standards (see also question 5B) or databases on in vivo measurements of joint loads [2,3]. Loading configurations generally represent high-frequency activities (e.g. walking or stair climbing). Low-frequency, high-impact loads (e.g. stumbling or a deep knee bend) can be applied to investigate overloading and the risk of peri-prosthetic fracture [18]. Comparison of strains in the native and reconstructed situations provides insight into the changes in load transfer following implant placement.

Bone strains can be measured using strain gauges, providing results at distinct locations around the reconstruction [19,20]. Digital image correlation allows for measurement of a continuous distribution of periprosthetic strains [20,21]. A more qualitative indication of periprosthetic bone strain can be obtained using photoelasticity techniques [22].

Testing can be performed either in cadaveric bones [18,19] or synthetic bones [20,21]. Testing in cadaveric bones provides the closest approximation of the mechanical properties of the bone and includes the anatomical variability that is present in the patient population. Synthetic bone constructs provide consistency between specimens, thereby minimizing inter-specimen variability and increasing repeatability.

The effect of joint reconstruction on physiological load transfer can be analyzed computationally by simulating bone strains in the native and reconstructed situations [21]. Mechanical validity of computational modelling of (periprosthetic) bone strains needs to be verified through comparison against physical testing. An advantage of computational strain analyses is that they provide information on the strain distribution in the full bone volume, rather than the surface strains that can be measured in physical testing. Analysis of bone strains provides insights in bone regions that are at risk of stress shielding, or periprosthetic fractures.

While analysis of bone strains provides information on the direct post-operative situation, models based on strain-adaptive remodeling algorithms [23,24] simulate the iterative process of bone formation and resorption. In strain adaptive remodeling simulations, the bone density (and stiffness) is updated based on a remodeling stimulus, for which generally the change in strain energy from the pre- to the post-operative situation is used.

Clinical relevance of test methods can be established by comparison of bone strains or remodeling outcomes against dual-energy X-ray absorptiometry (DEXA) studies. DEXA studies provide information on temporal changes in bone density at specific periprosthetic locations. Standardized regions of interest include the Gruen zones for hip stems [25], and Charnley zones for acetabular cups [26]. No standard regions of interest exist for knee reconstructions [27,28]. Results of clinical DEXA studies generally display a large variability, mainly due to relatively small number of patients that is included, which complicates validation and demonstration of clinical relevance of computational modelling. Analysis of populations of models, however, improves the robustness of predictions, and provides the possibility to identify parameters related to patient, implant design, and surgery, that may influence peri-prosthetic bone changes [29].

7. References

1. Kienapfel H, Sprey C, Wilke A, et al. **Implant fixation by bone ingrowth. J Arthroplasty 1999; 14: 355–68.**

2. G. Bergmann, A. Bender, J. Dymke, G. Duda, P. Damm. **Standardized Loads Acting in Hip Implants. PLoS ONE 11(5): e0155612. doi:10.1371/journal.pone.0155612**

3. G. Bergmann, A. Bender, F. Graichen, J. Dymke, A. Rohlmann, A. Trepczynski, M. O. Heller, I. Kutzner: **"Standardized Loads Acting in Knee Implants." PLoS ONE 2014 9(1): e86035. doi:10.1371/journal.pone.0086035**

4. Dion C, Yamomo G, Howard J, et al. **Revision total knee arthroplasty using a novel 3D printed titanium augment: A biomechanical cadaveric study. J Mech Behav Biomed Mater 2020; 110: 103944.**

5. Yang H, Bayoglu R, Renani MS, et al. **Validation and sensitivity of model-predicted proximal tibial displacement and tray micromotion in cementless total knee arthroplasty under physiological loading conditions. J Mech Behav Biomed Mater 2020; 109: 103793.**

6. Leuridan S, Goossens Q, Roosen J, et al. **A biomechanical testing system to determine micromotion between hip implant and femur accounting for deformation of the hip implant: Assessment of the influence of rigid body assumptions on micromotions measurements. Clin Biomech (Bristol, Avon) 2017; 42: 70–78.**

7. Morwood MP, Guss AD, Law JI, et al. **Metaphyseal Stem Extension Improves Tibial Stability in Cementless Total Knee Arthroplasty. J Arthroplasty 2020; 35: 3031–3037.**

8. Dabirrahmani D, Hogg M, Kohan L, et al. **Primary and long-term stability of a short-stem hip implant. Proc Inst Mech Eng H 2010; 224: 1109–19.**

9. Damm NB, Morlock MM, Bishop NE. **Friction coefficient and effective interference at the implant-bone interface. J Biomech 2015; 48: 3517–21.**

10. Berahmani S, Hendriks M, de Jong JJA, et al. **Evaluation of interference fit and bone damage of an uncemented femoral knee implant. Clin Biomech (Bristol, Avon) 2018; 51: 1–9.**

11. Wong AS, New AM, Isaacs G, et al. **Effect of bone material properties on the initial stability of a cementless hip stem: a finite element study. Proc Inst Mech Eng H 2005; 219: 265–75.**

12. Berahmani S, Janssen D, Verdonschot N. Berahmani S, Janssen D, Verdonschot N. **Experimental and computational analysis of micromotions of an uncemented femoral knee implant using elastic and plastic bone material models. J Biomech 2017; 61: 137–143.**

13. Berahmani S, Hendriks M, Wolfson D, et al. **Experimental pre-clinical assessment of the primary stability of two cementless femoral knee components. J Mech Behav Biomed Mater 2017; 75: 322–329**

14. Sánchez E, Schilling C, Grupp TM, et al. **The effect of different interference fits on the primary fixation of a cementless femoral component during experimental testing. J Mech Behav Biomed Mater 2021; 113: 104189.**

15. Navacchia A, Clary CW, Wilson HL, et al. **Validation of model-predicted tibial tray-synthetic bone relative motion in cementless total knee replacement during activities of daily living. J Biomech 2018; 77:115-123.**

16. Bah MT, Nair PB, Taylor M, et al. **Efficient computational method for assessing the effects of implant positioning in cementless total hip replacements. J Biomech 2011; 44: 1417-22.**

17. O'Rourke D, Bottema M, Taylor M. **Sampling strategies for approximating patient variability in population-based finite element studies of total hip replacement. Int J Numer Method Biomed Eng 2019; 35: e3168.**

18. Rupprecht M, Sellenschloh K, Grossterlinden L, et al. **Biomechanical evaluation for mechanisms of periprosthetic femoral fractures. J Trauma 2011; 70: E62-6.**

19. Decking R, Puhl W, Simon U, et al. **Changes in strain distribution of loaded proximal femora caused by different types of cementless femoral stems. Clin Biomech (Bristol, Avon) 2006; 21: 495-501.**

20. Correa TA, Pal B, van Arkel RJ, et al. **Reduced tibial strain-shielding with extraosseous total knee arthroplasty revision system. Med Eng Phys 2018; 62: 22-28.**

21. Chanda S, Dickinson A, Gupta S, et al. **Full-field in vitro measurements and in silico predictions of strain shielding in the implanted femur after total hip arthroplasty. Proc Inst Mech Eng H 2015; 229: 549-59.**

22. Arno S, Fetto J, Nguyen NQ, et al. **Evaluation of femoral strains with cementless proximal-fill femoral implants of varied stem length. Clin Biomech (Bristol, Avon) 2012; 27: 680-5.**

23. Huiskes R, Weinans H, Grootenboer HJ, et al. **Adaptive bone-remodeling theory applied to prosthetic-design analysis. J Biomech 1987; 20: 1135-50.**

24. Turner AW, Gillies RM, Sekel R, et al. **Computational bone remodelling simulations and comparisons with DEXA results. J Orthop Res 2005; 23: 705-12.**

25. Bodén HS, Sköldenberg OG, Salemyr MO, et al. **Continuous bone loss around a tapered uncemented femoral stem: a long-term evaluation with DEXA. Acta Orthop 2006; 77: 877-85.**

26. Anderl C, Mattiassich G, Ortmaier R, et al. **Peri-acetabular bone remodelling after uncemented total hip arthroplasty with monoblock press-fit cups: an observational study. BMC Musculoskelet Disord 2020; 21: 652.**

27. Minoda Y, Kobayashi A, Iwaki H, et al. **Comparison of bone mineral density between porous tantalum and cemented tibial total knee arthroplasty components. J Bone Joint Surg Am 2010; 92: 700-6.**

28. Järvenpää J, Soininvaara T, Kettunen J, et al. **Changes in bone mineral density of the distal femur after total knee arthroplasty: a 7-year DEXA follow-up comparing results between obese and nonobese patients. Knee 2014; 21: 232-5.**

29. Anijs T, Wolfson D, Verdonschot N, et al. **Population-based effect of total knee arthroplasty alignment on simulated tibial bone remodeling. J Mech Behav Biomed Mater 2020; 111: 104014.**

RESEARCH TOPIC 17

How to Assess Primary and Secondary Stability of Orthopaedic Joint Replacement Devices in a Clinical Setting Considering also How to Obtain/ Ensure Optimal Force Transmission into the Underlying Bone

Authors

Johan Kärrholm[1], Marcel Dreischarf[2], Richard T. Mayer[3], Rob G.H.H. Nelissen[4]

[1]Department of Orthopaedics, Institute of Clinical Sciences, Sahlgrenska Academy, University of Gothenburg, Gothenburg, Sweden
[2]Berlin Institute of Health, Julius Wolff Institute for Biomechanics and Musculoskeletal Regeneration, Charité–Universitätsmedizin Berlin, Germany
[3]Johnson & Johnson, Office of the Chief Medical Officer New Brunswick NJ, Zuchwil, Switzerland
[4]Department of Orthopaedics, Leids Universitair Medisch Centrum, Leiden, Netherlands

1. Abstract

Question: Implant Fixation

How can be clinically assessed if, using an implant, a reasonable primary and secondary stability, as well as a physiological application of force / force transmission into the underlying bone can be achieved?

1. What are the requirements for the study design, which evaluation criteria are relevant and are there special requirements for the imaging?
2. Which influencing factors exist in the clinical application, which cannot or can only poorly be depicted in the pre-clinical assessments?

Summary / Recommendation:

1. Observational studies suffer from lack of a control group, incomplete data collection, selection bias and confounding by indication. Nevertheless, they can signal inferior implant designs. But criteria on methodological quality and generalizability are mandatory for interpretation (e.g., AQUILA criteria).
2. Hip and knee Implant registries with 80-100% coverage and data completeness for primary and revision surgery are important to provide real world data. Also, in these large observational studies, which could include > 200.000 patient groups confounding by indication exists, thus no causal effects can be determined. But well organised registries can be used for nested prospective randomised trials, which should be preferred due to high external validity. Besides, registries with a lower coverage, e.g. the German EPRD, can provide important information to draw conclusions on poorly performing implants ("signalling function"). While they can help to identify substandard implants, to prove that an implant performs very good a high quality registry with very high coverage is required.
3. Implant fixation should be studied in prospective, preferably multi-centre, randomized studies with use of a well-documented standard implant as control, the benchmark. The benchmark implant should have a 10 year implant survival (or 1-survival: revision) of at least 90-95% depending on type of implant (i.e. revision of maximum 10 or 5%) based on data from implant registries (see for details on confidence interval and minimum number needed, the section on implant registries). These studies should ideally be constructed as nested trials within implant registries, using a non-inferiority design.
4. The clinical performance of an implant cannot or only poorly be predicted based on preclinical laboratory testing or with e.g. with finite element (FE) models. Since a multitude of factors related to the patient (e.g. bone morphology, biological and biomechanical adaptations to load, individual variations of inflammatory response to wear debris, activity) and surgical technique factors (choice of approach, implant positioning, soft-tissue damage) will affect the implant performance and fixation. In future improved knowledge about the influence of these factors may be used to improve different ways of mathematical modelling.

2. Level of Evidence

Narrative review, with reference list

3. Consensus Delegate Vote

90% – super majority, strong consensus (90% agree / 2.5% disagree / 7.5% abstain)

4. Graphical Abstract

n/a

5. Search Strategy

Narrative review

6. Rationale

Basically, three modes of implant failure can be detected: expected, early detected failures (1), expected, late detected failures (2) and unexpected failures (3). Expected, early detected failure modes must be detected in a pre-market setting (e.g., fatigue of metal, excessive liner wear). Expected, late detected failures are discovered in a post-market setting (e.g., excessive early migration of the implant). Unexpected failure modes can be present in both the early pre-market and late post-market phase, depending on the type of failure (e.g., excessive implant migration, adverse biologic response such as metal induced pseudotumors) and material corrosion or breakdown (e.g., implant fractures). In general, the longer the pre-market phase will last, the higher the likelihood of finding unexpected failures. Although evident failure modes will be detected in large patient groups after several years of implantation, the goal of a new implant introduction should be to prevent adverse events before massive introduction to the market. Clinical evidence needs to be free of bias (e.g., indication bias, selection bias) and based on randomly collected patient cohorts with minimum loss of follow up. If a new medical device (e.g., hip, knee, spine prosthesis/device) is to be evaluated, it should perform at least as good as the best performing on the market (non-inferiority trial design). Definition of the endpoint (e.g., revision surgery for implant loosening), data collection, quality of data (e.g., validity, completeness, representativeness), radiological and statistical methods within the comparison should be succinctly defined.

Clinical study designs for evaluation of implant fixation:

Observational cohort studies

Although observational studies are often used, they suffer from lack of a control group, incomplete data collection, selection bias and confounding by indication. Nevertheless, they can signal inferior implant designs since most of the failure is due to aseptic implant loosening. However, the validity of observational data should be related to the quality of methods used and generalizability of the results based on numbers included, lost to follow up and patient characteristics corresponding to the AQUILA (assessment of quality in lower limb arthroplasty) criteria.

Implant Registries

Basically, these are large observational studies from a hospital, a region, one or several nations. Information about coverage and completeness of the data including recording of the primary procedure and outcome of interest (e.g., revision) must be available to judge the validity of results. The latter is measured as implant survival, or 1-survival, indicating revision. Methodology of measurements are generally done using Kaplan-Meier analysis or a competing risk analysis. The latter is largely comparable to the KM analysis, with the difference to account for a larger number of deaths which could "compete" with presence of "revision" and thus effect the KM estimator. A larger number of deaths occur in the elderly population and at longer follow-up (> 10 years) (see consensus EFORT section on registries). Ideally implant registries should have a coverage and data completeness of 80-100%. These registries have a signal function, but have confounding by indication as well, so no causal effects can be determined. Radiographs are routinely not included in the data capture of national registers but may be included in nested prospective randomized trials. A validated automated radiographic assessment, supported by artificial intelligence, could close this gap in the future. Besides, registries with a lower coverage, e.g. the German EPRD, can provide important information to draw conclusions on poorly performing implants ("signalling function"). While they can help to identify substandard implants, to prove that an implant performs very good a high quality registry with very high coverage is required.

Randomised controlled trials with a non-inferiority design

Implant fixation should be studied in prospective randomized studies with use of a well-documented standard implant as control. Definition of the control group is important. Controls should belong to a demographically similar patient group and the control implant should be intended for the same medical indication as applicable for the study group. The definition of a standard reference implant or benchmark implant should preferably be based on high validity implant registries (e.g. hip implants with a mean survival at 10 years of at least 90 or 95%, depending on the "state of art" for the implant type of interest). The new implant should be non-inferior to that benchmark within the margin of the error. If registry data are not available, sufficient number of patients included to detect deviations with relation to the aim of the implant (e.g. improved fixation, reduced liner wear, better range of motion, better pain relief or improved bone consolidation/remodelling etc. The resolution of the measurement technique determines the number needed in a non-inferiority design with the benchmark. Theoretical model simulations studies show that at least 6000 implants (90% study power) are needed to proof non-inferiority to the benchmark (Sayers et al. 2017). Objective, validated, methods with high resolution to quantify migration or implant loosening (e.g., RSA, EBRA or low dose CT-based implant motion analysis) should be used. Since these methods have high accuracy only few patients are needed in a RCT study design. Newer methods like computer-assisted measurements of spine movement based on image matching (FXA, ACES GmbH, Esslingen, Germany) and wear measurement of hip liners on radiographs should be preferred over manual evaluations. All these measures minimize number of patients exposed to the potential hazard of a new and unproven implants. Although hierarchy on accuracy, applicability to different implants and ease-to-use exists between these measurement techniques. Consistency in comparing implant migration within and between studies is essential. A minimum requirement is that the first radiograph is taken within 1-3 days after surgery, with the same loading condition over time. Study data from observational studies (including large registry studies) and RCTs should be interpreted simultaneously. Observational studies (including large registry studies) may by their hypothesis-generating nature contribute to the interpretation of RCTs after evaluating the likelihood of bias in in either of them. The most important factors with influence on bias are: 1. In- and exclusion criteria used; 2. Patient characteristics 3. Surgical skill 4. Positive effect of participation in (RCT) study Hawthorne effect); 5. Quality of data (lost follow-up, validity measurements, see AQUILA); 6. (double) Blinding; 7. Confounding (ASA, postop regimen etc.).

A phased evidence-based introduction of new implants that examines every possible mode of expected and unexpected failure will be a challenge. A TOI, Toolbox Orthopaedic Implants, could be used both for existing and new implants.

Preclinical testing

The clinical performance of an implant cannot or only poorly be predicted based on preclinical testing with finite element (FE) models. Although knowledge by preclinical testing can be important to have an idea on in vitro liner wear after appropriate testing, FE models were also in contrast to real-world patient data (SHP, scientific Hip prosthesis) as well as with highly accurate studies on implant migration (RSA) in patients (Huiskes 1998, Stolk 2002, 2007, Broeke 2013). The impact of FE models on implant behaviour has plateaued (Taylor 2015). This is due to the complexity of a real patient. A multitude of factors such as patient activity, variations of bone anatomy, bone quality, bone metabolism and remodelling and individual variations of inflammatory response to wear debris will influence the probability of long-lasting fixation. Other factors such as soft tissue trauma during surgery and implant positioned are also important. Further advances in FE models appear to be limited either by a lack of patient data on loading, biological response to populate the models or the need to better understand the fundamentals of the mechanical and biological processes (Taylor 2015, Bergmann 2001, Kutzner 2017, Kanavaki 2017, Krezter 2009). Although FE are used in the design, development and pre-clinical testing of total joint replacements, model simulations must include all factors mentioned above to account for the real-world variability in implant performance and fixation. The

power of these FE modelling is the ability to run multiple simulations, but it should be noted that the necessary improvement of current models is impeded due to lack of data.

In the future a continuing input from real world data (e.g. high resolution migration studies, different types of radiographic/CT/DXA studies and data from registries) could all be used in a machine learning algorithm to constantly improve these FE models in the future to better simulate implant performance in patients.

7. References

1. Keurentjes JC, Pijls BG, Van Tol FR, Mentink JF, Mes SD, Schoones JW, Fiocco M, Sedrakyan A, Nelissen RG. **Which implant should we use for primary total hip replacement? A systematic review and meta-analysis. J Bone Joint Surg Am. 2014 Dec 17;96 Suppl 1:79-97. doi: 10.2106/JBJS.N.00397. Review.**

2. Malchau H. **On the importance of stepwise introduction of new hip implant technology. Thesis. Göteborg University, Göteborg, Sweden; 1995**

3. Pijls BG, Kärrholm J, Malchau H, Nieuwenhuijse MJ, Valstar ER. **RSA and registries: the quest for phased introduction of new implants. J Bone Joint Surg Am. 2011 Dec 21;93 Suppl 3:62-5.**

4. Nieuwenhuijse MJ, Nelissen RG, Schoones JW, Sedrakyan A. **Appraisal of evidence base for introduction of new implants in hip and knee replacement: a systematic review of five widely used device technologies. BMJ. 2014 Sep 9;349:g5133. doi: 10.1136/bmj.g5133. Review.**

5. Pijls BG, Valstar ER, Nouta KA, Plevier JW, Fiocco M, Middeldorp S, Nelissen RG. **Early migration of tibial components is associated with late revision: a systematic review and meta-analysis of 21,000 knee arthroplasties. Acta Orthop. 2012 Dec;83(6):614-24. doi: 10.3109/17453674.2012.747052. Epub 2012 Nov 9. Review.**

6. Price AJ, Alvand A, Troelsen A, Katz JN, Hooper G, Gray A, Carr A, Beard D. **Knee replacement. Lancet. 2018 Nov 3;392(10158):1672-1682. doi: 10.1016/S0140-6736(18)32344-4. Review.**

7. Sedrakyan A, Campbell B, Merino JG et al. **IDEAL-D: a rational framework for evaluating and regulating the use of medical devices. BMJ 2016: 353: i2372**

8. van der Voort P, Pijls BG, Nieuwenhuijse MJ, Jasper J, Fiocco M, Plevier JW, Middeldorp S, Valstar ER, Nelissen RG. **Early subsidence of shape-closed hip arthroplasty stems is associated with late revision. A systematic review and meta-analysis of 24 RSA studies and 56 survival studies. Acta Orthop. 2015;86(5):575-85. doi: 10.3109/17453674.2015.1043832. Review.**

9. Hasan S, Marang-van de Mheen PJ, Kaptein BL, Nelissen RGHH, Pijls BG. **RSA-tested TKA Implants on Average Have Lower Mean 10-year Revision Rates Than Non-RSA-tested Designs. Clin Orthop Relat Res. 2020 Jun;478(6):1232-1241. doi: 10.1097/CORR.0000000000001209.**

10. Cowan JB, Mlynarek RA, Nelissen RG, et al. **Evaluation of Quality of Lower Limb Arthroplasty Observational Studies Using the Assessment of Quality in Lower Limb Arthroplasty (AQUILA) Checklist. J Arthroplasty. 2015 Sep;30(9):1513-7. doi: 10.1016/j.arth.2015.03.020. Epub 2015 Mar 31.**

11. Pijls BG, Dekkers OM, Middeldorp S, et al. **AQUILA: assessment of quality in lower limb arthroplasty. An expert Delphi consensus for total knee and total hip arthroplasty. BMC Musculoskelet Disord. 2011 Jul 22;12:173. doi: 10.1186/1471-2474-12-173.**

12. McCulloch P, Altman DG, Campbell WB, et al. **No surgical innovation without evaluation: the IDEAL recommendations. Lancet. 2009 Sep 26;374(9695):1105-12. doi: 10.1016/S0140-6736(09)61116-8.**

13. Schulze, Trautwein, Vordemvenne et al., **A method to perform spinal motion analysis from functional X-ray images, J Biomech 2011, Volume 44, Issue 9, 3 June 2011, Pages 1740-1746**

14. Haversath, M., Klebingat, S., die VITAS-Gruppe. et al., **Endoprosthetic wear analysis using virtual CAD-based radiographs. Orthopäde 2018, 47, 811–819.**

15. Klebingat S, Jäger M., Bien T et al., **Highly Accurate Determination of Hip Implant Wear, Cup, Anteversion and Inclination Through an AI Automated 2D-3D Registration: An In-Vitro Validation. JOR 2021, in review.**

16. Pearson AM, Spratt KF, Genuario J, McGough W, Kosman K, Lurie J, et al. **Precision of lumbar intervertebral measurements: does a computer-assisted technique improve reliability? Spine. 2011;36: 572–80**

17. Hey et al. **Slump sitting X-ray of the lumbar spine is superior to the conventional flexion view in assessing lumbar spine instability, 2017 Spine Journal**

18. Sayers A, Crowther MJ, Judge A, Whitehouse MR, Blom AW. **Determining the sample size required to establish whether a medical device is non-inferior to an external benchmark..BMJ Open. 2017 Aug 28;7(8):e015397. doi: 10.1136/bmjopen-2016-015397.**

19. Derbyshire B, Prescott RJ, Porter ML. **Notes on the use and interpretation of radiostereometric analysis. Acta Orthop. 2009 Feb;80(1):124-30. doi: 10.1080/17453670902807474.**

20. Huiskes R, Verdonschot N, Nivbrant B. **Migration, stem shape, and surface finish in cemented total hip arthroplasty. Clin Orthop Relat Res. 1998 Oct;(355):103-12. doi: 10.1097/00003086-199810000-00011.PMID: 9917595**

21. Broeke RH, Harings SE, Emans PJ, Jutten LM, Kessels AG, Geesink RG. **Randomized comparison between the cemented Scientific Hip Prosthesis and Omnifit: 2-year DEXA and minimum 10-year clinical follow-up. J Arthroplasty. 2013 Sep;28(8):1354-61. doi: 10.1016/j.arth.2012.09.024. Epub 2013 Feb 27.**

22. Stolk J, Maher SA, Verdonschot N, Prendergast PJ, Huiskes R. **Can finite element models detect clinically inferior cemented hip implants? Clin Orthop Relat Res. 2003 Apr;(409):138-50. doi: 10.1097/01.blo.0000058882.03274.5e.**

23. Stolk J, Janssen D, Huiskes R, Verdonschot N. **Finite element-based preclinical testing of cemented total hip implants. Clin Orthop Relat Res. 2007 Mar;456:138-47. doi: 0.1097/BLO.0b013e31802ba491**

24. Taylor M, Prendergast PJ. **Four decades of finite element analysis of orthopaedic devices: where are we now and what are the opportunities? J Biomech. 2015 Mar 18;48(5):767-78. doi: 10.1016/j.jbiomech.2014.12.019**.

25. Bergmann G, Deuretzbacher G, Heller M, Graichen F, Rohlmann A, Strauss J, Duda GN. **Hip contact forces and gait patterns from routine activities. J Biomech. 2001 Jul;34(7):859-71. doi: 10.1016/s0021-9290(01)00040-9.**

26. Kutzner I, Richter A, Gordt K, Dymke J, Damm P, Duda GN, Günzl R, Bergmann G. **Does aquatic exercise reduce hip and knee joint loading? In vivo load measurements with instrumented implants. PLoS One. 2017 Mar 20;12(3):e0171972. doi: 10.1371/journal.pone.0171972.**

27. Kanavaki AM, Rushton A, Efstathiou N, Alrushud A, Klocke R, Abhishek A, Duda JL. **Barriers and facilitators of physical activity in knee and hip osteoarthritis: a systematic review of qualitative evidence. .BMJ Open. 2017 Dec 26;7(12):e017042. doi: 10.1136/bmjopen-2017-017042.**

28. Kretzer JP, Kleinhans JA, Jakubowitz E, Thomsen M, Heisel C. **A meta-analysis of design- and manufacturing-related parameters influencing the wear behavior of metal-on-metal hip joint replacements. .J Orthop Res. 2009 Nov;27(11):1473-80. doi: 10.1002/jor.20921**

RESEARCH TOPIC 18

Radiologic Methods and Parameters to Estimate Primary Stability of Implant Fixation to the Bone – Discussing Recommended Methods and Time Points for Evaluating Subsidence/Loosening of Implant Components and Evaluating Implant Fixation Depending on the Implant And Fixation Material

Authors

Johan Kärrholm[1], Marcel Dreischarf[2], José Cordero-Ampuero[3], Petra Heesterbeek[4], Richard T. Mayer[5], Rob G.H.H. Nelissen[6]

[1]Department of Orthopaedics, Institute of Clinical Sciences, Sahlgrenska Academy, University of Gothenburg, Gothenburg, Sweden

[2]Berlin Institute of Health, Julius Wolff Institute for Biomechanics and Musculoskeletal Regeneration, Charité-Universitätsmedizin Berlin, Germany

[3]Hospital Universitario de La Princesa, Madrid, Spain

[4]Sint Maartenskliniek, Nijmegen, the Netherlands

[5]Johnson & Johnson, Office of the Chief Medical Officer New Brunswick NJ, Zuchwil, Switzerland

[6]Department of Orthopaedics, Leids Universitair Medisch Centrum, Leiden, Netherlands

1. Abstract

Question: Implant fixation

What are the radiologic methods and parameters to estimate primary stability of implant fixation to the bone?

1. What are recommended time points for evaluating subsidence/loosening of implant components?
2. How can be decided which method for clinical examination is best for evaluating implant fixation depending on the implant and fixation material?

Summary / Recommendation:

1. Methods to estimate primary stability of implant fixation to bone are RSA, CTMA, EBRA, Conventional radiography, FXA, MRI. Methods measuring implant migration should have high resolution to minimize patients at risk when a new implant is evaluated. Early postoperative implant migration is predictive for mechanical loosening. Sequential Radiostereophotogrammetric (Radiostereometric) analysis (RSA) is the most validated and accurate technique measuring implant migration with a resolution down to 0.1-0.2 mm translation and 0.1-.0.2 degrees rotation. However, classic RSA requires placement of bone markers and a specific radiographic set-up. Recently developed CT-based migration measurements may be used without markers and pilot studies have shown similar resolution as classic RSA. CT has high radiation exposure, but doses are becoming increasingly lower with current developments. Standard radiographic examinations manually evaluated may be used to measure migration but have a poor resolution and is therefore not recommended to measure implant migration of new and unproven implants. Special, validated, computer-assisted methods can be used to substantially increase accuracy and precision of implant migration on conventional X-rays (e.g., EBRA (Einzel Bild Roentgen Analyse) or FXA (Functional X-ray Analysis), applicable to hip, knee, or spine implants. Implant loosening and interface quality (an indirect measure of loosening) are routinely evaluated on standard radiographs or CT by sequential measurement of radiolucent lines (RLL) or focal osteolysis about the implant. Progressive radiolucencies completely surrounding an implant is regarded to indicate loosening especially if combined with migration.

Additional methods such as CT will add information for instance about the three-dimensional extension of osteolysis. MRI can add information about the metabolism at different locations of the interface and/or adverse reaction caused by debris, particles, and inflammatory reactions. Choice of additional methods is optional and should be related to the questions in focus.

2. A baseline examination after surgery and before discharge is required. Further choice of time points for follow-up is dependent on the resolution of the method used. Minimum is baseline, 1 and 2 years, and 1 examination in between baseline and 1 year. The research question determines the follow-up regime. Long term follow-up is advised until more evidence has been collected.

3. Measurements of early implant migration with high resolution at least up to 2 years is a well-documented method to predict loosening of cemented and uncemented cups (hip arthroplasty) cemented stems (hip arthroplasty), cemented and uncemented tibial components (knee arthroplasty) as well as spinal devices (e.g., interbody cages or total disc replacements). If only conventional radiography is used, larger study cohorts will be needed. Using computer-assisted algorithms for an objective evaluation is recommended.

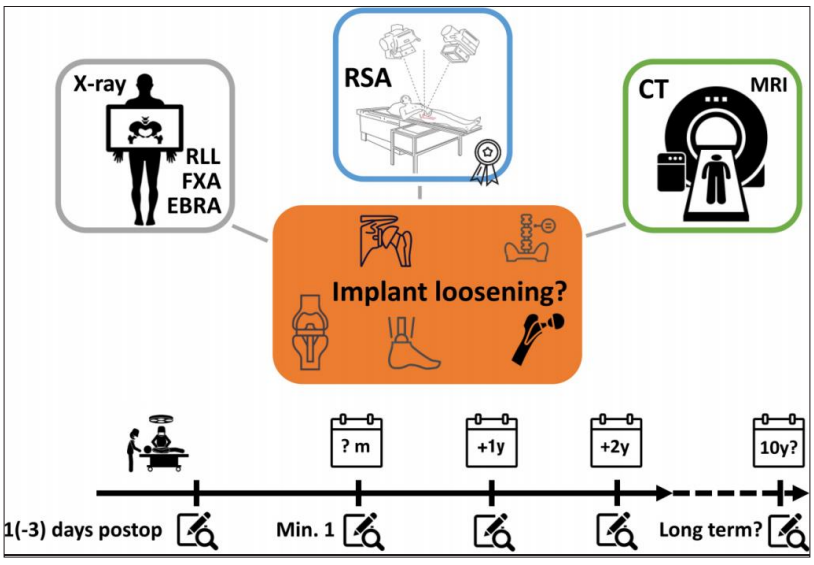
2. Level of Evidence

N/A

3. Consensus Delegate Vote

97% - unanimous, strongest consensus (97% agree / 0 disagree / 3% abstain)

4. Graphical Abstract



5. Search Strategy

N/A, narrative review.

6. Rationale

Conventional radiography

Target: All types of implants

Measurements of interest: Radiolucent lines, change of implant position (migration), osteolysis

Resolution: The development of progressive radiolucent lines (RLL) is considered an indirect radiographic sign of prosthetic loosening. If these lines surround the complete implant bone interface and are progressive in time, the implant is considered most likely loose (1, 2). It is important that contrast differences on radiographs between implant and bone can be an optical delusion (Mach sign). The development of radiolucent lines is a slow process that may last for years (3). In TKA, the Knee Society criteria use a width of >2mm as indicative of progression, no upper limit of RLL is defined.

Fluoroscopically assisted radiographs can improve the sensitivity of tibial component loosening compared to standard radiographs, but the clinical significance remains unclear.

In a comparative analysis using radiostereometric analysis (RSA) as reference, proximal-distal migration of the cup or the stem when measured manually on standard radiographs had an accuracy of 4-5 mm (4). The accuracy of conventional radiograph evaluations of acetabular cup migration can be improved down to 1-1.5 mm with use of specific techniques, such as EBRA (Einzel-Bild-Röntgen-Analyse). However, this technique needs standardized positioning of the hip between successive follow-ups (5, 6).

For spinal implants migration and loosening are typically assessed on conventional X-rays (lateral neutral and flexion/extension) under physiological loading. Thresholds for acceptable motions have been discussed with varying degree of consensus (e.g., < 2 ° range of motion, < 2 mm anterior-posterior migration) (7, 8). Computer-assisted methods (e.g., functional x-ray analysis - FXA) based on advanced image-registration techniques and AI have the potential to substantially improve the resolution (9) and even down to median offset of 0.1° and a maximal deviation of 0.31° (10).

Follow-up: Since implant migration is difficult to detect through manual measurements on conventional radiographs and secondary signs of migration often not become visible until 1-2 years have elapsed, this method is less suitable to evaluate new unproven implants.

Analysis should preferably be studied by comparison of radiographs taken directly postoperative and at subsequent follow up occasions. We suggest that examination with conventional radiography should be done postoperatively and at least at 2, 5, 10, 15 and 20 years. The projections/exposures chosen will vary depending on type of joint replacement studied.

Predictive value: There are studies on the predictive value of early migration measurements based on conventional radiography as regards risk of loosening (11-14). Such studies are scarce, probably due to the poor accuracy of conventional radiography to detect implant migration by manual evaluation. Thus, larger numbers of observations (>4000) are needed for valid conclusions. The clinical significance of RLL is well documented especially as regards cemented implants. Postoperative RLLs can be filled in during the postoperative 6 months or later, indicating that they are not predictive of loosening in early stages. One study (15) found that a novel percentage-based system (determining the percentage involvement of the tibial implant interface of any radiolucency at the bone-cement or cement-implant interface) was better able to predict implant loosening than the Knee Society criteria. Some studies point at the conclusion that RLLs are not predictable for later loosening especially as regards tibial components in TKRs, but the evidence is inconclusive.

Radiostereometric analysis (Roentgen stereophotogrammetric analysis), RSA

Target: All types of implants

Measurements of interest: Change of implant marker or implant position over time or at a single occasion with and without implant loading. Since marking of implants requires specific approval and may jeopardize the fatigue strength of the implant or its components the Model Based method (MB-RSA) using the 3D implant configuration is recommended. Markers are placed in bone, so that the position of the implant relative to the bone can be established precisely on dual X-rays using specialized RSA software at each examination. Changes of the implant position between examinations are calculated in three dimensions. Hence, RSA requires specialized equipment (software and reference box), markers, and experts to interpret results.

Micromotion of an implant can be expressed as translation along, and rotation around the three orthogonal axes. Summary measures total translation (TT =vectorial sum of translations), total rotation (TR =helical axis rotation) or the total translation of the most unstable point on the implant (Maximum Total Point Motion – MTPM). There is no strict consensus which parameters to report. Migration in the proximal-distal direction and MTPM are, however commonly used in studies of prediction of loosening.

Resolution: The precision and accuracy of radiostereometric analysis varies depending on implant size and configuration, scatter of bone markers (if used), bone configuration (if used) and radiographic set-up. For acetabular sockets, the reported precision values for translations varied between 0.06 and 0.62 mm, depending on implant geometry, and the direction of the analysed motions. Rotation varied within higher ranges, from 0.17 to 0.7. For stems, reported precision values for translation were 0.013 – 1.16, and for rotation 0.002 to 4.76, depending on implant, method (markers, femoral head centre (FHC) or use of MB-RSA or not), and direction of analysed motions (16-25). For the tibia component in TKA the precision values for translation vary between 0.03 and 0.4 mm depending on implant geometry and the direction of the analysed motions. Rotation varies within broader ranges, from 0.06 to 0.8 degrees (26). For the femoral component in TKA the precision values for translation vary between 0.15 and 0.42 mm, rotation between 0.20 – 0.67 degrees, depending on implant, and direction of analysed motions (27-29).

Validations of the model-based technique have revealed an accuracy or precision close to the one reported for marker-based RSA (30-33). Due to different geometry and smaller size, the accuracy and precision of upper extremity prostheses may differ (34). Based on a meta-regression-analysis of 23 studies these authors reported precision values for translations between 0.05 and 1.83 mm, depending on implant geometry, bone sizes studied and the direction of the analysed motions. Rotations varied within high ranges, starting from 0.1 degrees in the most optimum conditions. If analysed along the axis of symmetry for a certain type of shoulder component, for example, the error could increase to 10.7° and the precision of the determination of rotations became extremely poor on the same premise for one design of a trapeziometacarpal prosthesis. Fong et al. (35) reported high precision values of about 0.2 mm and 0.5 degrees for the talar component in an ankle joint prosthesis, which was the more asymmetrically shaped component, maximum errors of 2.75 mm for MTPM and 3.4 degrees for rotations when the more symmetrical tibial component was studied. Precision in a small cervical disc arthroplasty without markers can be challenging. In case of unacceptable precision on rotation parameters, only translation of the specific implant could be assessed (36). For certain components and directions of motions, the resolution of RSA when used to study the migration of small prostheses is of the same magnitude as that of larger implants.

In 2011, Sköldenberg and Odquist (37) evaluated the accuracy of marker-less RSA to study the Copeland humeral resurfacing head prosthesis. The algorithms developed by Börlin et al. (38) were used and marker-based RSA was used as a reference. The resolution of the marker-free RSA varied between 0.22 and 0.47 mm in terms of translations. The resolution of rotation measurements was surprisingly high (0.92° to 1.56°), probably because they included the small stem on this implant as an additional landmark. Ooms et al. (39) studied the precision (as defined above, presented as accuracy in the article) of RSA measurements of a trapeziometacarpal joint prosthesis, implanted into cadaveric hands. The maximum error for translation along the cardinal axes was 0.3 mm (longitudinal axis) for the metacarpal and 0.25 mm (transverse axis) for the trapezium component respectively. The corresponding errors for the determination of rotation were 3.3 and 1.8° (both longitudinal axis).

Follow-up: Most studies measured RSA directly postoperative (preferably the postoperative day, up to 3 days is acceptable as baseline), and 6 weeks, 3, 6, 12, and 24 months thereafter. Long-term follow-up is advised with further follow up at 5, 7.5 and 10 years.

Predictive value: The predictive value of implant component migration has been studied for hip arthroplasty (cemented and uncemented cups, cemented and uncemented stems – mainly composite beam type), knee arthroplasty (cemented and uncemented components) and osseointegrated implants for femoral amputees. These studies are based on long-term follow up of implants in single or compiled RSA studies or a combination of data from studies reporting early migration measured with RSA and long-term outcome according to clinical studies or national registers. All these studies have shown that increased early migration has a predictive value with varying resolution partly dependent on choice of motion parameter (Table 1).

Authors	Type of implant(s)	N[1]	Mean/median follow up in years, ±SD, *range*	Outcome	Measure	Estimated risk of failure
Hip replacements-primary cups						
Nieuwenhuijse et al. 2012 (54)	Exeter	41 C	9.4 ± 3.2, *3.1 - 12.0*	Revision cup loosening/radiographic loosening	Proximal migration at 2 years	HR = 19.9, *4.9-80* if exceeding 1.29 to 1.76 mm at two years
					Sagittal rotation (change of inclination) at 2 years	HR = 11.1, *2.83-43.9* if at least 1.85° to 2.53° at 2 years
Pijls et al. 2012 (55)	13 designs	(7) C, (6) UC	2 years: RSA studies 10 years: survival data	Revision rate due to loosening	Proximal migration at 2 years; HR per mm proximal migration	HR (unadjusted): 1.1, *1.05 – 1.14* HR adjusted for study quality: 1.06, *1.02-1.09* Acceptable: ≤0.2 mm; at risk: >0.2 – 1.0 mm; unacceptable: >1.0 mm
Johanson et al. 2017 (56)	4 designs	262 C	13.0 *2 – 24* and 15.3 *2-23*	Radiographic loosening and cup revision due to loosening	Proximal migration at 2 years; HR per mm subsidence	HR (adjusted for patient demographics and implant factors), radiographic loosening: 27, *5 – 154* HR (adjusted for patient demographics and implant factors), revision: 22, *3 - 161*
Hip replacements-revision cups						
Klerken et al. 2015 (57)	Various designs	244 C, 68 UC	*2-20*	Radiographic loosening and revision due to loosening	Proximal migration at 2 years; HR per mm proximal migration	HR (unadjusted): 1.37, *1.18 – 1.58* HR (adjusted for demographics, surgical factors and bone defects): 1.94 *1.31-2.88*
Hip replacements-primary stems						
Kärrholm et al. 1994 (58)	Lubinus SP	84 C	5.8	Stem revision due to loosening	Stem subsidence at 2 years	>50% if ≥1.2 mm, >95% if ≥2.4 mm
Johanson et al. 2016 (18)	Spectron EF Primary	244 C	14 *3-18*	Stem revision due to loosening or radiographic loosening	Stem subsidence at 2 years Stem retrotorsion at 2 years	HR (adjusted): 6.0 *2.5-15* per mm HR (adjusted): 5.1 *2.2-12* if ≥0.15 mm HR (adjusted): 1.7 *1.1-2.5* per degree
van der Voort et al. 2015 (59)	4 "shape-closed" designs	221 C (4) C	2 years: RSA studies 10 years: survival data	Aseptic revision rate or indication for revision surgery	Stem subsidence at 2 years; HR per 0.1 mm subsidence	HR (unadjusted): 1.04 *1.01- 1.07* HR (adjusted for study quality): 1.05 *1.03-1.08* Acceptable: ≤0.15 mm; at risk: >0.15 – 0.22 mm Unacceptable: >0.22 mm
Knee replacements-primary tibial components						
Ryd et al. 1995 (41)	7 uni- or bi-compartmental designs	143 C or UC	7, *2-11*	Revision due to loosening	MTPM between 1 and 2 years	MTPM >0.2mm between 1 and 2 years had a predictive power of 82% (sensitivity of 58% and specificity 93%)
Pijls et al. 2012 (40)	28 different designs	(18) C (10) UC	1 year: RSA studies 5 years: survival data	Revision rate due to loosening	MTPM at 1 year; HR per mm migration	HR (unadjusted): 1.07 *1.06 – 1.10* HR adjusted for study quality: 1.08, *1.06-1.09* Acceptable: ≤0.45 mm; At risk: >0.45 – 1.6 mm Unacceptable: >1.6 mm
Gudnason et al. 2017 (60)	3 designs	131 C	16 *14.8-17.4*	Revision due to loosening, of tibial component	Rotation around transverse axis (best predictive measure)	If ≥0.8°, specificity of 85%, sensitivity 50% (AUC:80%)

Table 1. Summary of studies that evaluated prediction of implant loosening resulting in revision based on RSA measurements of migration up to two years. Some of them have also included radiographic loosening. 1C=cemented, UC=uncemented, values in parenthesis refer to number of designs/types.

Measurement of interest: In acetabular cups proximal migration has been found to be the most sensitive parameter to predict later loosening of both cemented and uncemented implants (Table 1). For matte cemented stems (force closed) subsidence or distal migration is most sensitive (18). The predictive value of early migration measurements of polished cemented and uncemented stems has not been firmly established.

For tibial components in total knee arthroplasty the parameter MTPM is often used. Studies showed that most migration occurred during the first year, especially in the first 3 months after surgery and stabilized thereafter. In a systematic review Pijls et al. (40) demonstrated that for every mm increase in migration, there was an 8% increase in revision rate and that migration between 0.5 and 1.6 mm during the first postoperative five years were considered to indicate a revision rates higher than 5% at 10 years. Progressive migration between first and second year (>0.2mm) has also been considered to be predictive of later aseptic loosening (41), but the lower limit of 0.2 mm has been debated (42). Laende et al. (43, 44) showed that the migration at 10-year follow-up was strongly correlated to both one- and two-year migration. Thus, studies available suggests that both the magnitude of migration at 1 year and between 1 and 2 years could be used to predict loosening. There is however a debate within this field since some authors think that alternative measures to MTPM should be considered, since this parameter can be associated with large errors (42). Furthermore, for femoral components there are no studies of the significance and potential predictive value of early migration measurements.

CT-based implant motion analysis

Target: all types of implants

Measurement of interest: Radiolucent lines, implant loosening

Computed tomography (CT) has the inherent potential to be used for determining marker position, bone, and surface geometry in three dimensions. CT is used to evaluate radiolucency, degree of fusion or bony integration. Detection of radiolucent lines surrounding cementless femoral stems may be facilitated with use of digital tomosynthesis or CT with metal artifact reduction (45, 46), which may be used more often when low-dose CT scans reduce radiation dose to the level of conventional radiographs.

Resolution: In 2017, Otten et al. (47) compared CT with the standard RSA technique based on simultaneous exposures of two roentgen tubes in 46 cups supplied with tantalum markers. Implant and bone marker positions were located using both methods. The limits of agreement between the two methods varied between 0.56 and 0.74 mm and 1.05 and 2.17 degrees, which at that time was regarded as acceptable, since these errors were slightly above their precision using the standard technique. Later Brodén et al. (48) used a commercial image-registration software named "CT-based implant Motion Analysis" (CTMA, Sectra Lidköping) and studied the precision of cup migration measurements. Studies were done both with and without use of bone markers. Without bone markers the precision ranged between 0.07 and 0.31 mm depending on direction of the translation. The mean effective dose varied between 0.2 and 2.3mSv depending on CT scan used. In a later study, the precision of both CTMA and "classic RSA" were studied on 10 patients with a cemented cup and the authors found that there was no difference regardless if bone markers were used or not for the CTMA-based measurements. The effective radiation dose was however five times higher with use of CTMA. Brodén et al. (48) and Scheerlinck et al. (49) also developed a marker-free automated CT-based system to measure the migration of a hip stem (CTSA) by the separate segmentation of the implant and the surrounding bone. Under ideal circumstances, this method had about the same resolution as conventional RSA.

Cone-beam computerized tomography (CBCT) is a new technique dedicated to musculoskeletal imaging. Compared to conventional multidetector CT, CBCT has a better image quality, is less sensitive to artifacts, and has a lower radiation dose. It can be an added value with reliable and reproducible results for the diagnosis of implant loosening (50).

Follow-up: Same as for studies using classic or model-based RSA

Predictive value: Same as for studies using classic or model-based RSA

In the future, CT-based implant motion analysis might become a viable option to RSA-based analysis on convention radiographs. If the metallic artefacts can be sufficiently reduced, as appears to be the case, there will be no need for either markers or three-dimensional models. This or similar techniques will eliminate most of the problems associated with RSA evaluations, problems related to fitting a CAD or reversed engineered model, but they will not be able to compensate for problems related to implant symmetry. In addition, studying dynamic implant motion is complicated using CT, and the radiation dose is still an issue. In summary, recently presented techniques appear fascinating, but so far they have only been tested in solitary clinical studies and their future role needs to be more firmly established (51).

Magnetic Resonance Imaging (MRI)

Target: all types of implants

Measurements of interest: Tissue reactions

Evidence for the use of MRI for the diagnosis of implant loosening of the knee is scarce (52). This method can identify tissue reactions which may be consistent with component loosening, corrosion, or adverse reactions to metallic debris. Image-based component rotational analysis may be performed to determine the axial or transverse plane angulation between implant and bone. Suboptimal rotational alignment has for example been implicated in failure (53).

7. References

1. Barrack RL, Mulroy RD, Jr., Harris WH. **Improved cementing techniques and femoral component loosening in young patients with hip arthroplasty. A 12-year radiographic review. The Journal of bone and joint surgery British volume. 1992;74(3):385-9.**

2. Hodgkinson JP, Shelley P, Wroblewski BM. **The correlation between the roentgenographic appearance and operative findings at the bone-cement junction of the socket in Charnley low friction arthroplasties. Clinical orthopaedics and related research. 1988(228):105-9.**

3. Hultmark P, Kärrholm J, Strömberg C, Herberts P, Möse CH, Malchau H. **Cemented first-time revisions of the femoral component: prospective 7 to 13 years' follow-up using second-generation and third-generation technique. The Journal of arthroplasty. 2000;15(5):551-61.**

4. Malchau H, Kärrholm J, Wang YX, Herberts P. **Accuracy of migration analysis in hip arthroplasty. Digitized and conventional radiography, compared to radiostereometry in 51 patients. Acta orthopaedica Scandinavica. 1995;66(5):418-24.**

5. Krismer M, Bauer R, Tschupik J, Mayrhofer P. **EBRA: a method to measure migration of acetabular components. Journal of biomechanics. 1995;28(10):1225-36.**

6. Biedermann R, Krismer M, Stöckl B, Mayrhofer P, Ornstein E, Franzén H. **Accuracy of EBRA-FCA in the measurement of migration of femoral components of total hip replacement. Einzel-Bild-Röntgen-Analyse-femoral component analysis. The Journal of bone and joint surgery British volume. 1999;81(2):266-72.**

7. McAfee PC, Boden SD, Brantigan JW, Fraser RD, Kuslich SD, Oxland TR, et al. **Symposium: a critical discrepancy-a criteria of successful arthrodesis following interbody spinal fusions. Spine. 2001;26(3):320-34.**

8. Oshina M, Oshima Y, Tanaka S, Riew KD. **Radiological Fusion Criteria of Postoperative Anterior Cervical Discectomy and Fusion: A Systematic Review. Global spine journal. 2018;8(7):739-50.**

9. Pearson AM, Spratt KF, Genuario J, McGough W, Kosman K, Lurie J, et al. **Precision of lumbar intervertebral measurements: does a computer-assisted technique improve reliability? Spine. 2011;36(7):572-80.**

10. Schulze M, Trautwein F, Vordemvenne T, Raschke M, Heuer F. **A method to perform spinal motion analysis from functional X-ray images. Journal of biomechanics. 2011;44(9):1740-6.**

11. Grewal R, Rimmer MG, Freeman MA. **Early migration of prostheses related to long-term survivorship. Comparison of tibial components in knee replacement. The Journal of bone and joint surgery British volume. 1992;74(2):239-42.**

12. Freeman MA, Plante-Bordeneuve P. **Early migration and late aseptic failure of proximal femoral prostheses. The Journal of bone and joint surgery British volume. 1994;76(3):432-8.**

13. Abrahams JM, Kim YS, Callary SA, De Ieso C, Costi K, Howie DW, et al. **The diagnostic performance of radiographic criteria to detect aseptic acetabular component loosening after revision total hip arthroplasty. The bone & joint journal. 2017;99-b(4):458-64.**

14. Kim YS, Abrahams JM, Callary SA, De Ieso C, Costi K, Howie DW, et al. **Proximal translation of > 1 mm within the first two years of revision total hip arthroplasty correctly predicts whether or not an acetabular component is loose in 80% of cases: a case-control study with confirmed intra-operative outcomes. The bone & joint journal. 2017;99-b(4):465-74.**

15. Chalmers BP, Sculco PK, Fehring KA, Trousdale RT, Taunton MJ. **A Novel Percentage-Based System for Determining Aseptic Loosening of Total Knee Arthroplasty Tibial Components. The Journal of arthroplasty. 2017;32(7):2274-8.**

16. Aro HT, Alm JJ, Moritz N, Mäkinen TJ, Lankinen P. **Low BMD affects initial stability and delays stem osseointegration in cementless total hip arthroplasty in women: a 2-year RSA study of 39 patients. Acta orthopaedica. 2012;83(2):107-14.**

17. Flivik G, Kristiansson I, Kesteris U, Ryd L. **Is removal of subchondral bone plate advantageous in cemented cup fixation? A randomized RSA study. Clinical orthopaedics and related research. 2006;448:164-72.**

18. Johanson PE, Antonsson M, Shareghi B, Kärrholm J. **Early Subsidence Predicts Failure of a Cemented Femoral Stem With Minor Design Changes. Clinical orthopaedics and related research. 2016;474(10):2221-9.**

19. Kärrholm J, Malchau H, Snorrason F, Herberts P. **Micromotion of femoral stems in total hip arthroplasty. A randomized study of cemented, hydroxyapatite-coated, and porous-coated stems with roentgen stereophotogrammetric analysis. The Journal of bone and joint surgery American volume. 1994;76(11):1692-705.**

20. Kiernan S, Geijer M, Sundberg M, Flivik G. **Effect of symmetrical restoration for the migration of uncemented total hip arthroplasty: a randomized RSA study with 75 patients and 5-year follow-up. Journal of orthopaedic surgery and research. 2020;15(1):225.**

21. Luites JW, Spruit M, Hellemondt GG, Horstmann WG, Valstar ER. **Failure of the uncoated titanium ProxiLock femoral hip prosthesis. Clinical orthopaedics and related research. 2006;448:79-86.**

22. Mjöberg B, Selvik G, Hansson LI, Rosenqvist R, Onnerfält R. **Mechanical loosening of total hip prostheses. A radiographic and roentgen stereophotogrammetric study. The Journal of bone and joint surgery British volume. 1986;68(5):770-4.**

23. Onsten I, Carlsson AS, Sanzén L, Besjakov J. **Migration and wear of a hydroxyapatite-coated hip prosthesis. A controlled roentgen stereophotogrammetric study. The Journal of bone and joint surgery British volume. 1996;78(1):85-91.**

24. Shareghi B, Johanson PE, Kärrholm J. **Clinical evaluation of model-based radiostereometric analysis to measure femoral head penetration and cup migration in four different cup designs. Journal of orthopaedic research : official publication of the Orthopaedic Research Society. 2017;35(4):760-7.**

25. Thien TM, Thanner J, Kärrholm J. **Randomized comparison between 3 surface treatments of a single anteverted stem design: 84 hips followed for 5 years. The Journal of arthroplasty. 2010;25(3):437-44.e1.**

26. Pijls BG, Plevier JWM, Nelissen R. **RSA migration of total knee replacements. Acta Orthop. 2018;89(3):320-8.**

27. Henricson A, Wojtowicz R, Nilsson KG, Crnalic S. **Uncemented or cemented femoral components work equally well in total knee arthroplasty. Knee Surg Sports Traumatol Arthrosc. 2019;27(4):1251-8.**

28. Gao F, Henricson A, Nilsson KG. **Cemented versus uncemented fixation of the femoral component of the NexGen CR total knee replacement in patients younger than 60 years: a prospective randomised controlled RSA study. The Knee. 2009;16(3):200-6.**

29. Heesterbeek PJ, Wymenga AB, van Hellemondt GG. **No Difference in Implant Micromotion Between Hybrid Fixation and Fully Cemented Revision Total Knee Arthroplasty: A Randomized Controlled Trial with Radiostereometric Analysis of Patients with Mild-to-Moderate Bone Loss. The Journal of bone and joint surgery American volume. 2016;98(16):1359-69.**

30. Kaptein BL, Valstar ER, Stoel BC, Rozing PM, Reiber JH. **A new model-based RSA method validated using CAD models and models from reversed engineering. Journal of biomechanics. 2003;36(6):873-82.**

31. Valstar ER, de Jong FW, Vrooman HA, Rozing PM, Reiber JH. **Model-based Roentgen stereophotogrammetry of orthopaedic implants. Journal of biomechanics. 2001;34(6):715-22.**

32. Hurschler C, Seehaus F, Emmerich J, Kaptein BL, Windhagen H. **Accuracy of model-based RSA contour reduction in a typical clinical application. Clinical orthopaedics and related research. 2008;466(8):1978-86.**

33. Hurschler C, Seehaus F, Emmerich J, Kaptein BL, Windhagen H. **Comparison of the model-based and marker-based roentgen stereophotogrammetry methods in a typical clinical setting. J Arthroplasty. 2009;24(4):594-606.**

34. Ten Brinke B, Beumer A, Koenraadt KLM, Eygendaal D, Kraan GA, Mathijssen NMC. **The accuracy and precision of radiostereometric analysis in upper limb arthroplasty. Acta orthopaedica. 2017;88(3):320-5.**

35. Fong JW, Veljkovic A, Dunbar MJ, Wilson DA, Hennigar AW, Glazebrook MA. **Validation and precision of model-based radiostereometric analysis (MBRSA) for total ankle arthroplasty. Foot & ankle international. 2011;32(12):1155-63.**

36. van Hooff ML, Heesterbeek PJC, Spruit M. **Mechanical Stability of the Prodisc-C Vivo Cervical Disc Arthroplasty: A Preliminary, Observational Study Using Radiostereometric Analysis. Global spine journal. 2020;10(3):294-302.**

37. Sköldenberg O, Odquist M. **Measurement of migration of a humeral head resurfacing prosthesis using radiostereometry without implant marking: an experimental study. Acta orthopaedica. 2011;82(2):193-7.**

38. Börlin N, Röhrl SM, Bragdon CR. **RSA wear measurements with or without markers in total hip arthroplasty. Journal of biomechanics. 2006;39(9):1641-50.**

39. Ooms EM, ten Brinke B, Mathijssen NM, Blom IF, Deijkers RL, Kraan GA. **Feasibility of model-based Roentgen Stereophotogrammetric Analysis to evaluate early migration of the trapeziometacarpal joint prosthesis. BMC musculoskeletal disorders. 2015;16:295.**

40. Pijls BG, Valstar ER, Nouta KA, Plevier JW, Fiocco M, Middeldorp S, et al. **Early migration of tibial components is associated with late revision: a systematic review and meta-analysis of 21,000 knee arthroplasties. Acta orthopaedica. 2012;83(6):614-24.**

41. Ryd L, Albrektsson BE, Carlsson L, Dansgård F, Herberts P, Lindstrand A, et al. **Roentgen stereophotogrammetric analysis as a predictor of mechanical loosening of knee prostheses. The Journal of bone and joint surgery British volume. 1995;77(3):377-83.**

42. Niesen AE, Hull ML. **Propagation of registration errors into the change in maximum total point motion for determining stability of tibial baseplates. Computer methods in biomechanics and biomedical engineering. 2021:1-7.**

43. Laende EK, Astephen Wilson JL, Mills Flemming J, Valstar ER, Richardson CG, Dunbar MJ. **Equivalent 2-year stabilization of uncemented tibial component migration despite higher early migration compared with cemented fixation: an RSA study on 360 total knee arthroplasties. Acta orthopaedica. 2019;90(2):172-8.**

44. Laende EK, Richardson CG, Dunbar MJ. **Predictive value of short-term migration in determining long-term stable fixation in cemented and cementless total knee arthroplasties. The bone & joint journal. 2019;101-b(7_Supple_C):55-60.**

45. Gillet R, Teixeira P, Bonarelli C, Coudane H, Sirveaux F, Louis M, et al. **Comparison of radiographs, tomosynthesis and CT with metal artifact reduction for the detection of hip prosthetic loosening. European radiology. 2019;29(3):1258-66.**

46. Tang H, Huang X, Cheng X, Yang D, Huang Y, Zhou Y. **Evaluation of peri-prosthetic radiolucent lines surrounding the cementless femoral stem using digital tomosynthesis with metal artifact reduction: a cadaveric study in comparison with radiography and computed tomography. Quantitative imaging in medicine and surgery. 2020;10(9):1786-800.**

47. Otten V, Maguire GQ, Jr., Noz ME, Zeleznik MP, Nilsson KG, Olivecrona H. **Are CT Scans a Satisfactory Substitute for the Follow-Up of RSA Migration Studies of Uncemented Cups? A Comparison of RSA Double Examinations and CT Datasets of 46 Total Hip Arthroplasties. BioMed research international. 2017;2017:3681458.**

48. Brodén C, Sandberg O, Olivecrona H, Emery R, Sköldenberg O. **Precision of CT-based micromotion analysis is comparable to radiostereometry for early migration measurements in cemented acetabular cups. Acta orthopaedica. 2021:1-5.**

49. Scheerlinck T, Polfliet M, Deklerck R, Van Gompel G, Buls N, Vandemeulebroucke J. **Development and validation of an automated and marker-free CT-based spatial analysis method (CTSA) for assessment of femoral hip implant migration: In vitro accuracy and precision comparable to that of radiostereometric analysis (RSA). Acta orthopaedica. 2016;87(2):139-45.**

50. Dartus J, Jacques T, Martinot P, Pasquier G, Cotten A, Migaud H, et al. **The advantages of cone-beam computerised tomography (CT) in pain management following total knee arthroplasty, in comparison with conventional multi-detector CT. Orthopaedics & traumatology, surgery & research : OTSR. 2021;107(3):102874.**

51. Röhrl SM. **"Great balls on fire:" known algorithm with a new instrument? Acta orthopaedica. 2020;91(6):621-3.**

52. Schröder FF, Post CE, Wagenaar FBM, Verdonschot N, Huis In't Veld R. **MRI as Diagnostic Modality for Analyzing the Problematic Knee Arthroplasty: A Systematic Review. Journal of magnetic resonance imaging : JMRI. 2020;51(2):446-58.**

53. Koff MF, Burge AJ, Potter HG. **Clinical magnetic resonance imaging of arthroplasty at 1.5T. Journal of orthopaedic research : official publication of the Orthopaedic Research Society. 2020;38(7):1455-64.**

54. Nieuwenhuijse MJ, Valstar ER, Kaptein BL, Nelissen RG. **Good diagnostic performance of early migration as a predictor of late aseptic loosening of acetabular cups: results from ten years of follow-up with Roentgen stereophotogrammetric analysis (RSA). The Journal of bone and joint surgery American volume. 2012;94(10):874-80.**

55. Pijls BG, Nieuwenhuijse MJ, Fiocco M, Plevier JW, Middeldorp S, Nelissen RG, et al. **Early proximal migration of cups is associated with late revision in THA: a systematic review and meta-analysis of 26 RSA studies and 49 survivalstudies. Acta orthopaedica. 2012;83(6):583-91.**

56. Johanson PE, Furnes O, Ivar Havelin L, Fenstad AM, Pedersen AB, Overgaard S, et al. **Outcome in design-specific comparisons between highly crosslinked and conventional polyethylene in total hip arthroplasty. Acta orthopaedica. 2017;88(4):363-9.**

57. Klerken T, Mohaddes M, Nemes S, Kärrholm J. **High early migration of the revised acetabular component is a predictor of late cup loosening: 312 cup revisions followed with radiostereometric analysis for 2-20 years. Hip international : the journal of clinical and experimental research on hip pathology and therapy. 2015;25(5):471-6.**

58. Kärrholm J, Borssén B, Löwenhielm G, Snorrason F. **Does early micromotion of femoral stem prostheses matter? 4-7-year stereoradiographic follow-up of 84 cemented prostheses. The Journal of bone and joint surgery British volume. 1994;76(6):912-7.**

59. van der Voort P, Pijls BG, Nieuwenhuijse MJ, Jasper J, Fiocco M, Plevier JW, et al. **Early subsidence of shape-closed hip arthroplasty stems is associated with late revision. A systematic review and meta-analysis of 24 RSA studies and 56 survival studies. Acta orthopaedica. 2015;86(5):575-85.**

60. Gudnason A, Adalberth G, Nilsson KG, Hailer NP. **Tibial component rotation around the transverse axis measured by radiostereometry predicts aseptic loosening better than maximal total point motion. Acta orthopaedica. 2017;88(3):282-7.**

RESEARCH TOPIC

JOINT STABILITY AND KINEMATICS

Research Topic 19

Assessment of Functional Joint Stability and Movement Performance after Total Joint Arthroplasty Regarding the Ability of an Implant to Enable the Reconstruction of a Functionally Satisfying and Stable Joint, Including an Appropriate Range of Motion and Best Possible Preservation / Restoration of Kinematics

RESEARCH TOPIC 19

Assessment of Functional Joint Stability and Movement Performance after Total Joint Arthroplasty Regarding the Ability of an Implant to Enable the Reconstruction of a Functionally Satisfying and Stable Joint, Including an Appropriate Range of Motion and Best Possible Preservation / Restoration of Kinematics

Authors

William R. Taylor[1], Bernardo Innocenti[2], Georg N. Duda[3], Thomas M. Grupp[4,5], Matthias Woiczinski[5]

[1]Institute for Biomechanics, ETH Zürich, Zürich, Switzerland
[2]BEAMS Department, Université Libre de Bruxelles, Belgium
[3]Julius Wolff Institute and Center for Musculoskeletal Surgery, Berlin Institute of Health at Charité - Universitätsmedizin Berlin, Germany
[4]Aesculap AG Research & Development, Tuttlingen, Germany
[5]Ludwig Maximilians University Munich, Department of Orthopaedic and Trauma Surgery, Musculoskeletal University Center Munich (MUM), Campus Grosshadern, Munich, Germany

1. Abstract

Question: Joint stability and kinematics

How can one prove that an implant enables the reconstruction of a functionally satisfying and stable joint, including an appropriate range of motion and best possible preservation / restoration of kinematics?

One of the challenges facing implant designers, clinicians, and researchers alike for the improved development of implants that are subject to the Medical Device Regulation is the assessment and understanding of stability and functional outcome of the joint. Such insight is essential for providing key perspectives on each individual's clinical outcome, and therefore for the long-term success of the product. Here, capturing the movement and loading patterns of a joint during functional activities and then understanding the consequences of these dynamic measurements on clinical outcome, however, is inherently challenging. Hence, the focus of this white paper is to present a consensus statement on the current state-of-the-art in assessing functional joint stability, and movement performance after total joint arthroplasty, specifically addressing the question: "How can one prove that an implant enables the reconstruction of a functionally satisfying and stable joint, including an appropriate range of motion and best possible preservation / restoration of kinematics?"

Summary / Recommendation:

With novel "at home systems" (e.g. wearables / phone cameras etc.) now commonplace, as well as kinematic assessment equipment (gait analysis and mobile fluoroscopy systems etc.) becoming more widely available, the monitoring of patient functional outcome, joint stability and mobility has become accessible for the clinical assessment of joint performance. Combined with frequent and regular data collection using digital technologies and smart device applications, such functional metrics (gait speed, movement symmetry, limb coordination, heel strike impact, range of joint motion, mobility, number of steps, frequency of movement etc.) can be easily collected, incorporated into open registries, and used for clinical evaluation of implant functional performance. The capture of digital datasets on mobility and functionality are scalable and can lay the foundations for not only killing-off ineffective or inept products, but also provide strong evidence for demonstrating functional outcomes and differences between performance in implant designs.

2. Level of Evidence

n/a

3. Consensus Delegate Vote

98% - unanimous, strongest consensus (98% agree / 2% disagree / 0% abstain)

4. Graphical Abstract

n/a

5. Rationale

The term "joint stability" generally encompasses the condition from passively stiff to lax joints, and can explain why some joints are more prone to dislocation or injury than others. Since pain, overloading, and degeneration are key consequences of an overly stiff or lax joint, the stability of a joint provides a fundamental basis driving clinical decision-making and treatment. Particularly, the interactions between articular contact, passive soft tissue structures, and active neuromuscular control play key roles in governing the constraints on the force and moment balance within the joint. Moreover, after joint arthroplasty, the implant design and its surgical implantation parameters alter the articular contact mechanics and natural strain environment of the soft tissues. On the other hand, the term "kinematics" addresses the movement patterns of the articulating components within a joint, and provides a key metric for understanding the overall functional ability of the joint. Importantly, after total joint arthroplasty (TJA), the active and passive stability and kinematic interactions therefore become critical for producing stable functional movement patterns without overloading the implant-bone or implant-implant interfaces [2], or the surrounding soft tissues throughout different activities of daily living [3-5].

In order to comprehend the success or otherwise of a specific implant design from a stability and functional perspective, it is critical to both assess and understand joint stability throughout activities of daily living, where patients often report difficulties at certain ranges of motion, or when undertaking specific tasks. Here, it is important to consider the active muscle forces that induce considerable loads and translations within the joint [6-9]. On the other hand, passive structures (tendon, ligaments, menisci, capsule etc.) generate increasing forces only when translated or deformed from their relaxed state. As a result, clinical testing of a passive joint will not generate the same conditions as those achieved in vivo during dynamic and functional joint movements. Consequently, there is now a drive to capture and evaluate joint kinematics and kinetics throughout functional activities of daily living, such that loading and unloading events, as well as precise assessment of the changing six degree of freedom joint movement patterns, are all considered within the evaluation of functional outcome.

Although the target of joint replacement is to reduce pain and restore the dynamic functionality of a pathologic joint, the current approach in the Med-Tech industry to demonstrate suitability of a new implant is to demonstrate "substantial equivalence" to other similar products that are available on the market. However, this approach completely ignores any evaluation of kinematics of the joint, and dynamic functional outcome of the patient. As such, a key requirement for medical device regulation critically needs to address the functional needs of a patient and how an implant performs in a dynamic environment. The critical question is therefore: how is it possible to provide sufficient evidence of a suitably stable

joint reconstruction and satisfactory functional outcomes in pre-clinical testing, including features such as appropriate range of motion or best possible bone stock preservation? This challenge is further exacerbated by the need for design innovation to be promoted rather than suppressed by excessive regulation. Here, "substantial equivalence" to existing implant designs that leave some 30% of patients dissatisfied should not be the goal. The answers to this challenge lie in collaborative partnerships between the regulatory bodies and the key medical stakeholders, including clinicians, the orthopaedic industry, and research partners, in order to ensure that all aspects of design, innovation, materials, biocompatibility, longevity, and functionality are appropriately addressed. This process undoubtedly requires iterative feedback from clinical outcomes back to industry to promote relevant incremental design improvement as well as kill-off ineffective or inept products. However, such processes critically need to be based on objective testing, standardised evaluation metrics, and clear and transparent reporting procedures. In return, medical device regulations need to provide sufficient space for outliers or special cases in order to support innovation.

One important aspect is that a manufacturer's product-specific claims on outcomes should only be made once sufficient scientific evidence on the pre-clinical development and testing processes have been provided. Here, it is essential that functional testing is not only appropriate but also sufficiently strong to back up any claims made. While claims based on longevity, fatigue, component wear etc. tend to be based on modelling and mechanical testing studies, the evaluation of clinical outcome and patient functional satisfaction tends to be based on clinical appraisal and self-reported data. However, the assessment, evaluation, and reporting of functional kinematics and joint stability is far more challenging, and has not yet found consensus in the clinical or research communities.

A number of emerging technologies exist for assessing dynamic functionality and joint kinematics. These include the use of gait laboratory facilities for functional testing, which generally measure the movement of reflective markers attached to the subject's skin using an opto-electronic motion capture system. Here, general metrics of gait speed, stride time and length etc., as well as characteristics that are less visible to the naked eye such as variability, balance, coordination, symmetry etc. can all be monitored throughout complete cycles of daily activities, but the technique is time consuming and subject to artefact resulting from the skin moving relative to the underlying skeletal segments. Access to accurate internal joint translation and rotations is altogether more challenging, however, with current methods typically limited to ultrasound imaging, or video fluoroscopy. Both of these techniques are limited spatially in their restricted field of view, but are now being further developed through navigated approaches and mobile tracking units to allow the assessment of joints throughout complete cycles of functional activities. Such technologies should be embraced as methods for supporting and informing clinical evaluation, implant performance, reporting and feedback processes.

In vivo, in vitro and in silico scientific investigations into joint motion have demonstrated different kinematics for different implant designs [10]. Of importance to note, however, is that different implant designs does not mean different company designs (or company brands), but rather the functional restriction of different prosthesis geometries in the joint, which have been developed and iterated over decades. As a result, implant manufacturers continually develop new concepts and features to gain further benefits for improved stability and kinematic performance outcomes. The final in vivo functionality of a joint, however, is critically dependent on many additional factors beyond the implant itself. It is generally accepted that the relationship between design, pre-operative joint status, surgical implantation, and post-operative functional outcome is not well understood. While the assessment and reporting of joint kinematics and stability outcomes is the key focus of this consensus document, it is important to acknowledge the contribution of additional techniques for understanding joint motion:

In vitro testing

Joint Testing Rigs

In a first pre-clinical development scenario, different in vitro approaches are useful for determining the likely outcome of a newly proposed implant design, including artificial or cadaveric bone implantations, using loading and boundary conditions imposed on quasi-static test rigs. Due to the limited transferability of the results from these simple biomechanical setups, such tests are most useful at the early stages of implant development.

After the development process is completed, but if possible before the total design freeze, dynamic in vitro and in silico studies offer useful approaches for analysing knee kinematics, and therefore also assessments of e.g. implant-bone interface stability. For in vitro studies, dynamic weight bearing test-rigs with human cadaveric joints can offer several advantages. Soft tissue support, anatomical variation, and the local mechanical environment can all be studied under comparable conditions to those found in vivo. The most important benefit of such testing under weight-bearing conditions, however, is that different implant designs or surgical techniques can be tested within an individual specimen to allow an understanding of the resulting joint kinematics. To represent the boundary conditions as closely as possible to the in vivo joint [11], a few general points have to be taken into account:

For knee joint testing particularly, most weight bearing test set-ups are based on the concept of the Oxford Knee Rig but have variations regarding the boundary conditions within different research groups. Importantly, the knee joint should possess six degrees of freedom to avoid artificial constraining forces, which would result in misleading knee kinematics. Moreover, relative mediolateral translation (ankle or hip) should be carefully considered, at least at the beginning of the test, to avoid incorrect varus/valgus adaptation of the knee. As most knee rigs are muscle controlled, at least the quadriceps should be actively loaded throughout the whole flexion of the knee. Further muscle representation (e.g. hamstrings) should at least be statically loaded or, even better, actively represented [9]. If possible, test rigs should be set-up by a proficient technician/engineer, while specimen preparation should be performed using original instruments and by an experienced surgeon or anatomist in order to ensure a consistent and repeatable test environment.

Within such test rigs, joint kinematics should be measured using a 3D measurement system and based on established coordinate systems relative to anatomical landmarks [12]. Care should be taken if data sets generated using different landmarks or calculation methods are to be compared. For example, the flexion facet centre is a good anatomical landmark for tibio-femoral joint because it remains consistent throughout flexion, due to its relationship with the anatomical epicondyles of the femur. However, it is important to note that this rarely represents the true functional flexion axis of the joint during dynamic activities of daily living in vivo [13, 14], and will therefore likely introduce secondary errors in the estimated translation within the joint.

Weight bearing knee rigs are a powerful tool to study different knee characteristics in a standardized environment. Currently, most rigs are limited to knee bending activities with only limited muscle load control, which is a likely reason for their poor replication of functional knee kinematics. In general, research should aim to minimise these issues through e.g. integrating the loading conditions generated using validated computer models.

In silico models

Another technique that should be implemented to support implant design/development is computational modelling and simulation in silico, with the goal to provide useful information on implant kinematics, kinetics, and stresses in the components. In general, computational models used in orthopaedic biomechanics can be classified into: (1) musculoskeletal models, based on multibody dynamics, and (2) finite element (FE) models based on the equations of continuum mechanics. Musculoskeletal models are typically used to calculate muscular and joint contact forces based on previously recorded activities, while the FE method can be employed to investigate local stresses and strains, as well as contact mechanics between tissues and implant components, including damage and failure. High quality models should attempt to apply physiological boundary conditions [11] and mimic realistic loading scenarios [6, 8].

The development of a FE knee model (both native and replaced) requires dealing with a high number of assumptions related to geometry, material properties, loading and boundary conditions, and the limitations of the model need to be suitably acknowledged and reported to be able to provide clinically relevant results. In addition, thorough model validation should be performed, reporting detailed information about its accuracy against experimental and/or in vivo measurements [15].

Kinematic and Joint Stability Tests in vivo

In general, a successful functional outcome for joint replacement is pain free and stable joint motion that provides sufficient range of motion to allow the patient to undertake all desired activities, accounting for different expectations between patients. Current testing in clinical environments generally relies on clinician experience in manual palpation of the joint in question, in order to provide an assessment of joint stability. Functional testing is often based on self-reported questionnaires, but early adoption of e.g. digital wearable technologies and gait evaluation, is now permeating into the field. However, a host of novel technologies to evaluate both joint stability and functional performance have recently emerged, possibly presenting high-quality options for objective testing. Moreover, since many of these approaches are based on digital technologies, new methods for reporting and storage of personalised data now become available, thereby also improving the comparability of data at different follow-up time points. The resulting solutions are hence able to facilitate digital platforms for efficient, low-cost and improved early diagnosis, screening, and monitoring of functional performance.

Based on the fundamental challenges introduced by the new Medical Device Regulation in the design, development, and testing of new implants to ensure high quality functional outcome, we therefore provide a number of consensus recommendations for the testing of joint function and stability, as well as feedback reporting, in order to promote implant innovation:

In Vitro Testing

- Standardized test rig designs that can apply muscle forces and boundary conditions in a consistent manner between experimental facilities (i.e. at different companies)
- Application of standardized boundary conditions [11], as well as kinematics / loading conditions based on high fidelity in vivo measurements such as the CAMS-Knee or Grand Challenge datasets [16, 17]
- Reporting of joint kinematics and laxity in consistent coordinate systems [12, 18, 19]

In Silico Models

- Simulation software should undergo rigorous testing and verification for simple problems where an analytical solution is known to ensure the coding of the equations of motion and continuum equations are correctly implemented [20]
- Use established benchmark datasets (e.g. CAMS-Knee [17], Grand Challenge [16]) such that consistent experimental / boundary conditions are applied as inputs to enable model predictions to be compared across simulation platforms and implant designs
- Rigorous validation should be performed against gold standard in vivo measurements (joint kinematics from fluoroscopy, joint contact forces from instrumented implants) [21, 22]
- Completion of thorough sensitivity analyses to quantify uncertainty in model predictions [15, 23]
- Reporting of models and simulation frameworks in consistent open-source formats containing complete model details that enable computation predictions to be reproduced [20, 24, 25]

In Vivo Measurements

- Measurements of high-fidelity datasets (e.g. biplane fluoroscopy) should be performed on a limited number of patients to precisely evaluate in vivo joint function for novel implant designs
- Assessment and reporting should be performed in not only good, but also poor outcomes to evaluate potential deficits in patients with stiff/painful/unstable joints
- Functional movement and mobility should be regularly collection as standard using "at home systems" (e.g. wearables / phone cameras) to monitor patient outcomes for extended periods following surgery
    - Gait symmetry, limb coordination, heel strike impact, range of joint motion ...
    - Gait speed

    - Mobility, number of steps, frequency of movement
    - Metrics of variability, if the resolution of the technology allows
- Development of digital data collection technologies and apps for mobile and smart devices for the regular collection of patient reports of pain, stability, and function at multiple time points, both pre- and postoperatively
- Support should be sought from insurance companies to fund the collection and sharing of this invaluable data source for objective evaluation of functional performance over long follow-up periods

Feedback reporting

Ideally functional metrics collected from digital technologies should be incorporated into registries in standardized formats that are well organized and easily parsed such that machine learning approaches can be applied to identify trends and evaluate implant performance.

- Together with surgical technique and implant type, sizes etc, metrics of joint stability and functional outcome data should be entered into open access registries
    - pre-operative status, including range of motion
    - use of standardised metrics and units for reporting stability, where available
    - Report stiffness in both internal and external rotation relative to a physiologically relaxed neutral position

When related specifically to product design, claims of novelty or uniqueness in terms of functionality should not be made on the results of computational predictions, but rather only after they have been verified within patients.

- What is the specific formulation of the intended use of a new product?
- If that intended use addresses kinetics and kinematics, then formulate these claims in sufficiently specific manner to be quantifiable as patient benefit
- Suggest a means on how to quantify this patient benefit. This could be larger knee mobility assessed by wearable technology or deep knee flexion under loading assessed in a fluoroscopic measurement
- Provide a clear frame of reference, ideally in existing clinical performance and with known data from post-market registry data

6. References

1. Fink, M. and B. Akra. **Regulatory clearance: How are outcome measurements critical? Injury, 2020. 51 Suppl 2: p. S67-s70.**

2. Baxmann, M., et al., **The influence of contact conditions and micromotions on the fretting behavior of modular titanium alloy taper connections. Medical Engineering & Physics, 2013. 35(5): p. 676-683.**

3. Hosseini Nasab, S.H., et al., **Loading Patterns of the Posterior Cruciate Ligament in the Healthy Knee: A Systematic Review. PLoS One, 2016. 11(11): p. e0167106.**

4. Hosseini Nasab, S.H., et al., **Length-Change Patterns of the Collateral Ligaments During Functional Activities After Total Knee Arthroplasty. Ann Biomed Eng, 2020. 48(4): p. 1396-1406.**

5. Huber, C., et al., **Properties and Function of the Medial Patellofemoral Ligament: A Systematic Review. Am J Sports Med, 2020. 48(3): p. 754-766.**

6. Bergmann, G., et al., **Standardized Loads Acting in Hip Implants. PLoS One, 2016. 11(5): p. e0155612.**

7. Heinlein, B., et al., **ESB Clinical Biomechanics Award 2008: Complete data of total knee replacement loading for level walking and stair climbing measured in vivo with a follow-up of 6-10 months. Clin Biomech (Bristol, Avon), 2009. 24(4): p. 315-26.**

8. Kutzner, I., et al., **Loading of the knee joint during activities of daily living measured in vivo in five subjects. J Biomech, 2010. 43(11): p. 2164-73.**

9. Victor, J., et al., **The influence of muscle load on tibiofemoral knee kinematics. Journal of Orthopaedic Research, 2010. 28(4): p. 419-428.**

10. Dupraz, I., et al., **Impact of femoro-tibial size combinations and TKA design on kinematics. Arch Orthop Trauma Surg, 2021.**

11. Speirs, A.D., et al., **Physiologically based boundary conditions in finite element modelling. J Biomech, 2007. 40(10): p. 2318-23.**

12. Wu, G., et al., **ISB recommendation on definitions of joint coordinate system of various joints for the reporting of human joint motion—part I: ankle, hip, and spine. Journal of Biomechanics, 2002. 35(4): p. 543-548.**

13. Ehrig, R.M., et al., **A survey of formal methods for determining the centre of rotation of ball joints. J Biomech, 2006. 39(15): p. 2798-809.**

14. Ehrig, R.M., et al., **A survey of formal methods for determining functional joint axes. J Biomech, 2007. 40(10): p. 2150-7.**

15. Anderson, A.E., B.J. Ellis, and J.A. Weiss, **Verification, validation and sensitivity studies in computational biomechanics. Computer Methods in Biomechanics and Biomedical Engineering, 2007. 10(3): p. 171-184.**

16. Fregly, B.J., et al., **Grand challenge competition to predict in vivo knee loads. J Orthop Res, 2012. 30(4): p. 503-13.**

17. Taylor, W.R., et al., **A comprehensive assessment of the musculoskeletal system: The CAMS-Knee data set. J Biomech, 2017. 65: p. 32-39.**

18. Wu, G., et al., **ISB recommendation on definitions of joint coordinate systems of various joints for the reporting of human joint motion—Part II: shoulder, elbow, wrist and hand. Journal of Biomechanics, 2005. 38(5): p. 981-992.**

19. Hull, M.L., **Coordinate system requirements to determine motions of the tibiofemoral joint free from kinematic crosstalk errors. J Biomech, 2020. 109: p. 109928.**

20. Erdemir, A., et al., **Considerations for reporting finite element analysis studies in biomechanics. Journal of Biomechanics, 2012. 45(4): p. 625-633.**

21. Hicks, J.L., et al., **Is My Model Good Enough? Best Practices for Verification and Validation of Musculoskeletal Models and Simulations of Movement. Journal of Biomechanical Engineering, 2015. 137(2).**

22. Petrella, A., et al., **How Good is Good Enough? Lessons in Musculoskeletal Model Validation With the Anybody Modeling System. Journal of Medical Devices, 2013. 7(4).**

23. Innocenti, B., et al., **Contact forces in several TKA designs during squatting: A numerical sensitivity analysis. Journal of Biomechanics, 2011. 44(8): p. 1573-1581.**

24. Lund, M.E., et al., **On validation of multibody musculoskeletal models. Proceedings of the Institution of Mechanical Engineers, Part H: Journal of Engineering in Medicine, 2012. 226(2): p. 82-94.**

25. Derrick, T.R., et al., **ISB recommendations on the reporting of intersegmental forces and moments during human motion analysis. J Biomech, 2020. 99: p. 109533.**

RESEARCH TOPIC

TRANSFERABILITY OF RESULTS

RESEARCH TOPIC 20

Transferability of Pre-Clinical/Clinical Results of a Product to Another Device Taking into Consideration General Criteria of Equivalence Devices/ Reference Products and Limitations for the Transferability of Results Across Different Variants of One Implant

Authors

Dario Bergadano[1], Joshua Bridgens[2], Thomas M. Grupp[3,4], Arndt-Peter Schulz[5,6]

[1]Medacta International SA, Castel San Pietro, Switzerland
[2]DePuy Synthes, Leeds, United Kingdom
[3]Aesculap AG Research & Development, Tuttlingen, Germany
[4]Ludwig Maximilians University Munich, Department of Orthopaedic and Trauma Surgery, Musculoskeletal University Center Munich (MUM), Campus Grosshadern, Munich, Germany
[5]Department of Trauma Surgery, Orthopedic and Sports Traumatology, BG Klinikum Hamburg, Hamburg, Germany
[6]Fraunhofer Research Institution for Individualized and Cell-Based Medical Engineering, Lübeck, Germany

1. Abstract

Question: Transferability of results

To what extent can pre-clinical/clinical results of a specific product be transferred to another device?

1. Which general (not product-specific) criteria must equivalence devices/reference products fulfil to ensure transferability of results to other products (in relation to the relevant parameters in a specific case)?
2. Which limitations exist for the transferability of pre-clinical/clinical results across different variants of one implant (e.g. regarding different sizes, different materials, combination possibilities)?

Summary / Recommendation:

Pre-clinical and clinical results of a specific product can be transferred to another device if the two devices are EQUIVALENT. To prove EQUIVALENCE, the criteria specified in MDR, Annex XIV Part A must be fulfilled. As a guidance, some criteria specific to the field of large joint replacement devices are listed here below. ALL criteria must be fulfilled.

The EQUIVALENCE criteria proposed here are compatible with those required by TGA (Clinical Evidence Guidelines, Medical devices, 2020) and by FDA (Final Guidance - The New 510 (k) Paradigm - Alternate approaches to demonstrating substantial equivalence in premarket notifications, 1998) to demonstrate substantial equivalence.

The new device can benefit from updates derived from an adjourned state of the art, provided the new state-of-the-art features have been clinically validated in a SIMILAR application.

Variants to the original device, such as different dimensions or dimensional proportions can always be included in the transfer of clinical results as long as said variants constitute a harmonized distribution within the product range and are demonstrably required to better fit the intended patient population. In case the variants represent a mechanical worst-case, all foreseen preclinical tests must be carried out before conformity can be claimed.

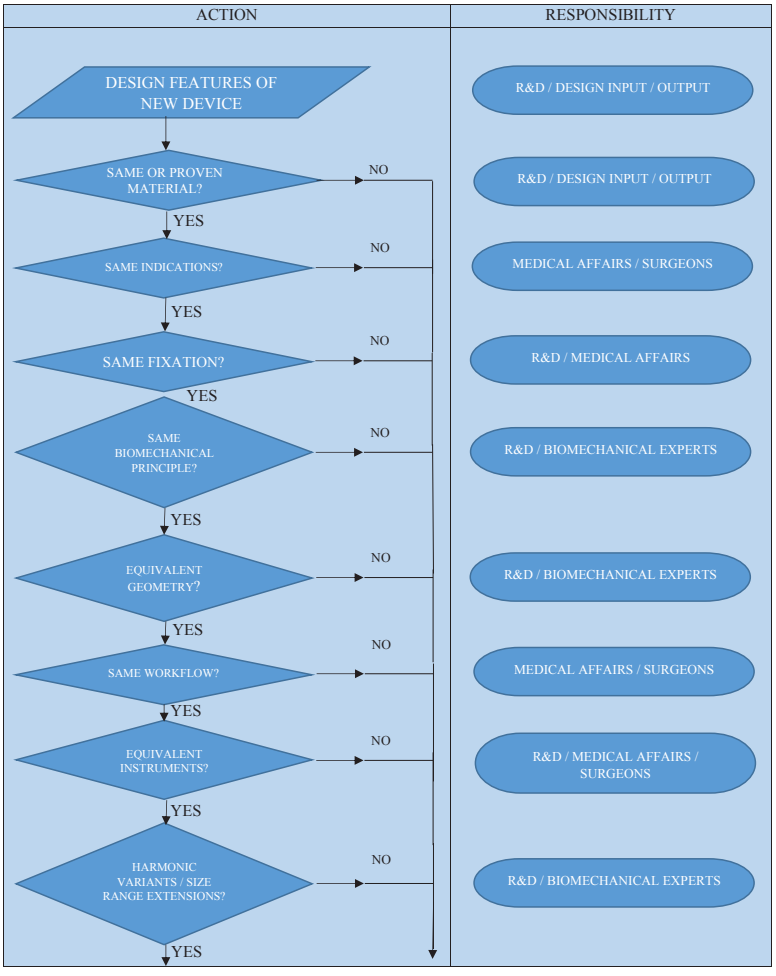
2. Level of Evidence

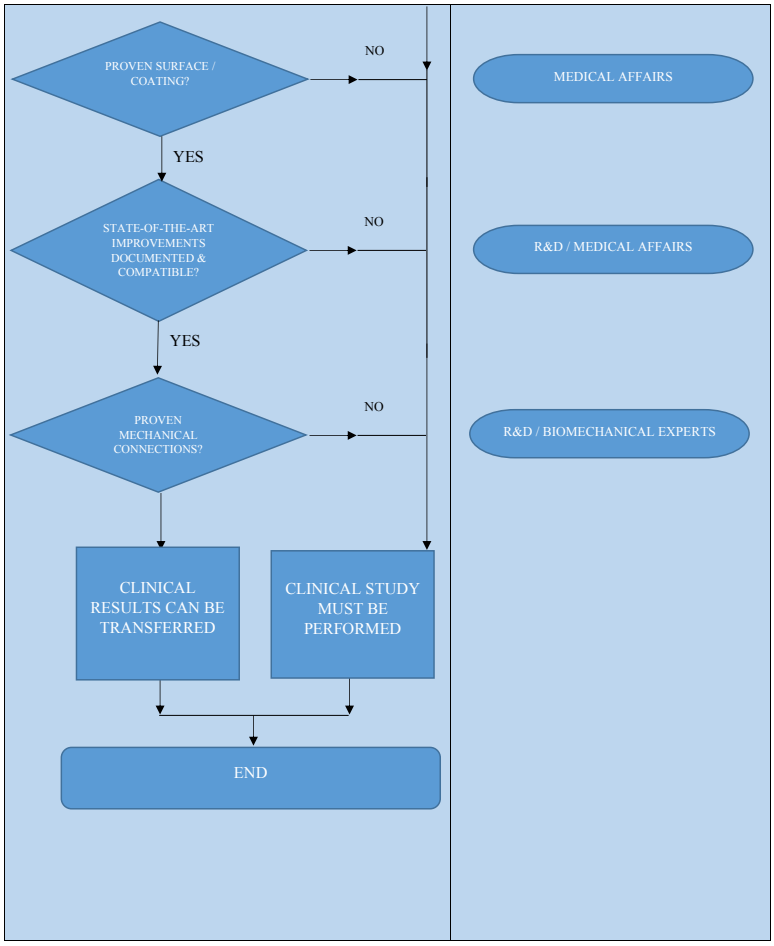
Moderate

3. Consensus Delegate Vote

92% – super majority, strong consensus (92% agree / 8% disagree / 0% abstain)

4. Graphical Abstract





5. Search Strategy

Research focus: equivalence in knee and hip prosthesis

- Inclusion criteria:
  - orthopedic prosthesis hip/knee joint
  - comparative study
  - regulatory topic: equivalence demonstration
- Exclusion criteria
  - Languages different from English
  - Congress/event Proceeding
  - Duplicate paper
  - Non-inherent

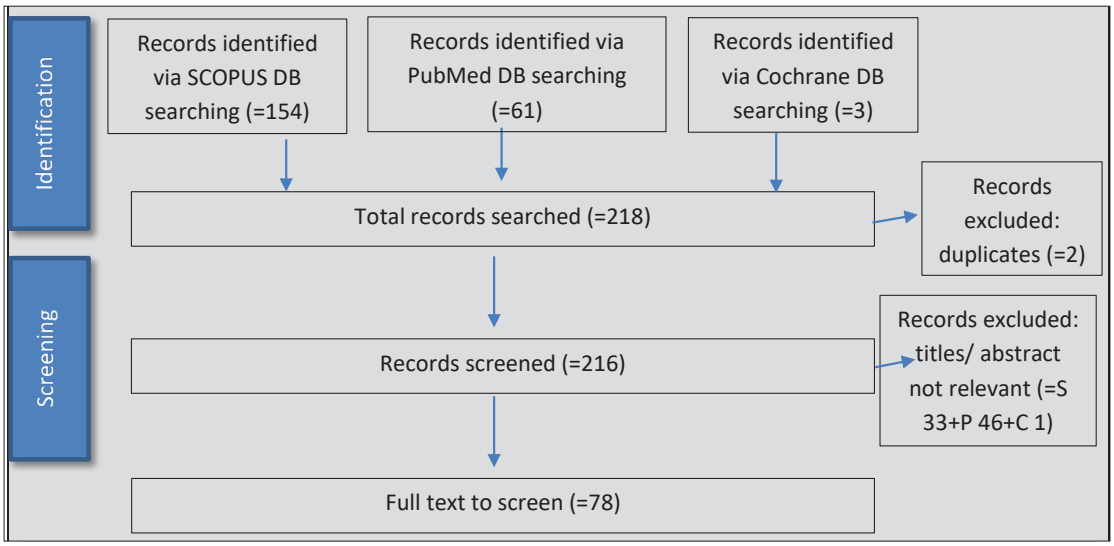
Search platforms included:

- SCOPUS
- PubMed
- Cochrane

Search platform: SCOPUS (S)
Search terms:
SUBJAREA (medi) AND (predicate OR predecessor) AND KEY (prosthesis) AND (LIMIT-TO (LANGUAGE, "English"))
Timeframe: ALL

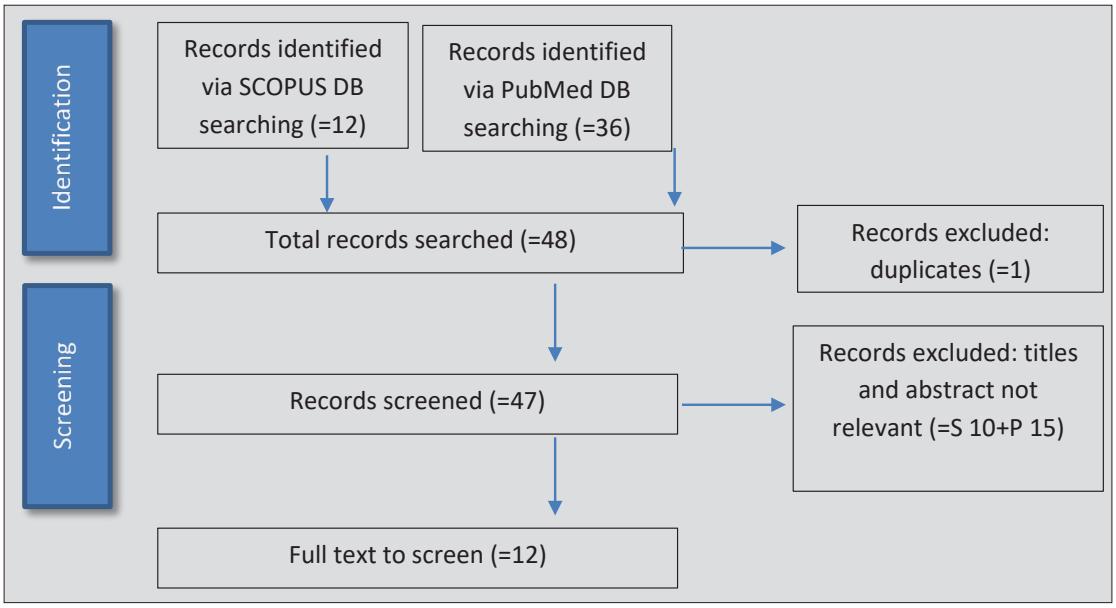
Search platform: PubMed (P)
Search terms:
(((knee arthroplasty[MeSH Major Topic]) OR (hip arthroplasty[MeSH Major Topic]))) AND (((predicate OR predecessor OR similar OR variant*) AND (device AND equiv*)))
Timeframe: ALL

Search platform: Cochrane (C)
Search terms:
#1 MeSH descriptor: [Arthroplasty, Replacement, Hip] explode all trees 1900
#2 MeSH descriptor: [Arthroplasty, Replacement, Knee] explode all trees 2552
#3 ("predecessor"):ti,ab,kw 99
#4 (#1 OR #2) AND #3



Search platform used: SCOPUS (S)
Search terms:
ALL ("size range" OR "size extension" OR "surface finish") AND (equivalent OR similar OR variant) AND (hip OR knee) AND (prosthesis) AND (LIMIT-TO (SUBJAREA, "ENGI") OR LIMIT-TO (SUBJAREA, "MATE")) AND (LIMIT-TO (LANGUAGE, "English")) AND (LIMIT-TO (SUBJAREA, "MEDI"))
Timeframe: ALL

Search platform used: PubMed (P)
Search terms:
("size range" OR "size extension" OR "surface finish") AND (equivalent OR similar OR variant) AND (hip OR knee) AND (prosthesis)
Timeframe: ALL



6. Rationale

The substantial equivalence between two joint replacement devices can be demonstrated if the following requirements are ALL fulfilled:

1. The devices are made out of the same material group. Alternatively, the material of the new device has been used successfully in the same specific application. E.g.: a Cobalt-alloy version of a tibial tray previously made in Titanium alloy. Commonly used material groups are:
   - Ti alloys,
   - Co alloys,
   - Stainless steels,
   - Ultra High Molecular Weight Polyethylene
   - Alumina,
   - Zirconia,
   - TZA ceramics,
   - ATZ ceramics,
2. The devices are intended for the same indications
3. The devices rely on the same fixation principle (cemented/ cementless)
4. The devices are based on the same "biomechanical principle". Some examples are supplied here below:
   - Cementless femoral stem in THA: metaphyseal / diaphyseal fixation,
   - Cemented femoral stem in THA: thin / uniform cement mantle, polished / porous / rough / macrostructured surface,
   - Cementless pressfit acetabular cup: hemispherical / elliptical / conical / Threaded / DualMobility,
   - TKA. Fixed / Mobile bearing
   - TKA: CR / PS / UC / MP
5. The devices are designed based on the same geometry. Relevant geometrical differences should be assessed by technical experts, thereby meaning design engineers with adequate documented experience. Some examples (non-exhaustive list):
   - Femoral stem in THA: rounded distal / rectangular cross section,
   - Coverage angle in THA acetabular cups,
   - Single /variable radius in TKA femoral components,
   - Keel and macrostructures in TKA tibial components
6. The devices are implanted following a similar surgical workflow. Some examples:
   - Reamed hip stem / broach-only
   - TKA: extramedullary / Intramedullary alignment, bone referencing / ligament balancing / kinematic alignment
7. The dedicated instruments are conceptually the same though their technical design and materials may vary
8. The new device may include some new variants as long as this is not in contradiction with §4 (e.g., THA stem: shorter or anteverted neck, thinner distal section; THA cups: screw holes, thickness, TKA femoral component: different anterior shield length / Q angle / posterior condyles, TKA tibial component: different stem / keel heights)
9. The new device may cover an extended size range as well as a different size distribution, as long as the added sizes follow the same criteria used for sizing in the original device
10. The new device may have a different surface or coating, as long as the new features have been clinically proven in the same application and the difference does not alter the biomechanical principles (e.g. rough/polished cemented stems are NOT equivalent)
NOTE: It has been shown in the past that femoral hip stems of the same design can react sensitive to substantial surface modifications of the surface which change the functional mechanism of stem fixation. Thus, a thorough justification of any modifications is mandatory.
11. The new device may benefit from an applicable evolution in the technical state of the art (e.g.: a new/improved manufacturing technology, provided a risk-based approach is maintained and documented; a new treatment to enhance material properties, if duly documented in pre-clinical tests)
12. Mechanical connections should be identical to the original device's or demonstrated to be non-inferior in comparative mechanical tests

7. References

Abdullah AH, Jaafar EF, Nasir N, Abdul Latip EN, Tardan G. **Influences of prosthesis stem lengths in cementless total hip arthroplasty [Internet]. 2011. 2088 p. (Applied Mechanics and Materials; vols. 52–54). Available from: https://www.scopus.com/inward/record.uri?eid=2-s2.0-79954527026&doi=10.4028%2fwww.scientific.net%2fAMM.52-54.2088&partnerID=40&md5=7c9ef088332b29b813692c6916ba94d1**

Alsheghri AA, Alageel O, Amine Mezour M, Sun B, Yue S, Tamimi F, et al. **Bio-inspired and optimized interlocking features for strengthening metal/polymer interfaces in additively manufactured prostheses. Acta Biomaterialia. 2018;80:425–34.**

Bai F, Zhang J, Wang Z, Lu J, Chang J, Liu J, et al. **The effect of pore size on tissue ingrowth and neovascularization in porous bioceramics of controlled architecture in vivo. Biomedical Materials [Internet]. 2011;6(1). Available from: https://www.scopus.com/inward/record.uri?eid=2-s2.0-79551480846&doi=10.1088%2f1748-6041%2f6%2f1%2f015007&partnerID=40&md5=cc95b67f0cd31dad97a4d440a50aacd4**

Bermúdez-Castañeda A, Igual-Muñoz A, Mischler S.; **A Crevice Corrosion Model for Biomedical Trunnion Geometries and Surfaces Feature. ; Materials (Basel). 2021 Feb 20; 14(4):1005. doi: 10.3390/ma14041005.; PMID: 33672713 Free PMC article.**

Bishop NE, Höhn J-C, Rothstock S, Damm NB, Morlock MM. **The influence of bone damage on press-fit mechanics. Journal of biomechanics. 2014;47(6):1472–8.**

Bryant M, Farrar R, Brummitt K, Freeman R, Neville A. **Fretting corrosion of fully cemented polished collarless tapered stems: The influence of PMMA bone cement. Wear. 2013;301(1–2):290–9.**

Caruana J, Janssen D, Verdonschot N, Blunn GW. **The importance of a thick cement mantle depends on stem geometry and stem-cement interfacial bonding. Proceedings of the Institution of Mechanical Engineers, Part H: Journal of Engineering in Medicine. 2009;223(3):315–27.**

Collis DK, Mohler CG.; **Loosening rates and bone lysis with rough finished and polished stems. ; Clin Orthop Relat Res. 1998 Oct; (355):113-22. doi: 10.1097/00003086-199810000-00012.; PMID: 9917596**

Cristofolini L, Teutonico AS, Savigni P, Erani P, Viceconti M. **Preclinical assessment of the long-term endurance of cemented hip stems. Part 1: Effect of daily activities - A comparison of two load histories. Proceedings of the Institution of Mechanical Engineers, Part H: Journal of Engineering in Medicine. 2007;221(6):569–84.**

Damron LA, Kim D-G, Mann KA. **Fatigue debonding of the roughened stem-cement interface: Effects of surface roughness and stem heating conditions. Journal of Biomedical Materials Research - Part B Applied Biomaterials. 2006;78(1):181–8.**

Datir SP, Kurta IC, Wynn-Jones CH.; **Ten-year survivorship of rough-surfaced femoral stem with geometry similar to Charnley femoral stem. ; J Arthroplasty. 2006 Apr; 21(3):392-7. doi: 10.1016/j.arth.2004.11.018.; PMID: 16627148**

Datir SP, Wynn-Jones CH.; **Staged bilateral total hip arthroplasty using rough and smooth surface femoral stems with similar design: 10-year survivorship of 48 cases. ; Acta Orthop. 2005 Dec; 76(6):809-14. doi: 10.1080/17453670510045417.; PMID: 16470434**

Della Valle AG, Zoppi A, Peterson MG, Salvati EA.; **A rough surface finish adversely affects the survivorship of a cemented femoral stem. ; Clin Orthop Relat Res. 2005 Jul; (436):158-63. doi: 10.1097/00003086-200507000-00024.; PMID: 15995435**

Endo MM, Barbour PS, Barton DC, Wroblewski BM, Fisher J, Tipper JL, Ingham E, Stone MH.; **A comparison of the wear and debris generation of GUR 1120 (compression moulded) and GUR 4150HP (ram extruded) ultra high molecular weight polyethylene. ; Biomed Mater Eng. 1999; 9(2):113-24.; PMID: 10524294**

European Commission, Clinical evaluation – equivalence MDCG 2020-5, April 2020, available at https://ec.europa.eu/health/sites/health/files/md_sector/docs/md_mdcg_2020_5_guidance_clinical_evaluation_equivalence_en.pdf

Food & Drug Law Institute, Shapiro J.K., **Substantial Equivalence Premarket Review, Food and Drug Law Journal, Vol. 9, No. 3 (2014), pp.365-398, available at https://www.jstor.org/stable/10.2307/26661024**

Galvin A, Jennings LM, McEwen HM, Fisher J. **The influence of tibial tray design on the wear of fixed-bearing total knee replacements. Proceedings of the Institution of Mechanical Engineers, Part H: Journal of Engineering in Medicine. 2008;222(8):1289–93.**

Galvin A, Jennings LM, McEwen HM, Fisher J.; **The influence of tibial tray design on the wear of fixed-bearing total knee replacements. ; Proc Inst Mech Eng H. 2008 Nov; 222(8):1289-93. doi: 10.1243/09544119JEIM434.; PMID: 19143421**

Germain MA, Tipper JL, Ingham E, Makarem R, Wroblewski BM, Fisher J.; **Analysis of debris from brushing the femoral canal with a plastic brush--a potential cause of loosening in total hip replacement. ; Proc Inst Mech Eng H. 1999; 213(6):503-6. doi: 10.1243/0954411991535112.; PMID: 10635699**

Gheduzzi S, Miles AW. **A review of pre-clinical testing of femoral stem subsidence and comparison with clinical data. Proceedings of the Institution of Mechanical Engineers, Part H: Journal of Engineering in Medicine. 2007;221(1):39–46**.

González Della Valle A, Rana A, Nestor B, Bostrom M, Westrich G, Salvati EA.; **Metallic shedding, surface finish changes, and extensive femoral osteolysis in the loose Spectron EF stem. ; Clin Orthop Relat Res. 2006 Jan; 442:165-70. doi: 10.1097/01.blo.0000181145.01306.f9.; PMID: 16394756**

Goodman SB, Mihalko WM, Anderson PA, Sale K, Bozic KJ. **Introduction of New Technologies in Orthopaedic Surgery. JBJS Rev. 2016 May 24;4(5).**

Gravius S, Wirtz DC, Siebert CH, Andereya St, Mueller-Rath R, Maus U, et al. I**n vitro interface and cement mantle analysis of different femur stem designs. Journal of Biomechanics. 2008;41(9):2021–8.**

Gravius S, Wirtz DC, Siebert CH, Mueller-Rath R, Maus U, Mumme T. **In vitro interface and cement mantle analysis of different femur stem designs. Journal of biomechanics. 2008;41(9):2021–8.**

Hamadouche M, Baqué F, Lefevre N, Kerboull M.; **Minimum 10-year survival of Kerboull cemented stems according to surface finish. ; Clin Orthop Relat Res. 2008 Feb; 466(2):332-9. doi: 10.1007/s11999-007-0074-6. Epub 2008 Jan 10.; PMID: 18196414 Free PMC article. Clinical Trial.**

Haynes DR, Rogers SD, Hay S, Pearcy MJ, Howie DW.; **The differences in toxicity and release of bone-resorbing mediators induced by titanium and cobalt-chromium-alloy wear particles. ; J Bone Joint Surg Am. 1993 Jun; 75(6):825-34. doi: 10.2106/00004623-199306000-00004.; PMID: 8314823**

Hinrichs F, Kuhl M, Boudriot U, Griss P.; **A comparative clinical outcome evaluation of smooth (10-13 year results) versus rough surface finish (5-8 year results) in an otherwise identically designed cemented titanium alloy stem. ; Arch Orthop Trauma Surg. 2003 Jul; 123(6):268-72. doi: 10.1007/s00402-003-0515-y. Epub 2003 May 13.; PMID: 12743716**

Kedgley AE, Takaki SE, Lang P, Dunning CE. **The effect of cross-sectional stem shape on the torsional stability of cemented implant components. 2007;**

Kedgley AE, Takaki SE, Lang P, Dunning CE. **The effect of cross-sectional stem shape on the torsional stability of cemented implant components. Journal of Biomechanical Engineering. 2007;129(3):310–4.**

Lachiewicz PF, Kelley SS, Soileau ES.; **Survival of polished compared with precoated roughened cemented femoral components. A prospective, randomized study. ; J Bone Joint Surg Am. 2008 Jul; 90(7):1457-63. doi: 10.2106/JBJS.G.01043.; PMID: 18594093 Clinical Trial.**

Lai Y-S, Wei H-W, Chang T-K, Cheng C-K. **The effects of femoral canal geometries, stem shapes, cement thickness, and stem materials on the choice of femoral implant in cemented total hip replacement. Journal of the Chinese Institute of Engineers, Transactions of the Chinese Institute of Engineers,Series A/Chung-kuo Kung Ch'eng Hsuch K'an. 2009;32(3):333–41.**

Łapaj Ł, Mróz A, Kokoszka P, Markuszewski J, Wendland J, Helak-Łapaj C, Kruczyński J.; **Peripheral snap-fit locking mechanisms and smooth surface finish of tibial trays reduce backside wear in fixed-bearing total knee arthroplasty. ; Acta Orthop. 2017 Feb; 88(1):62-69. doi: 10.1080/17453674.2016.1248202. Epub 2016 Oct 26.; PMID: 27781667 Free PMC article.**

Maak TG, Wylie JD. Medical Device Regulation: **A Comparison of the United States and the European Union. J Am Acad Orthop Surg. 2016 Aug;24(8):537–43.**

Nam D, Lawrie CM, Salih R, Nahhas CR, Barrack RL, Nunley RM. **Cemented Versus Cementless Total Knee Arthroplasty of the Same Modern Design: A Prospective, Randomized Trial. J Bone Joint Surg Am. 2019 Jul 3;101(13):1185–92.**

Norman TL, Shultz T, Noble G, Gruen TA, Blaha JD. **Bone creep and short and long term subsidence after cemented stem total hip arthroplasty (THA). Journal of Biomechanics. 2013;46(5):949–55.**

Ong A, Wong KL, Lai M, Garino JP, Steinberg ME.; **Early failure of precoated femoral components in primary total hip arthroplasty. ; J Bone Joint Surg Am. 2002 May; 84(5):786-92. doi: 10.2106/00004623-200205000-00014.; PMID: 12004022**

Pajarinen J, Lin T-H, Sato T, Yao Z, Goodman SB. **Interaction of materials and biology in total joint replacement-successes, challenges and future directions. Journal of Materials Chemistry B. 2014;2(41):7094–108.**

Puloski SK, McCalden RW, Macdonald SJ, Bourne RB, Rorabeck CH.; **Surface analysis of posterior stabilized femoral components used in total knee arthroplasty. ; J Arthroplasty. 2003 Oct; 18(7):822-6. doi: 10.1016/s0883-5403(03)00327-9.; PMID: 14566734**

Rahman HSA, Choudhury D, Osman NAA, Shasmin HN, Bakar WA, Abas W. **In vivo and in vitro outcomes of alumina, zirconia and their composited ceramic-on-ceramic hip joints. Journal of the Ceramic Society of Japan. 2013;121(1412):382–7.**

Ramos A, Simões JA. **The influence of cement mantle thickness and stem geometry on fatigue damage in two different cemented hip femoral prostheses. Journal of biomechanics. 2009;42(15):2602–10.**

Rudol G, Wilcox R, Jin Z, Tsiridis E.; **The effect of surface finish and interstitial fluid on the cement-in-cement interface in revision surgery of the hip. ; J Bone Joint Surg Br. 2011 Feb; 93(2):188-93. doi: 10.1302/0301-620X.93B2.24759.; PMID: 21282757**

Tatani I, Megas P, Panagopoulos A, Diamantakos I, Nanopoulos P, Pantelakis S. **Comparative analysis of the biomechanical behavior of two different design metaphyseal-fitting short stems using digital image correlation. Biomedical engineering online. 2020;19(1):1–18.**

Therapeutic Goods Administration, Clinical evidence guidelines – medical devices, Version 1.1 November 2020, available at https://www.tga.gov.au/node/913447

Utzschneider S, Paulus A, Datz JC, Schroeder C, Sievers B, Wegener B, Jansson V.; **Influence of design and bearing material on polyethylene wear particle generation in total knee replacement. ; Acta Biomater. 2009 Sep; 5(7):2495-502. doi: 10.1016/j.actbio.2009.03.016. Epub 2009 Mar 24.; PMID: 19375997**

Verdonschot N, Tanck E, Huiskes R. **Effects of prosthesis surface roughness on the failure process of cemented hip implants after stem-cement debonding. Journal of Biomedical Materials Research. 1998;42(4):554–9.**

Verdonschot N, Tanck E, Huiskes R.; **Effects of prosthesis surface roughness on the failure process of cemented hip implants after stem-cement debonding. ; J Biomed Mater Res. 1998 Dec 15; 42(4):554-9. doi: 10.1002/(sici)1097-4636(19981215)42:4<554::aid-jbm11>3.0.co,2-k.; PMID: 9827679**

Zuckerman D, Brown P, Das A. **Lack of publicly available scientific evidence on the safety and effectiveness of implanted medical devices. JAMA Intern Med. 2014 Nov;174(11):1781–7.**

RESEARCH TOPIC 21

Transferring Pre-Clinical Results of Joint Replacement Devices into the Clinical Setting

Authors

Marcus Jäger[1,2], Francesco Traina[3,4], Alexander Giurea[5], Thomas M. Grupp[6,7], Sabine Rusch[8]

[1]Department of Orthopaedics, Trauma & Reconstructive Surgery, St. Marien Hospital Mülheim an der Ruhr, Mülheim, Germany

[2]Chair of Orthopaedics and Trauma Surgery, University of Duisburg-Essen, Essen, Germany

[3]Ortopedia-Traumatologia e Chirurgia protesica e dei reimpianti d'anca e di ginocchio, IRCCS Istituto Ortopedico Rizzoli, Bologna, Italy

[4]DIBINEM, Università di Bologna, Bologna, Italy

[5]Department of Orthopedics and Trauma Surgery, Medical University of Vienna, Vienna, Austria

[6]Aesculap AG Research & Development, Tuttlingen, Germany

[7]Ludwig Maximilians University Munich, Department of Orthopaedic and Trauma Surgery, Musculoskeletal University Center Munich (MUM), Campus Grosshadern, Munich, Germany

[8]Aesculap AG Medical Scientific Affairs, Tuttlingen, Germany

1. Abstract

Question: Transferability of results

To what extent can pre-clinical test results of a product be transferred into the clinical setting?

Summary/Recommendation:

Pre-clinical testing is an important phase of product development and medical device registration. To transfer per-clinical tests in a clinical scenario the prediction rate of a defined preclinical evaluation method has to be evaluated. Here, scientific literature and data from registries might help to predict the reliability of a defined method.

On this basis we recommend to select an appropriate basic test method, explore the differences between test and reality and define potential effects. Afterwards an assessment of the testing method is recommended (unrestricted transfer possible, limited applicability, non-applicability). Further proceeding includes the adjustments of methodology, the planning of complementary tests and the specification of open topics including potential risk for patients and users.

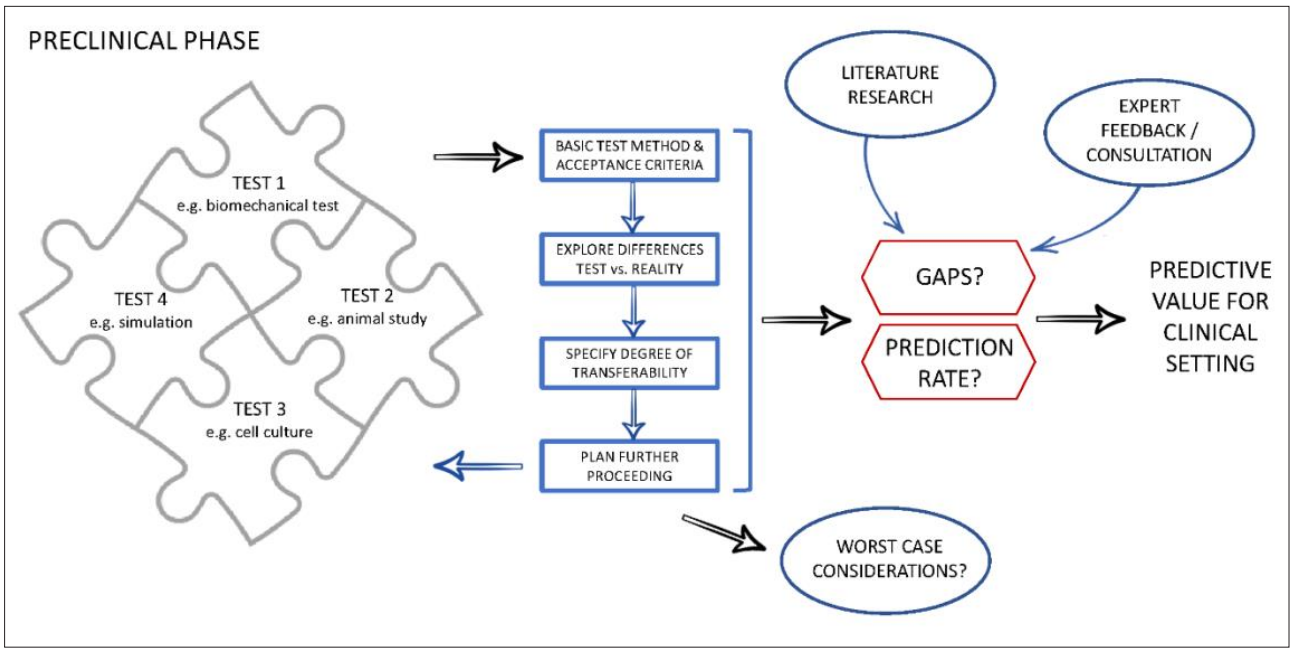
2. Level of Evidence

Limited

3. Consensus Delegate Vote

90% – super majority, strong consensus (90% agree / 2% disagree / 8% abstain)

4. Graphical Abstract



5. Search Strategy

A systematic literature search was conducted in May 2021 to identify literature with the potential to contribute to the answer of the above-mentioned research question.

The electronic databases PubMed and EMBASE were explored using a combination of the following search-terms combined with the Boolean operator:

- "Arthroplasty" AND "preclinical test" AND "transfer" AND "outcome": 395 hits
- "TKA / /THA", AND "preclinical test" AND "transfer" AND "outcome": 12 hits
- Total knee replacement "AND "preclinical test" AND "transfer" AND "outcome": 592 hits
- "Total hip replacement" AND "preclinical test" AND "transfer" AND "outcome": 590 hits

Search words were adapted to the specific terminology of the database and were searched including different spelling and synonyms.

Article language was restricted to English, French, German or Italian language.

Because of the aim of the process only studies from 1980 to 2021 were included to provide insight into current practice and state of the art procedures.

Publications citing which did not focus on the transfer of pre-clinical data into the clinical scenario were excluded as those papers mainly reported non-prosthetic materials such as bone substitutes, antibiotic testing etc.

The selection was not further restricted regarding basic study design.

Duplicate articles were removed.

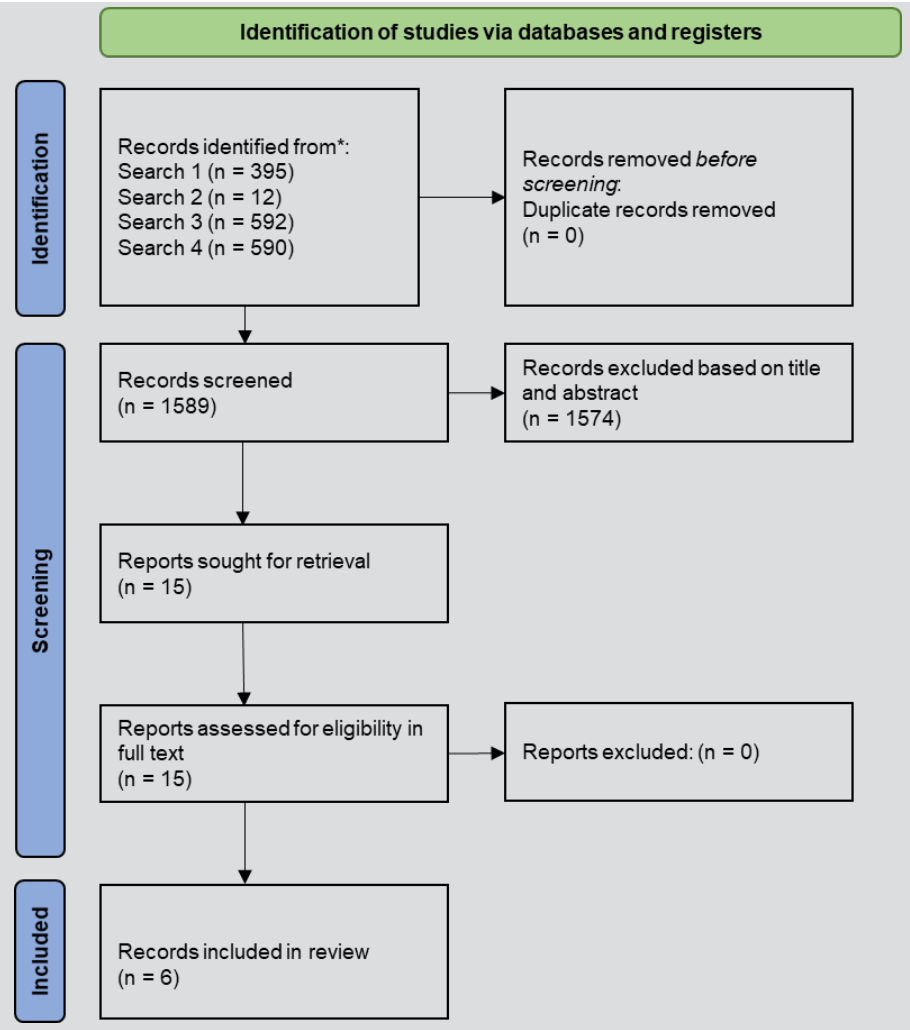
Abstracts from scientific meetings were not included in the review as sufficient data is not provided.

Case reports and case series were also excluded due to the nature of the contained information.

Basic research articles with no correlation to clinical data were also excluded.

The title and abstract of all articles were screened to select articles contributing to the topic. Potentially eligible articles were obtained for detailed assessment and were screened for applicability in detail.

After the selection process, 6 articles in total were included in the review.



6. Rationale

Which general (not product-specific) criteria must products and the pre-clinical testing methods fulfil to guarantee transferability of results into the clinical setting?

Pre-clinical testing is an important phase of product development and medical device registration. The transferability of pre-clinical testing methods into the clinical data has the potential to reduce the efforts of distributing manufacturers to approve a medicinal product (time, costs). Furthermore, there are also associated ethical aspects such as a decrease of animal experiments, minimise the risk (i.e. open questions) of a new product for clinical study patients, reduce time consuming clinical studies while patients become deprived of fundamental improvements of technical progress as well as a waste in human resources and stress in the clinical setting (1).

However, pre-clinical testing can only in rare cases consider every aspect related to clinical safety and performance of an endoprosthetic implant (1). In accordance to the MDR a critical risk-benefit evaluation has to be applied to pre-clinical testing to avert unnecessary damage to patients, users, third parties and assets. To transfer per-clinical tests in a clinical scenario the prediction rate of a defined

preclinical evaluation method is a key point. We assume that in some evaluation techniques, the prediction rate is even not known or can only be estimated approximately. In these cases, appropriate knowledge of empiric data from the literature can be a solid base for decision making (2). Also, a consultation of clinical experts and feedback by senior orthopaedic surgeons might be helpful to select the appropriate pre-clinical evaluation method and for the reliability of the assumed predictive value. Here, clinical experience and clinical outcome data will help to improve laboratory accuracy and prediction. It can provide valuable guidelines for pre-clinical test protocols.

To basically allow an assessment how or to what extent pre-clinical test results can predict the expression of specific aspects under clinical circumstances, the test methodology must fulfil certain requirements. Frequently, complementary testing using different stages of testing (e.g. basic cell culture testing, biomechanical or bench testing, animal experiments) are required to address all relevant aspects (1, 2, 3). The following process is recommended in order to optimally define the adequate proceeding and strengthen the prediction rate (1, 3):

1. Select an appropriate basic test method and acceptance criteria for pass/fail:
   - standardized test methods: those should be used as they allow the interpretation of results based on broad experience with established products and acceptance criteria are defined (2)
   - established test methods confirmed by clinical success: preferably validity of test results is confirmed by clinical success of similar products, acceptance criteria can thus be defined
   - comparative testing with established products or materials (4): comparative testing following distinct test methods can be valuable for specific questions, e.g. concerning material differences
   - newly developed test methods: those need extensive validation also regarding acceptance criteria (3, 5), e.g. based on testing of established products but can be very valuable
2. Explore differences between test and reality and define potential effects:
   - controllable differences of known circumstances: those should be minimized, e.g. temperature, used lubricants and physiological fluids (1, 6)
   - not controllable differences of known circumstances: those must be considered and represented as best as possible (3, 5, 6), e.g. peak forces
   - potentially unexpected differences incl. possible effects: it must be kept in mind that living systems may include unpredictable reactions or processes which may have an impact on the actual performance of products in humans (1, 4–6)
3. Specify assessable aspects based on 1. and 2.:
   - aspects with full applicability of results: unrestricted transfer of obtained results into the clinical application is possible
   - aspects with limited applicability: gaps in pertinence restrict the transfer possibilities, certain questions may remain open and should be considered further (3)
   - aspects which cannot be assessed preclinically: those need to be addressed further or may remain for clinical assessment
4. Plan further proceeding:
   - adjust test methodology: select tests which may be better suited or adapt the test situation, e.g. load direction, different animal model
   - plan complementary tests to address gaps: where aspects cannot be addressed adequately the steps 1. to 4. should be repeated for those gaps (4)
   - specify open topics and justify risks for clinical studies/assessment (2)

Advantages of this proceeding are:

- clear documentation and robust reasoning
- structured identification of gaps which can possibly be closed
- complexity is broken down to make gaps manageable and allow them to be addressed by literature research, expert opinion, or supplementary testing
- defined gaps can be closed
- focused clinical studies can be planned to lower the associated risk for patients and limit required resources
- results of subsequent clinical studies will be more valuable due to focused research questions

Thus, it is possible to choose a test methodology or test array with a high prediction rate that can reliably address all effects on clinical safety and performance which can be assessed preclinically (1).

Which problems exist regarding transfer of worst-case testing results, i.e. of extreme implant combinations or constellations, margin implant sizes, rare articulation materials (e.g. metal-ceramic), to the clinical use of the complete implant system in standard cases?

The lack of knowledge in the prediction rates of a defined pre-clinical evaluation methods is not the only aspect which limits the transferability of pre-clinical tests to the clinical scenario. There are also other associated challenges which might complicate the transferability into clinical settings. Especially the outliers (best and worst-case testing results) are lacking in forecasting power. For these cases we also recommend a stepwise procedure, comparable to the above-mentioned workflow, which includes not only one but also other alternative pre-clinical testing methods such as animal experiments.

Typical worst-case testing results can occur by the simulation of situations in the following domains:
Implant related with regard to intended or reasonably expectable use:

- extreme implant sizes, e.g. smallest implant sizes with regard to body weight
- extreme geometric implant combinations, e.g. tibial and femoral size differences in TKA
- the combination of different implant systems from different manufacturers, e.g. not unlikely in arthroplasty revision surgery
- patient-specific / custom-made implants

Related to test methodology to respect patient related effects:

- movement patterns of the articulating partners which do not occur in preclinical settings (e. g. high speed tribology) (3)
- extreme load transfers to material (e. g. to simulate the effects of adipose patients to an implant)
- implant failure and associated effects (6)
- significant modifications of lubricants in biotribology (1) (e. g. changes of the pH or viscosity for the simulation of pre-clinical infections)

It is important to consider those situations and describe or document them accordingly in the test documentation with regard to the following questions (3, 5, 6):

1. Which aspects can be addressed by worst-case testing and which clinical parameters need to be considered?
2. Which circumstances need to be controlled in the test scenario?
3. Is the chosen worst-case truly valid for the assessed parameter?
4. Which aspects may be assessed by additional simulation techniques?
5. Which aspects cannot be assessed preclinically?

Furthermore, has to be evaluated how to proceed with the obtained test result:

- worst-case testing positive (passed): for which aspect(s) exactly can the result be transferred onto the complete implant range/ all combination possibilities /patient characteristics?
- worst-case testing negative (failed): testing should proceed with further sizes or repeated simulation to establish whether adaptations of the basic design are required, small adjustments of the worst-case are sufficient (but is this then further the worst-case?) or whether specific limitations (e.g. weight limitations for small sizes) may be adequate.

Moreover, when clinical applied for the first time, these scenarios should be limited to a very small cohort of patients if doubts regarding the forecasting power of pre-clinical test results still remain even after application of all pertinent steps.

7. References

1. Affatato S, Ruggiero A. A Perspective on Biotribology in Arthroplasty: **From In Vitro toward the Accurate In Silico Wear Prediction. Applied Sciences 2020; 10(18):6312.**

2. Aitchison GA, Hukins DWL, Parry JJ, Shepherd DET, Trotman SG. **A review of the design process for implantable orthopedic medical devices. Open Biomed Eng J 2009; 3:21–7.**

3. Ali AA, Mannen EM, Rullkoetter PJ, Shelburne KB. **Validated Computational Framework for Evaluation of In Vivo Knee Mechanics. J Biomech Eng 2020; 142(8).**

4. Harrison N, Field JR, Quondamatteo F, Curtin W, McHugh PE, Mc Donnell P. **Preclinical trial of a novel surface architecture for improved primary fixation of cementless orthopaedic implants. Clin Biomech (Bristol, Avon) 2014; 29(8):861–8.**

5. Cristofolini L, Varini E, Viceconti M. **In-vitro method for assessing femoral implant-bone micromotions in resurfacing hip implants under different loading conditions. Proc Inst Mech Eng H 2007; 221(8):943–50.**

6. Colic K, Sedmak A. **The current approach to research and design of the artificial hip prosthesis: a review; 2016. (vol 1).**

RESEARCH TOPIC

EVALUATION OF INSTRUMENTS AND USABILITY

Research Topic 22

Pre-Clinical and Clinical Evaluation of Instruments and Usability with Regard to Handling, Workflow and Functionality

RESEARCH TOPIC 22

Pre-Clinical and Clinical Evaluation of Instruments and Usability with Regard to Handling, Workflow and Functionality

Authors

Alexander Giurea[1], Francesco Benazzo[2,3], Ashley Blom[4,5], Massimiliano Bernardoni[6], Christoph Schilling[7], Francesco Traina[8,9], Richard T. Mayer[10], Søren Overgaard[11,12,13,14]

[1]Department of Orthopedics and Trauma Surgery, Medical University of Vienna, Vienna, Austria

[2]Sezione di Chirurgia Protesica ad Indirizzo Robotico, Unità di Traumatologia dello Sport, U.O.C. Ortopedia e Traumatologia Fondazione Poliambulanza, Brescia, Italy

[3]IUSS, Istituto Universitario di Studi Superiori, Pavia, Italy

[4]National Institute for Health Research Bristol Biomedical Research Centre, University Hospitals Bristol and Weston NHS Foundation Trust and the University of Bristol, Bristol, UK

[5]Translational Health Sciences, Bristol Medical School, Musculoskeletal Research Unit, University of Bristol, Learning & Research Building, Southmead Hospital, Bristol, UK

[6]Medacta International SA, Castel San Pietro, Switzerland

7Aesculap AG Research & Development, Tuttlingen, Germany

[8]Ortopedia-Traumatologia e Chirurgia protesica e dei reimpianti d'anca e di ginocchio, IRCCS Istituto Ortopedico Rizzoli, Bologna, Italy

[9]DIBINEM, Università di Bologna, Bologna, Italy

[10]Johnson & Johnson, Office of the Chief Medical Officer New Brunswick NJ, Zuchwil, Switzerland

[11]Department of Clinical Medicine, Copenhagen University, Copenhagen, Denmark

[12]Department of Orthopaedic Surgery and Traumatology, Bispebjerg and Frederiksberg Hospital, University of Copenhagen, Copenhagen, Denmark

[13]Orthopaedic Research Unit, Department of Orthopaedic Surgery and Traumatology, Odense University Hospital, Odense, Denmark

[14]Department of Clinical Research, University of Southern Denmark, Odense, Denmark

1. Abstract

Question: Evaluation of instruments and usability

Part 1

How can be assessed from a pre-clinical point of view if the handling of an implant including the implant-specific instruments is uncomplicated and if the workflow runs smoothly, achieves the desired results and does not lead to undue stress for patient and surgeon?

Part 2

How can be assessed from a clinical point of view if the handling of an implant including the implant-specific instruments is uncomplicated and if the workflow runs smoothly, achieves the desired result and does not lead to undue stress for patient and surgeon?

Summary / Recommendation:

Part 1

The proper handling of implant-specific instruments used for TKA/THA is an essential requirement and a crucial need to ensure that the desired surgical outcome is guaranteed.

Usability of implant-specific instruments and implants in orthopedics is indicated if the instruments are intuitive, easy to use and not overengineered. No harm neither to the user (surgeon, nurse) nor to the patient is to be expected.

To reach this purpose it is mandatory to consider adequate methods and involve experienced users during development. An analysis of methods and their purpose, as well as the effect of use environment with special focus on implants and implant-specific instruments will be part of this consensus statement. A minimum of 4 (with result saturation) from experienced to non-experienced users is recommended.

Part 2

Beside all described standards and methods about pre-clinical usability testing there is a gap between pre-clinical setup and the clinical practice. This article address parameters from clinical point of view which are necessary for a usability study with implants and implant specific instruments in TKA and THA to reduce or close these gaps. In summary, expert reviews can close the gap by means of systematic questionnaires avoiding usability issues.

2. Level of Evidence

High

3. Consensus Delegate Vote

Part 1

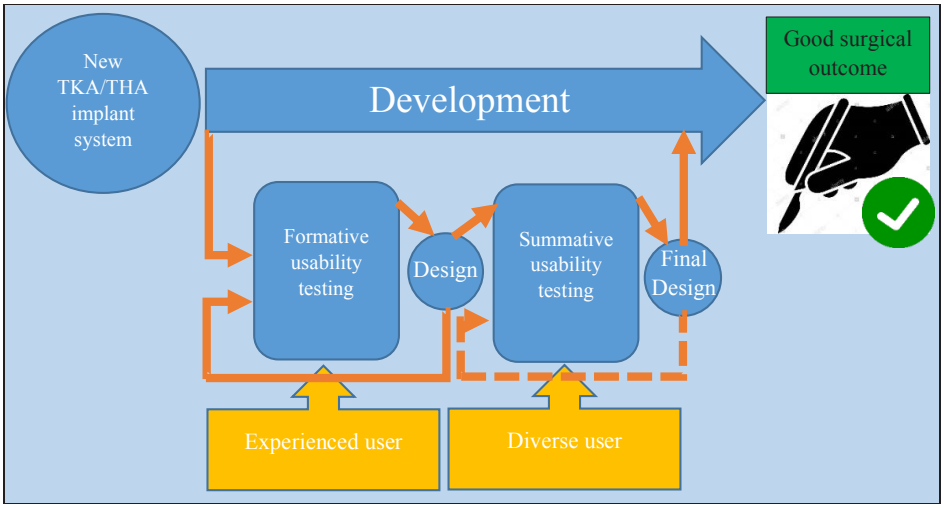
95% – unanimous, strongest consensus (95% agree / 0% disagree / 5% abstain)

Part 2

97% – unanimous, strongest consensus (97% agree / 0% disagree / 3% abstain)

4. Part 1

Graphical Abstract



b. Search Strategy

A systematic literature search was conducted in January 2021 to identify literature with the potential to contribute to the answer of the above mentioned research question.

The electronic databases PubMed and EMBASE were explored using the combination of the following search-terms combined with the Boolean operator "AND":

orthopaedic/endoprosthetic instruments/equipment, hip arthroplasty, knee arthroplasty, medical devices, assessment, testing, study, evaluation

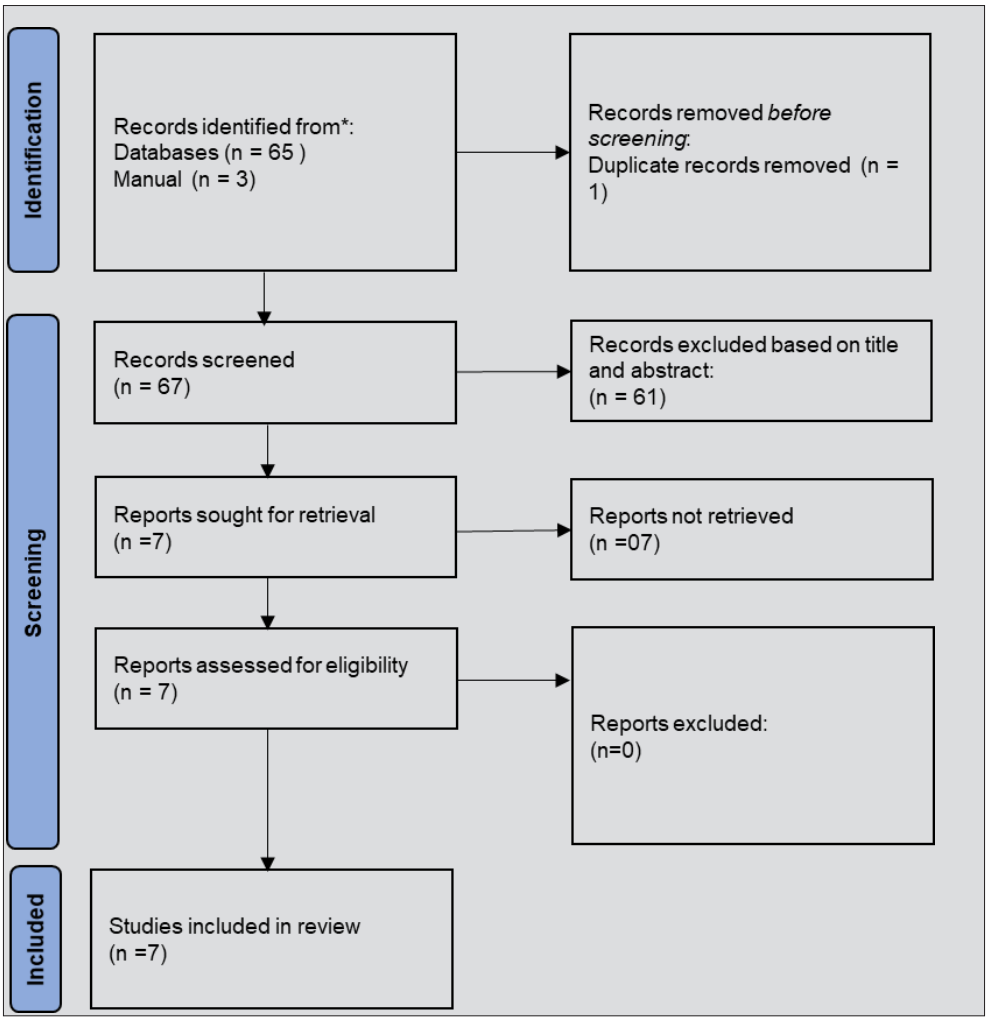
Search words were adapted to the specific search terms of the database and were searched including different spelling and synonyms. Article language was restricted to English, French, German or Italian language. Because of the aim of the process only studies from 2001 to 2021 were included to provide insight into current practice and state of the art procedures. Publication type was limited to papers for which a full text version was available. Abstracts from scientific meetings were not included in the current review as sufficient data cannot be extracted. Where necessary, the results were also restricted regarding study type to guarantee the inclusion of high quality literature.

The title and abstract of all articles were screened to select articles contributing to the topic.

Potentially eligible articles were obtained for detailed assessment and were in detail screened for applicability. The selection was not further restricted regarding basic study design. Duplicate articles were removed.

Abstracts from scientific meetings were not included in the current review as sufficient data cannot be extracted.

65 publications met the search criteria. Manual search and inclusion of relevant Medical Device Safety Standard contributed with additional 3 references. After the selection process,7 articles were included in the review.



c. Rationale

The usability engineering for medical devices described in IEC 62366-1 is one of the common methods described in standards. A Differentiation is made between formative and summative usability studies (1). Similar test methods are used in Human factors engineering (2). Both type of studies must be planned and organized in a manner that relevant topics can be addressed by a suitable method to obtain realistic results.

The formative evaluation is intended to explore user interface design, strength, weaknesses, and unanticipated use errors. Instead of this intention the summative evaluation is conducted at the end of the development, to obtain objective evidence that the product can be used safely (1).

To our opinion implant-specific instruments and implant usability fulfill all criteria mentioned above, so it is necessary that both types of usability evaluations are included in a development process.

Formative studies may have not explicit acceptance criteria for pass or fail. Trial and error on models together with experienced users are able to address following topics (3):

- Acquire use-requirements: At the beginning of a development it is essential to understand the user needs and the context of use. During a usability session with experienced users (maybe on saw-bone with predecessor or competitor implant system) the pain points and possible use error could be detected and addressed in the requirements list. Use of the "think aloud" / "think loud" (2) technique is common practice.
- Collect potential use-errors: During formative usability testing on early prototypes or predecessor / competitor systems it could be helpful to identify and collect (potential) use errors. This is a good addition to the theoretical approach by risk analysis.
- Improve the user interface for the intended use: Formative tests are intended to improve user interfaces, which represents the interface between the user and the medical device. An iterative process is the main idea behind. The user is familiar with the environment and therefore he is able to recognize limitations and potential for improvements.
- Evaluation of the effectiveness and efficiency of the instrumen: It is recommended to observe the effectiveness of the instrument during usability testing. Observable criteria (task fulfillment) are good indicators. E.g.: "User is able to perform tibia cut with the fixated cutting guide and the oscillating saw".

  Efficiency is more subjective based, but in all cases linked to a noticeable reduction of time, force, or resources. E.g.: "User needs 50% less time to align the mechanical axis in comparison to the competitor system.
- Recognition of ergonomic issue: Formative evaluation includes the assessment of ergonomic aspects of the user, preferably in an adequate simulated use environment like OR room. If the handling of the instruments leads the user to uncomfortable positions for longer time, this could affect users' concentration and may result in potential harm of the user and the patient.
- Interfaces and or interactions to other instruments: If there is an interaction between instruments following the surgical workflow it is helpful to observe if the order of use is in a logical and intuitive way. Struggling with the correct order of using the instruments could lead to use-errors.

The summative evaluation is performed with defined formal acceptance criteria in a measurable format. Therefore, a fixed test protocol with defined observations and a questionnaire is the best combination to fulfill this requirements.

In this evaluation phase the general functional principle of instruments and implants must be ensured and is not in the focus of the usability test. A defined test scenario is required with a clear order. A test leader is strongly recommended to guide through the test scenario. In addition, some transcript writer can assess time during the test and ensure an adequate documentation. Audio and video records could support the analysis of the test results and will be helpful to show use errors to the designer team.

Potential use-errors must be investigated. Diversity of users is recommended to cover different anthropometric issues, as well as different mental models. The number of users necessary for valuable test results vary in the literature from 4-5 up to 49 persons (4). For summative studies FDA requirements define 15-20 persons per user group. There are also some statistical data available which can be considered to evaluate the individual need for test purposes (5). In order to fulfill both scientific as well as practical/economic aspects for usability testing of THA and TKA implants and instruments, we recommend a sufficient number of at least 4 test users and until saturation of the results is achieved, spanning the spectrum of experienced and non-experienced users in a two phase approach.

An important factor in the context of usability testing in orthopedics is the use environment (6). Usability studies must be covering the workflow under realistic OR conditions

Based on detailed indicators and questions the adequate setup must be chosen. The planning must include a dedicated evaluation of the test setup environment. Depending on the test question it could be necessary to perform tests on human specimen, if testing on saw-bone or other kind of models will not be able to reach the desired result.

The use environment could affect the handling of instruments and implants in different ways:

- Limitations based on the sterile OR field
- Conflicts of instruments with each other and other equipment on the OR table
- Conflicts depending on the anatomical situation or the surgical approach

A combination of model-testing, cadaveric testing and "dry run" could be necessary to achieve relevant feedback.

- Model-testing: Artificial anatomical models like saw-bones, help to understand and simulate the anatomical situation. This setup can be chosen to test early (plastic) prototypes or check the compatibility between different devices. E.g.: Pins or K-Wires for fixation of cutting guides and the limitation during reaming or sawing.

- Cadaveric testing: Testing on human specimen allows to evaluate the handling of instruments under more realistic anatomical conditions e.g.: in the visualization of the situs. Also, different surgical approaches with influence of the soft-tissue and anatomic structures can be simulated. At least there is a possibility to simulate forces needed for operating the instruments like hammering, chiseling, screwing, reaming, or rasping.
- Dry run: This method is helpful if tests with nurses or CSSD-staff is needed. Topics for assembling or disassembling can be addressed by this method. No test models or specimen are used, only the instrument under evaluation.

Even after summative usability testing there are sometimes open issues, that lead to a change of the design or to clinical evaluation (7). The decision to repeat the usability study is depending on different aspects:

1. Is the risk-benefit ratio still acceptable?
2. Are the necessary adjustments in design (user interface) minor changes?
3. Can these changes be exactly addressed in detail and completely fulfilled without the necessity of user verification?

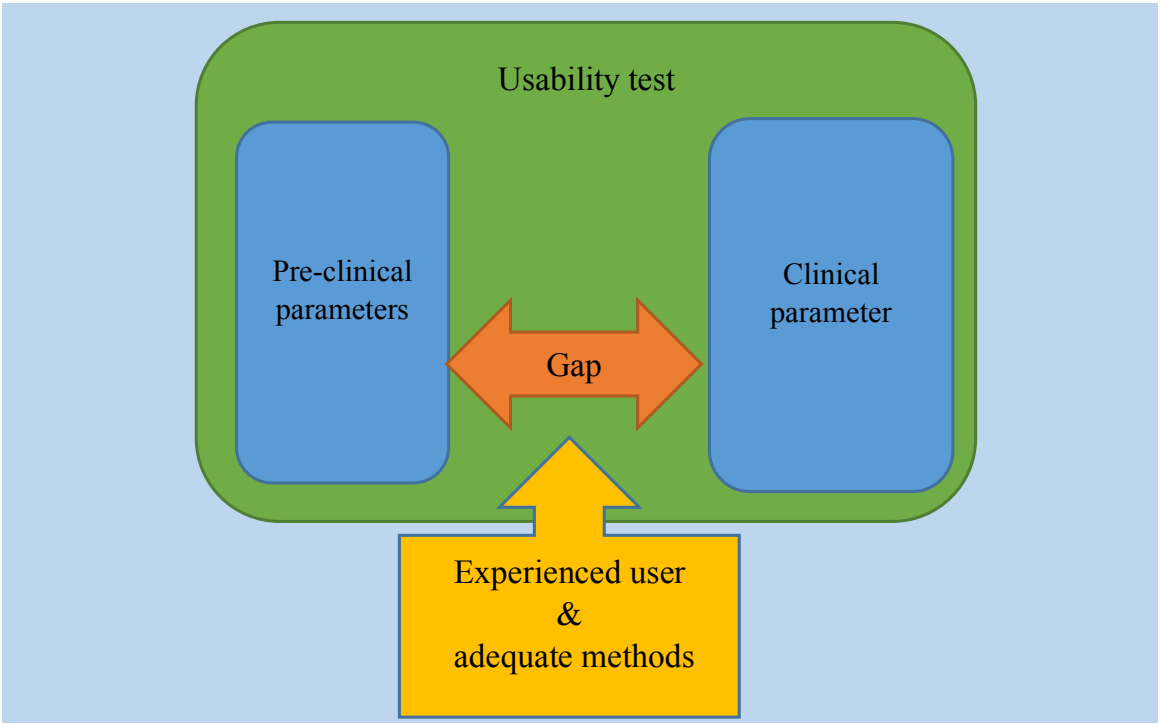
If one of these questions can be answered with "No", repetition of the summative usability study is necessary.

Beside all described aspects of pre-clinical testing there are limitations of all simulated methods and models. There may be individual aspects which must be figured out and taken into account by clinical assessment. Such points are for example, but not limited to:

- behavior of instruments / handling in terms of the individual anatomical situation of patients
- soft tissue quality,
- muscle relaxation
- lack/ presence of tourniquet

5. Part 2

a. Graphical Abstract



b. Search Strategy

A systematic literature search was conducted in January 2021 to identify literature with the potential to contribute to the answer of the above mentioned research question.

The electronic databases PubMed and EMBASE were explored using the combination of the following search-terms combined with the Boolean operator "AND":

orthopaedic/endoprosthetic instruments/equipment, hip arthroplasty, knee arthroplasty, medical devices, assessment, testing, study, evaluation.

Search words were adapted to the specific search terms of the database and were searched including different spelling and synonyms.

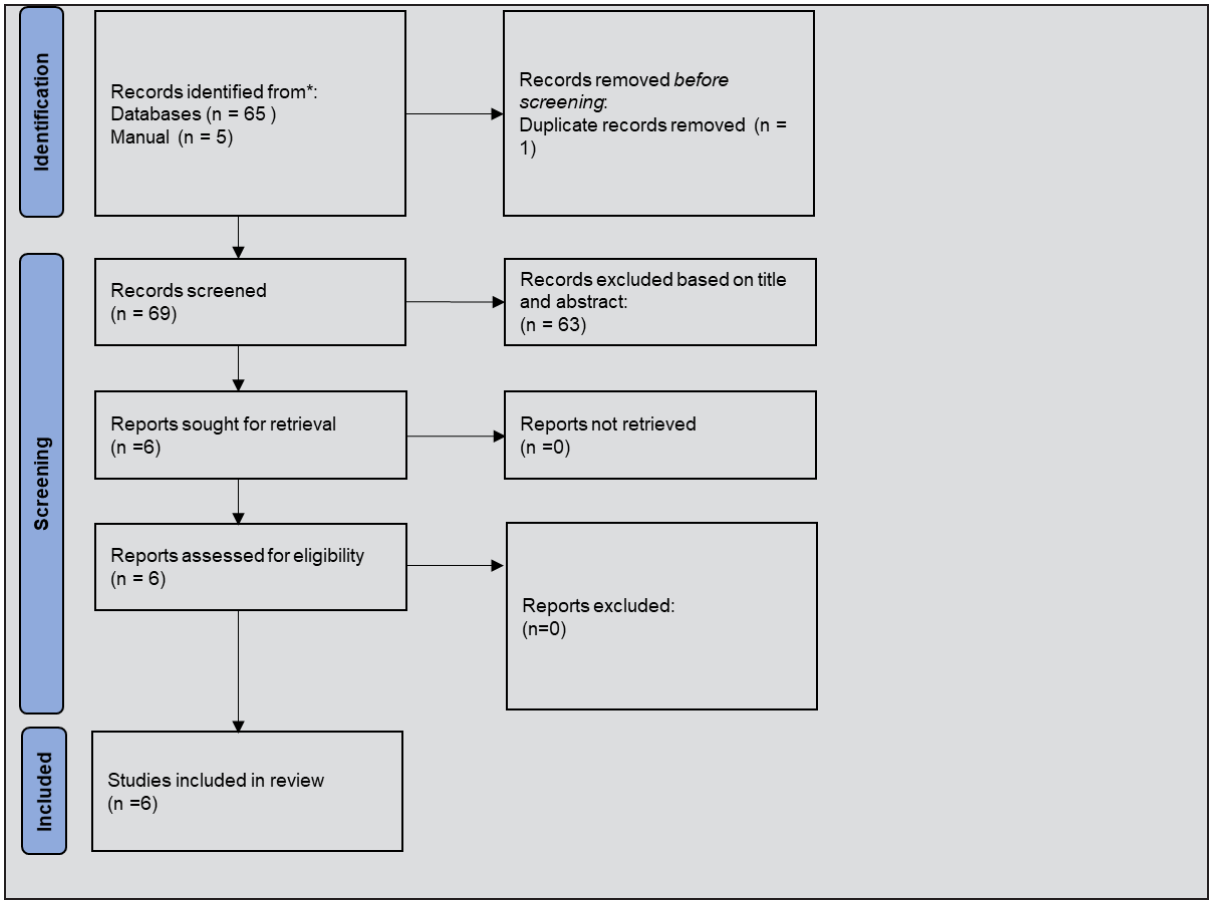
Article language was restricted to English, French, German or Italian language. Because of the aim of the process only studies from 2001 to 2021 were included to provide insight into current practice and state of the art procedures. Publication type was limited to papers for which a full text version was available. Abstracts from scientific meetings were not included in the current review as sufficient data cannot be extracted. Where necessary, the results were also restricted regarding study type to guarantee the inclusion of high quality literature.

The title and abstract of all articles were screened to select articles contributing to the topic.

Potentially eligible articles were obtained for detailed assessment and were in detail screened for applicability. The selection was not further restricted regarding basic study design. Duplicate articles were removed.

Abstracts from scientific meetings were not included in the current review as sufficient data cannot be extracted.

65 publications met the search criteria. Manual search and inclusion of relevant Medical Device Safety Standard contributed with additional 5 references. After the selection process, 6 articles were included in the review.



c. Rationale

In addition to the standards and requirements for usability engineering from a pre-clinical point of view presented in Part 1 of this document, there are also some additional aspects from clinical point of view which are not addressed in standards or literature in detail.

Human factors and usability engineering are focused on the intended user group in order to obtain a medical product which can be used by the intended user group with a high probability that no user-error occur that lead to a high risk for the patient, users or third persons.

Methods for usability test are described as well as requirements for the use environment. Test cases based on hazardous use-scenarios which are evaluated by risk management or in other systematic methods (1, 2).

Clinical aspects relevant for a usability test are addressed by safety and performance indicators inside the clinical evaluation (8). The use environment and the aspects of the OR conditions are relevant parameters such as: Light conditions, operating-height, visibility of situs, infrastructure. Some more important parameters were evaluated by Surma-aho at al. (6).

The different knowledge of users and their experience with (common) reference design must be considered by the evaluation of usability. If users are familiar with a special design and the mental model the interaction will be different (9). If the design of the user interface of an instrument is totally different to (common) comparable designs, a special focus on use-errors must be assigned.

But which aspects should be considered in addition for implants and implant-specific instruments? These are not described in the literature.

An important factor that should be considered in usability tests is the surgical approach. For TKA or THA different surgical approaches are used. The use of instruments must be safe and effective in all intended different approaches. Thus, evaluation of instruments must include the appropriateness for the different intended surgical approaches. If approach-specific variants of these instruments exist, it is also helpful to evaluate if the user is able to identify them.

When clinical workflows have to be investigated the test setup must consider that. It could be sufficient to split different intended surgical workflows to independent tests. In early-stage instruments can be tested separately for their intended use, but in a later phase the interaction of all instruments and implants is essential.

The test environment can be accessed as a simulated clinical environment. To ensure that all relevant clinical parameters are considered, several parameters must be included in the test setup:

- Influence of the bone quality (normal vs. osteoporotic)
- Influence of (known) anatomical defects (hip dysplasia, Genu varum or valgum etc…)
- Influence of anthropologic sizes

If all parameters would be included, the amount of testing would be increase to a level, that no practical realization (i.e. unavailability of specific human specimen for each parameter) or economic development would be possible. (high amount of human specimen, lots of test cases).

In addition, there is a gap between cadaveric testing on human specimen and application in patients. Differences are:

- Visibility due to bleeding
- tension of muscles and soft-tissue, muscles relaxation
- presence/ lack of tourniquet
- sterile operating conditions and behavior
- impact of patients and anesthetic condition on operating time
- stress level for operating team

Another gap which is also very important during usability testing in a simulated use environment is based on the usability test itself. The user is aware about the test situation and about the test scenario. This is different to a real situation in the OR. There is no (alive) patient, there is no mental pressure, there is (maybe) no noise of other devices. This situation leads to the effect that stress situation will not occur during usability testing. On the other hand, the awareness and the attention of the user is not comparable to the real situation in the OR. To close this gap, some unexpected situations can be included in a test scenario. E.g.: loss of sterility of main instruments. This could be excluded in the final evaluation, but it is essential to get the full awareness (back) of the user.

These limitations cannot be addressed in a usability study completely using test cases and covering all questions by test result. As an appropriate technique the expert review (2) could solve this issue. Surgeons (especially experienced surgeons) are familiar with most of the described situations. Therefore, their opinion could be systematically asked and evaluated without having detailed test scenarios.

In summary usability issues in pre-CE studies should be strongly avoided. The main purpose of the usability and human factor engineering is the common way to evaluate the interaction of the medical product with the intended user in a simulated use environment in a pre-clinical study.

For implants and implant-specific instruments which are equal to existing (well known) designs, it is possible to evaluate the user-interface with existing data. Adequate method is listed in IEC 62366-1 (1). There are limitations in this method, especially if parts are different to the comparable design.

6. References

1. ISO - International Organization for Standardization. **IEC 62366-1 Medical devices: Part 1: Application of usability engineering to medical devices; 2015 2015.**

2. FDA - U.S. Food and Drug Administration - **CDRH. Applying Human Factors and Usability Engineering to Medical Devices: Guidance for Industry and Food and Drug Administration Staff; 2016.**

3. Bitkina OV, Kim HK, Park J. **Usability and user experience of medical devices: An overview of the current state, analysis methodologies, and future challenges. International Journal of Industrial Ergonomics 2020; 76:102932.**

4. Bastien JMC. **Usability testing: a review of some methodological and technical aspects of the method. Int J Med Inform 2010; 79(4):e18-23.**

5. AAMI - Association for the Advancement of Medical Instrumentation. **ANSI/AAMI HE75:2009 (R2018) - Human Factors Engineering - Design Of Medical Devices; 2009 2009.**

6. Surma-Aho A, Hölttä-Otto K, Nelskylä K, Lindfors NC. **Usability issues in the operating room – Towards contextual design guidelines for medical device design. Applied Ergonomics 2021; 90:103221.**

7. van der Peijl J, Klein J, Grass C, Freudenthal A. **Design for risk control: the role of usability engineering in the management of use-related risks. Journal of Biomedical Informatics 2012; 45(4):795–812.**

8. MEDDEV 2.7/1 revision 4 CLINICAL EVALUATION: A GUIDE FOR MANUFACTURERS AND NOTIFIED BODIES UNDER DIRECTIVES 93/42/EEC and 90/385/EEC: MEDDEV 2.7/1.

9. Schmettow M, Schnittker R, Schraagen JM. **An extended protocol for usability validation of medical devices: Research design and reference model. Journal of Biomedical Informatics 2017; 69:99–114.**

RESEARCH TOPIC

MODIFICATIONS/ ADJUSTMENTS

Research Topic 23

Requirements and Considerations for the Implementation of Modifications During the PMCF Phase of a Joint Replacement Device Based on the Functional Relevance of the Adjustment

RESEARCH TOPIC 23

Requirements and Considerations for the Implementation of Modifications During the PMCF Phase of a Joint Replacement Device Based on the Functional Relevance of the Adjustment

Authors

Joshua Bridgens[1], Massimiliano Bernardoni[2], Christoph Schilling[3], Sabine Rusch[4], Philippe Massin[5], Ricardo Larrainzar-Garijo[6], Francesco Traina[7,8], Volkmar Jansson[9]

[1]DePuy Synthes, Leeds, United Kingdom
[2]Medacta International SA, Castel San Pietro, Switzerland
[3]Aesculap AG Research & Development, Tuttlingen, Germany
[4]Aesculap AG Medical Scientific Affairs, Tuttlingen, Germany
[5]CMC Ambroise-Paré-Hartmann, Neuilly-sur-Seine, France
[6]Orthopaedics and Trauma Department, Medicine School, Universidad Complutense, Hospital Universitario Infanta Leonor, Madrid, Spain
[7]Ortopedia-Traumatologia e Chirurgia protesica e dei reimpianti d'anca e di ginocchio, IRCCS Istituto Ortopedico Rizzoli, Bologna, Italy
[8]DIBINEM, Università di Bologna, Bologna, Italy
[9]Ludwig Maximilians University Munich, Department of Orthopaedic and Trauma Surgery, Musculoskeletal University Center Munich (MUM), Campus Grosshadern, Munich, Germany

1. Abstract

Question: Modifications/Adjustments

What are additional requirements to implement function-relevant / non-function-relevant modifications/adjustments during the PMCF phase of a device?

Summary / Recommendation:

Based on cumulative experience it can be difficult for manufacturers, surgeons or regulators to be confident whether modifications to devices (implants or instruments) will affect function or not.1-3 We cannot therefore recommend specific requirements based on whether modifications are function relevant or not. This is due to the inherent, and historically proven, difficulty in predicting what the effect of changes will be, regardless of their intent. In light of this challenge we suggest a risk-based approach to assessment of modifications, the level of assurance that can be gained from pre-clinical testing and therefore the implication for subsequent post-market clinical follow up (PMCF) activities.

In cases where there is a high degree of confidence that there will not be any functional effect of the modification no change in PMCF activities is required. Pre-clinical testing is sufficient to provide the assurance that there will be no functional effect.

Any change falling within the scope of this guidance will, by definition, be similar enough to the original implant that it can be considered equivalent and anticipated to have the same safety and performance characteristics. However, there may be cases where, despite being considered equivalent, there is a lower degree of certainty that any changes will not affect function. Pre-clinical testing will help indicate sufficient similarity for equivalence to be accepted, but enough difference for there to be a lower degree of confidence that there will be no change in function. In these cases specific PMCF activities should be specified which will ensure that the clinical safety and performance are unchanged.

It should be noted that any substantial change to design, where there would be the anticipation of altered function with a likely consequence for clinical safety and performance would necessarily require, as a minimum, a new device application on the basis of equivalence. This recommendation only applies to modifications or adjustments which are more minor than this.

Recommendation: In view of historical experience, and the limited scope for modification allowed by the MDR, we suggest a risk-based approach to assess and decide if any modifications or adjustments to implants or instruments are likely to be function relevant and, dependent on this assessment, what the implications for further PMCF activities would be.

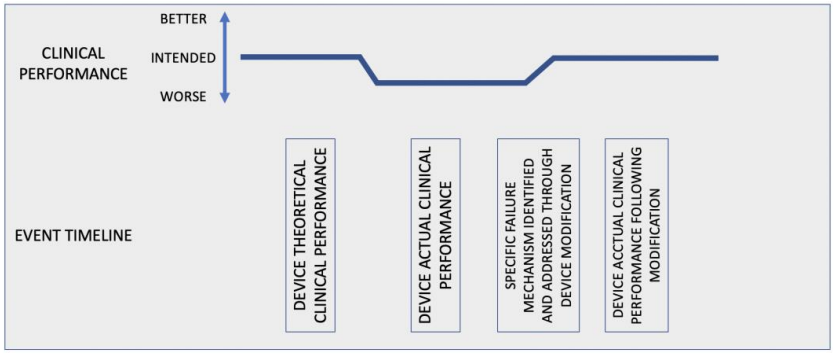
2. Level of Evidence

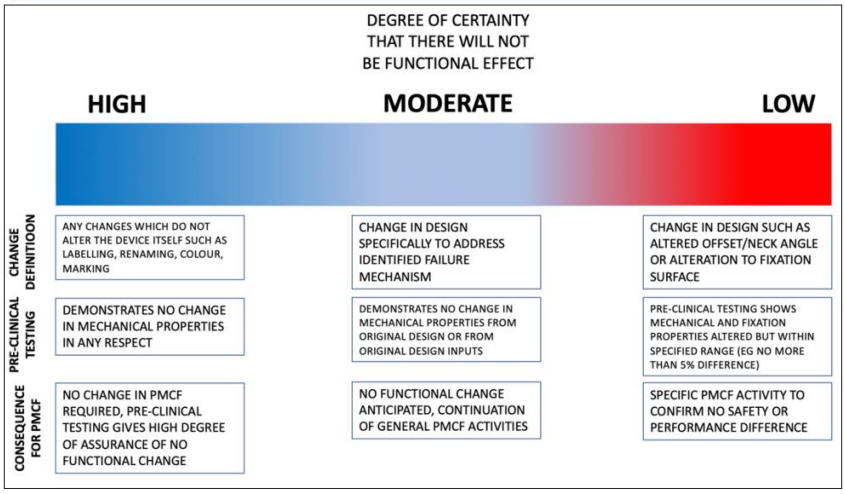
Low

3. Consensus Delegate Vote

95% - unanimous, strongest consensus (95% agree / 2.5% disagree / 2.5% abstain)

4. Graphical Abstract





5. Search Strategy

No formal literature search was conducted on the basis that the question was so specific to the detailed requirements of the EU MDR that no publication would address these specifics. Informal searches were conducted which confirmed this expectation. There are publications which detail the potential unintended consequences of design modifications but due to the experience and awareness of these within the expert group it was felt that a formal search focusing only on this area would be duplicative and redundant. If was not felt that appropriate search terms could be developed or that, even if it were possible, they would give any further useful evidence in relation to the technical question posed.

6. Rationale

Modifications or adjustments to devices already on the market ("in the PMCF phase") have historically been common. They have been made to address an identified failure mechanism (to improve the design of the device) or to increase the utility of the device to user surgeons. This may have been through increasing the range of sizes or altering certain characteristics such as offset length of a hip stem.

It is generally recognised that sometimes what appear to be relatively minor changes can have significant and unanticipated effects on the performance of the implant. One of the most well known examples of this is the change of the Exeter stem from a polished to a matt finish.3 Rockborn and Olsson, in their 1993 publication, commented that "The introduction of the matt stem was not preceded by biomechanical analysis or by testing and if this design feature is the main reason for the inferior results, it is an illustration of the fact that seemingly minor alterations may have serious and unforeseen clinical consequences."

These observations are applicable to instruments as well as implants as we have used the term 'device' to refer to both.

It is therefore clear that robust pre-clinical testing and justification for any modifications or adjustments that may affect device safety or performance ("function-relevant") is required. The problem is that it is not easy to determine if any change is likely to be function-relevant, particularly if understanding of the devices' mechanism of action is incompletely understood.

It is also necessary to take into consideration the regulatory requirements presented by the EU MDR relevant to the question posed. Any significant modification or adjustment to a device in the PMCF phase would only be 'allowable', i.e. the device could only maintain the same CE mark, if it were able to be considered to be 'equivalent' (in the regulatory meaning of the word) to the original device. This mandates that the design "shall be similar to the extent that there would be no clinically significant difference in the safety and clinical performance of the device". The scope for any 'function-relevant' modifications is therefore limited by the scope provided by the expectation that the modification, although changing the device function, would not lead to any "clinically significant difference in the safety and clinical performance of the device".

We therefore felt that one of the most likely reasons for modification of a device would be where a manufacturer has identified, through their post-market surveillance, a failure mechanism which is occurring more frequently than anticipated in their initial design risk assessment. This would mean that the function, and potentially performance of the device, was compromised from that which was originally intended. A design change would then be made, not with the intention of altering safety or performance, but with the intention to return it to the originally intended level. This concept is demonstrated graphically in Figure 1 below.

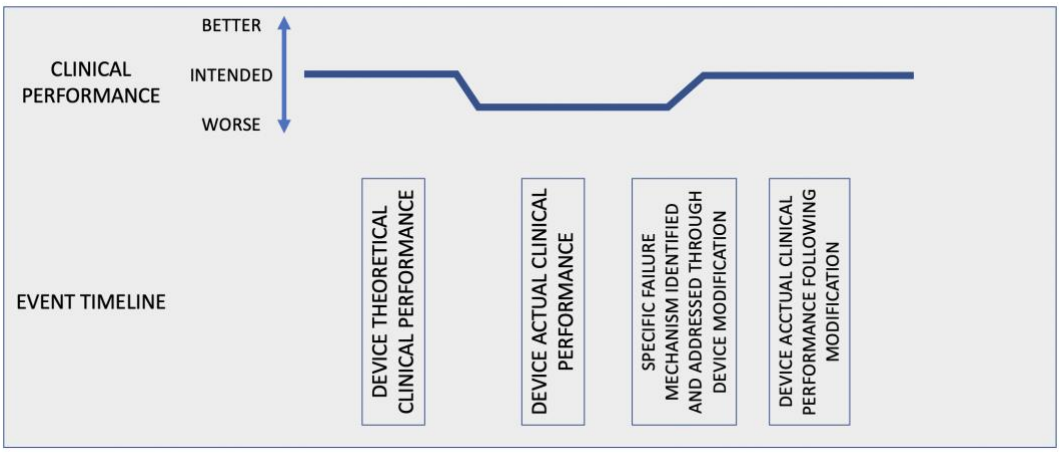


Figure 1

In view of historical experience, and the limited scope for modification allowed by the MDR, we suggest that the most important question is how to assess and decide if any modifications or adjustments are likely to be function relevant and, dependent on this assessment, what the implications for further PMCF activities would be. It must be remembered that this is all within the envelope of a change that would not be expected to lead to any "clinically significant difference in the safety and clinical performance of the device" even if device function has been altered.

We have developed an algorithm that may be of use for manufacturers and regulators in considering if a modification is likely to be function-relevant (Figure 2).

With regard to the information in Figure 2, we will consider each risk level and the implications of this separately.

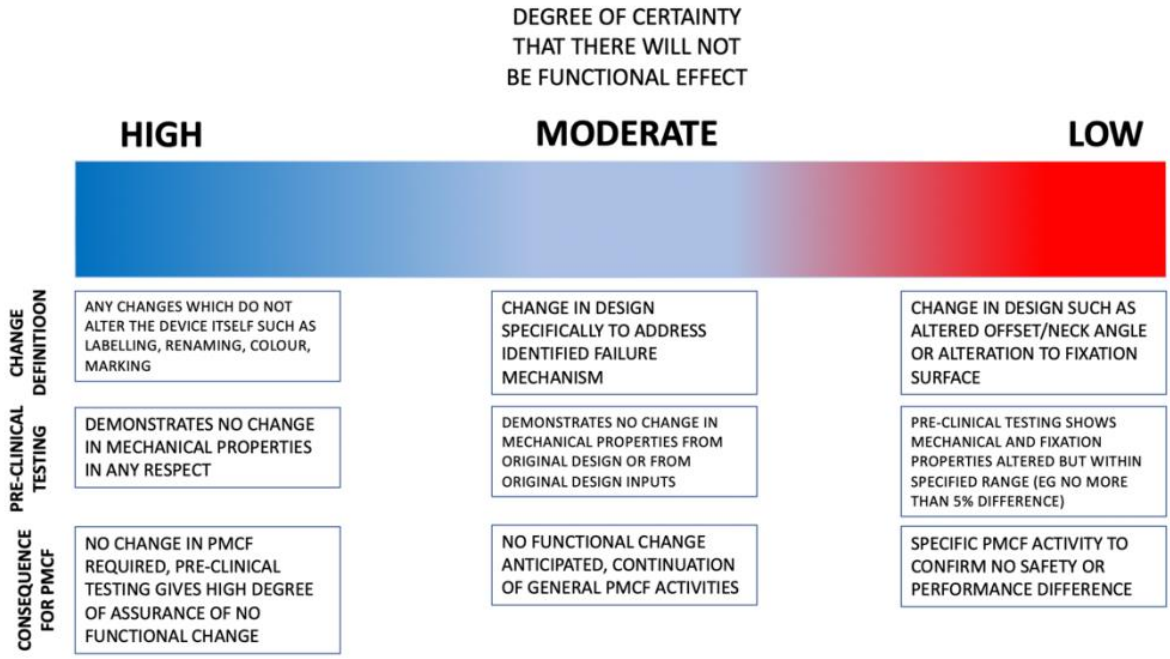


Figure 2

High degree of certainty that there will not be a functional effect

We felt that there could be changes to a device where there could be a high degree of certainty that they would not have a functional effect. These changes are summarized in our "change definition" as "any changes which do not alter the device itself such as labelling, renaming, colour, marking".

Even here there is a requirement for pre-clinical testing as some of the changes included within the change definition, such as a change in marking, while not intended to have any effect on the implant could conceivably do so (for example laser etching leading to a stress riser in the implant). It is therefore still necessary for pre-clinical testing to demonstrate that the mechanical properties of the implant have been unchanged "in any respect". If these conditions are met, then there will be a high degree of confidence that there would be no functional effect and so no change in subsequent clinical follow up (PMCF) will be required.

Moderate degree of certainty that there will not be a functional effect

In this case the change definition refers to a "change in design specifically to address an identified failure mechanism". We felt that this was the most likely scenario whereby a change would be made to a device which would not be anticipated to alter its clinical safety or performance, and so would fall within the scope of equivalence. An example scenario could be total knee replacement tibial tray locking mechanism. The original design met design inputs on original testing but, when used in clinical practice, post-market surveillance showed a higher failure rate than anticipated. The overall performance of the device still fell within acceptable limits but it was felt that the design of the locking mechanism could be improved in order to meet the original design inputs. The new design was not intended to alter clinical safety or performance, but to return these to the originally intended level (see Figure 1).

In this scenario, pre-clinical testing should show "no change in mechanical properties from the original design" or no change "from original design inputs". As long as this is the case no functional change, from that originally intended, would be anticipated. Continuation of originally planned PMCF activities would therefore be appropriate.

Low degree of certainty that there will not be a functional effect

In this case the change definition refers to a "change of design" which would be clearly recognized as such by users, for example change in a femoral neck offset or a total knee component fixation surface.

Pre-clinical testing would demonstrate that there is a difference in mechanical properties but this is within a narrow range of the original design (for example, no more than 5% difference). The exact reasonable numerical degree of difference cannot be defined generally as it will vary according to the specifics of the individual case and must be justified based upon solid considerations.

For these types of design changes it is not necessarily the case that there will be an effect on clinical safety or performance, in which case the device would no longer be equivalent to its original design. This is because the design change may be to accommodate patient anatomy (eg the femoral offset change) and so improve performance in a group of patients with particular anatomy, or to improve fixation in a particular patient group, for example the elderly with more osteoporotic bone.

However, as the degree of certainty is low, there should be a specific PMCF activity put in place to monitor the safety and performance of device with the particular design change. The PMCF activity should be focused on outputs of relevance to the design change. For example, if only the fixation surface has been altered PMCF outputs should focus on device fixation and survival, rather than other aspects of function, such as patient pain relief or range of motion, which are very unlikely to be affected.

This risk-based approach is considered reasonable and can be useful to manufacturers, user surgeons and regulators in their assessment of design changes during the PMCF phase of device release.

7. References

1. Lewis PL, Graves SE, de Steiger RN, Campbell DG, Peng Y, Hatton A, Lorimer M. **Does Knee Prosthesis Survivorship Improve When Implant Designs Change? Findings from the Australian Orthopaedic Association National Joint Replacement Registry. Clin Orthop Relat Res. 2020 Jun;478(6):1156–1172**

2. Janssen D, Aquarius R, Stolk J, Verdonschot N. **Finite-element analysis of failure of the Capital Hip designs. J Bone Joint Surg Br. 2005 Nov;87(11):1561–7**

3. Rockborn P, Olsson SS. **Loosening and bone resorption in exeter hip arthroplasties. Review at a minimum of five years. J Bone Joint Surg Br. 1993 Nov;75(6):865–8**

RESEARCH TOPIC

PRE-CE STUDIES/SAFETY STUDIES

RESEARCH TOPIC 24

Potentials and Limitations of a Pre-CE Study (or Safety Study) in the Field of Arthroplasty

Authors

Ashley Blom[1,2], Dario Bergadano[3], Ina Wüstefeld[4], Justin Cobb[5], Faris Haddad[6], Marcus Jäger[7,8], Hassan Achakri[9], Matthias Fink[10], Arndt-Peter Schulz[11,12]

[1]National Institute for Health Research Bristol Biomedical Research Centre, University Hospitals Bristol and Weston NHS Foundation Trust and the University of Bristol, Bristol, UK

[2]Translational Health Sciences, Bristol Medical School, Musculoskeletal Research Unit, University of Bristol, Learning & Research Building, Southmead Hospital, Bristol, UK

[3]Medacta International SA, Castel San Pietro, Switzerland

[4]Aesculap AG Medical Scientific Affairs, Tuttlingen, Germany

[5]Imperial College London, MSk Lab, Sir Michael Uren Biomedical Engineering Research Hub, Imperial College London, White City Campus, London, UK

[6]University College London Hospitals, The Princess Grace Hospital, and The NIHR Biomedical Research Centre at UCLH, London, UK

[7]Department of Orthopaedics, Trauma & Reconstructive Surgery, St. Marien Hospital Mülheim an der Ruhr, Mülheim, Germany

[8]Orthopaedics and Trauma Surgery, University of Duisburg-Essen, Essen, Germany

[9]Zimmer Biomet, Clinical Affairs, Zug - Switzerland

[10]TÜV Süd group, Munich, Germany

[11]Department of Trauma Surgery, Orthopedic and Sports Traumatology, BG Klinikum Hamburg, Hamburg, Germany

[12]Fraunhofer Research Institution for Individualized and Cell-Based Medical Engineering, Lübeck, Germany

1. Abstract

Question: Pre-CE studies / safety studies

What are potentials and what are limitations of a pre-CE study (or safety study) in the field of arthroplasty?

1. Which aspects can be assessed reliably? (surgical approaches, indications, broad application etc.)?
2. What amount of work and effort generates a pre-CE study for clinicians and in the hospital?
3. Which study methods are generally considered reasonable to yield reliable results?
4. What are alternatives to conventional/traditional pre-CE studies or are there additional ways to acquire supplementary data/ information?

Summary / Recommendation:

Benefits:

- Safety: An implant migration study coupled with a clinical trial using a number of surgeons and centres should be able to demonstrate safety with an acceptable non-inferiority margin.
- Efficacy: A clinical trial with 2 years follow up and adequate confidence intervals should be able to demonstrate efficacy and usability in a number of centres with a number of surgeons, with an acceptable margin of non-inferiority.

Some of the limitations include:

(i) Some of the clinical outcomes are designed exclusively for regulatory application purposes and may not be clinically relevant representations of patient outcomes

(ii) Each clinical trial has a unique composite end point, which make comparisons between studies difficult.

(iii) Due to their regulatory designs, it is not practical to keep patients blinded with regards to treatment received during the postoperative follow-up when conducting radiographic evaluations as part of outcome assessments, which could potentially cause patient bias.

(iv) Most of the studies have 24 months postoperative endpoint follow-up. Any clinical benefit of a procedure will be evident by this time, but longer term disadvantages may not have become evident by then.

(v) health economists do not routinely attribute more than 5 years benefit to any given procedure. When this is the case, the longer term benefits of arthroplasty will be underestimated.

(vi) Pre-CE studies are limited with respect to sample size of patients and surgeons participating as investigators. The results from the studies may not be representative for all patients population and surgeon skills. They are thus at risk of not being generalisable.

(vi) Extended time for surgeons and patients' access to improved but not innovative devices and/or proven technologies (e.g. not novel highly cross-linked polyethylene).

1. Several aspects can be reliably assessed which include preoperative preparations, operative surgeon(s), surgical approaches and techniques, implants and materials, clinical outcomes, and postoperative instructions. There is no direct data on which aspects can be assessed more reliably. Some of the clinical outcomes are designed exclusively for regulatory application purposes and may not be clinically relevant representations of patient outcomes. Pre CE studies must be small and highly regulated, focussed on non-inferiority, although superiority studies are also permissible, with existing techniques and technologies. Extrapolation from these data to include a variety of approaches and indications will properly be the role of the 'controlled release' advocated by 'Beyond Compliance' mechanism in the UK for instance as well as national registries.
2. Pre-CE studies of novel joint replacements are usually, but not exclusively, sponsored and run by the company manufacturing the device, who will usually pay for the trial to be run by a clinical trial centre or organisation. These costs, and any additional costs to the clinicians and hospital beyond the standard of care, should properly be funded by the company. In outline these tasks include:
    - Initial preparations, patient recruitment, consenting and trial related follow up
    - Surgeon, study nurse and theatre team training and ongoing support
    - Adverse events documentation and reporting
    - Ongoing audit of trial compliance Interim and final report generation
    - File preparation, archiving and audits/inspections
3. Provided the regulators believe that there is a prima facie case that the new device is substantially similar to an existing device, pre-clinical establishment of equivalence followed by a post-market clinical follow-up preferably nested within a national registry is acceptable. For novel devices, a single arm first-in-human safety study should be undertaken first, to confirm that the procedure can be completed safely with the instruments as intended. A performance study may be single arm, but should be multi-surgeon and multi-centre to demonstrate usability and increase generalisability. Randomized control trials remain the highest standard for documenting clinical safety and efficacy of a class III device. In certain circumstances, such as for innovative high risk devices, they may have a role in pre-CE evaluation. They may be double-blinded, partial blinded or open. Prospective data compared with historical controls can also be used.
4. If a safety and efficacy study has been undertaken in another well managed jurisdiction, such as the USA, there is a good argument to be made that the data submitted for FDA approval should be acceptable for submission to the appropriate notified body, to satisfy the MDR.

2. Level of Evidence

Strong

3. Consensus Delegate Vote

95% - unanimous, strongest consensus (95% agree / 0% disagree / 5% abstain)

4. Graphical Abstract

N/A

5. Search Strategy

Data Sources: MEDLINE; clinicaltrials.gov; systematic reviews of relevant studies in arthroplasty that have worked with the FDA to identify summaries and effectiveness of relevant premarket application trials;(1, 2) and reference lists of relevant studies.

6. Context

Implantable Orthopaedic devices are designed for a number of treatment aims including aiding the healing of fractures and alleviating pain and functional impairment due to arthritis. The treatment is efficacious if it achieves the aims for which it was designed. This should not be confused with longevity/ implant survivorship, which is a measure of how long an implant remains in situ before being removed and/or replaced. Implantable devices need to first and foremost be safe and efficacious. It is also desirable that they have longevity and cost effectiveness.

New implantable devices may range from being completely novel in their design or substantially similar to existing devices. The more they differ from existing implants, the greater the risk that they will behave in a substantially different manner. They are not associated with greater risk per se, but due to their novelty are more likely to behave in a less predictable manner. A greater level of scrutiny would thus be warranted.

A balance needs to be struck whereby treatments are safe, yet innovation is not stifled. Greater regulatory burden may improve safety at the expense of innovation. This may ultimately be at the detriment of patients. Conversely insufficient regulatory control can lead to extensive patient harm.

Furthermore, the context in which implantable devices are used is important. The severity of the condition and the efficacy of existing treatments needs to be considered. For example if a new device is developed to treat a severe condition for which there are no existing efficacious treatments then the level of acceptable risk may be higher than for a device developed to treat a mild condition for which efficacious treatments already exist.

7. Background

Both in Europe and the US, medical devices are categorized into different classes based on the risk they pose to patients, with class III devices representing the highest risk category. Orthopaedic implants used for arthroplasty are classified as class III devices.(3) Before medical devices are allowed onto the market, they undergo premarket (pre-CE in Europe, pre-CA in the UK) evaluation. The premarket regulations of the U.S. (Food and Drug Administration, FDA) and Europe used to be very different despite the creation of a Global Harmonization Task Force for device regulations in 1992 (www.ghtf.org).(3) The European system (applicable in EU, Norway, Iceland, Liechtenstein and Switzerland) is now called the Medical Device Regulation (2017/745/EU) and like the U.S. system requires the premarket demonstration of safety and performance.(3)

In the US, the regulatory pathways used by the FDA include the 510(k) premarket notification and the premarket approval (PMA). A 510(k) is a premarket submission made to FDA to demonstrate that the device to be marketed is as safe and effective, that is, substantially equivalent, to a legally marketed device.(4) The submitter may market the device immediately after 510(k) clearance is granted. It is by far the one most commonly used regulatory pathway. Historically, this proof of substantial equivalence allowed manufacturers of medical devices to avoid clinical trials, which is a hallmark of the approval process for new drugs or new innovative devices. As a result, medical devices which were by definition not innovative were approved without clinically demonstrating safety or effectiveness.(5) A PMA is the required process for any innovative medical device which would include any novel joint replacement. This comprises scientific review to ensure the safety and effectiveness of Class III devices. It typically requires clinical trials with human participants preceded by laboratory testing. Innovative class III devices such as orthopaedic implants typically undergo a PMA process. In the US, premarket trials are registered on clinicaltrials.gov. The summaries of safety and effectiveness for all premarket application trials are made available on the FDA PMA database.

Historically the Medical Device Directive 93/42/EEC required a pre-market clinical investigation for implantable and high risk (Class III) medical devices. For Class III medical devices (e.g. hip and knee prostheses) the clinical data was assessed by a Notified Body before the issueingt of the CE (Conformité Européenne) mark. The demonstration of equivalence to an already marketed device was less stringent under the Medical Device Directive (MDD) compared to the now applicable Medical Device Regulation. This led to a high number of medical devices being placed on the European market for which clinical data on the device itself was only collected in the post-market phase.(3)

This premarket evaluation of a device by a Notified Body is performed only once for the entire European market. It includes the assessment and verification of the clinical evaluation, defined as the assessment and analysis of clinical data pertaining to a medical device to verify the clinical safety and performance of the device when used as intended by the manufacturer. With the introduction of the Medical Device Regulation (MDR), a process begun in 2017, and finalised in May 2021, the EU now requires that the manufacturer demonstrates safety and

performance and the device has an overall positive benefit-risk ratio when compared with the State-of-the-Art for similar devices. However, there is no published standard in the MDR, so manufacturers have to provide their own evidence from a pre CE marked clinical evaluation. Premarket clinical trials and data in Europe were not mandatorily made public. The European Clinical Trial Register does not provide information on clinical trials for medical devices.(3) Under the MDR, this information will be publicly available in the so called "Summary of Safety and Clinical Performance" report that will be published in the European database on medical devices (Eudamed).

Under the MDR manufacturers are required to send a study report to the national CA of the member states in which the study was conducted. However, overall, only half of the studies of high-risk medical devices are published and there are often discrepancies in important study features between the summary provided to the regulatory agencies and the eventual publication in peer-reviewed journals.(6–8) Most of the publications available are based on post-marketing studies.

In the United States, since 2007, full transparency is assured by the obligatory registration of trial protocols and study results, both for drugs and devices. There have been calls for more transparency in Europe; clear need for a public trial registry (9) and a publicly available summary describing the basis for granting a CE mark.(10) This will be the case now under MDR using Eudamed.

It is noteworthy that other regions do have alternative regulatory processes with components which could be considered in the European context. For example, in Australia there is a restriction on manufacturers charging for the use of implants as opposed to restriction on the use of implants whilst still under regulatory review.

8. Rationale

In a prospective, randomised, multicentre FDA investigational device study of the ProDisc-L total disc replacement compared with circumferential arthrodesis for the treatment of two-level lumbar degenerative disc disease, the authors reported several design limitations in addition to the relatively short duration of 2 years.(11) For the purposes of FDA evaluation, a composite, binary end point consisting of ten criteria was established for each treatment group. The authors acknowledged that this endpoint was designed exclusively for regulatory purposes and was not a clinically relevant representation of patient outcomes. They also stated that it was not practical to keep patients blinded postoperatively regarding the treatment received because of the use of radiographic evaluation and the posterior incision required by posterolateral arthrodesis.(11) In another FDA investigational device study, the authors acknowledged that FDA criteria for overall success was a mathematical measure and did not directly measure clinical success.(12)

1. Many aspects of a pre-CE study can be controlled. Previous studies have controlled aspects including preoperative preparations (e.g., templating), operative surgeon(s), surgical approaches and techniques, implants and materials, and postoperative instructions. Delamarter and colleagues acknowledged in their study that the composite endpoint was designed exclusively for regulatory purposes and was not a clinically relevant representation of patient outcomes.(11)
2. The PMA process for orthopaedic devices (class III) requires clinical trials,(13) which can be costly and time-consuming. Properly conducted definitive trials are more difficult to run with orthopaedic interventions/implants compared with pharmacological agents and other interventions.(14) They are labour intensive, expensive, and they have a late response given the demand for long-term follow-up. Most premarket trials of arthroplasty devices have reported follow-ups of 24 months postoperatively(11, 12, 15, 16). The use of migration data as a surrogate for longer term revision rate for aseptic loosening, now enables this endpoint to be used. The average submission-to-approval time for medical devices through the PMA route has been reported to be 27 months.(13, 17) The regulations in the EU differ to those in the US. According to EU Regulation 2017/745, the manufacturer can "specify and justify the level of clinical evidence necessary to demonstrate conformity with the relevant general safety and performance requirements."(18, 19) In the case of novel orthopaedic implants, which are class III devices, clinical investigations are required, with the exception of the equivalence pathway in MDR, Art. 61.4 and 61.5 .Simulators can be used in preclinical testing; however, simulators make several large assumptions about the mechanics when compared to in vivo use, given the heterogenous patterns of human use of joints; therefore, extrapolation from simulation to human use is currently often a major approximation.(20) It can be challenging to assess how long it will take a modified material to structurally survive, or how long it will take to establish that completely unexpected complications will arise in the preclinical setting.(20) This leaves a dilemma in the advancement of implants. These are not evaluated in premarket trials. National registries are used to report information on implant survival. They are not intended to estimate device benefits but function as a post-market safety surveillance mechanism. Hence, they are considered only in the systematic appraisal of harms.(1) When registries include Patient Reported Outcome Measures, they can be used to assess clinical benefit.
3. In both Europe and the US, medical devices are categorized into different classes based on the risk they pose to patients, with class III devices representing the highest risk category. Orthopaedic implants used for arthroplasty are classified as class III devices. (3) Before medical devices are allowed onto the market, they undergo premarket (pre-CE in Europe) evaluation. The premarket regulations of the United States (FDA) and Europe remain very different despite the creation of a Global Harmonization Task Force for device regulations in 1992 (www.ghtf.org). The European system is based on the demonstration of safety and performance, whereas the US system requires the premarket demonstration of safety and efficacy/effectiveness. This leads to different clinical trials. For the demonstration of device performance, an RCT is neither necessary nor appropriate, whereas it is the most robust method of demonstrating clinical efficacy in a controlled way.(3) In the US, the FDA approves novel, high-risk medical devices through the PMA process, which requires clinical trials with human participants preceded by laboratory testing.(3) Prospective data compared with historical controls can also be used in the absence of a clinical trial. In an FDA investigational device exemption study to compare survival, Harris hip scores, and osteolysis and component migration between a new ceramic-on-ceramic bearing and a ceramic-on-polyethylene bearing, an initial prospective safety study was conducted followed by a prospective, randomized study.(21) There is no requirement in Europe to demonstrate through randomised trials the clinical efficacy of high-risk devices

in the premarket phase. The net result was an earlier market introduction of devices in Europe compared with the United States, but at the risk of insufficiently documented efficacy and safety. Now, following the introduction of the MDR, trials used for premarket evaluation in Europe are based on typically single-arm trials which have to demonstrate safety and "performance". Trials with less than 100 patients are at risk of not sufficiently documenting efficacy and safety for the devices.(3) Sauerland and colleagues conducted a cross-sectional analysis to assess the methodological quality of pre-market clinical studies performed on medical devices in Europe.(22) They analysed 122 study applications of which the proportion of studies on class I, IIa, IIb and III devices was 10%, 15%, 28% and 39%, respectively. In 70 (57%) of the applications, the planned design was a double or partially blinded RCT, with the remaining studies being controlled (non-randomised) and studies with no control groups. No orthopaedic studies were evaluated in this analysis.(22) Randomised control trials are the highest standard for documenting clinical safety and demonstrating any incremental efficacy for class III devices. High quality RCTs are possible even if these trials involve no or only the partial blinding of patients, physicians, or outcome assessors. (23) Randomized controlled trials with sham (or placebo) surgery as a comparator for proving device efficacy may sometimes be considered unethical. It may be more appropriate to document efficacy by comparing to the reference treatment (watchful waiting, another device, drug treatment, conventional surgery, etc.).(3) To not expose too many patients unnecessarily to a high-risk intervention, the conduct of an RCT should not be delayed.(3) Nieuwenhuijse and colleagues recently conducted a systematic review to evaluate effectiveness and safety evidence concerning the introduction of five innovative, relatively recent, already widely implemented device technologies used in major total joint replacement.(1) These included ceramic-on-ceramic bearings, modular femoral necks, and uncemented monoblock (not metal-on-metal) acetabular cups in total hip replacement, and high flexion implants and gender specific implants in total knee replacement. The authors worked with the FDA to identify the summaries of safety and effectiveness for all premarket application trials. National registries were used as sources of safety data. Most of the comparative effectiveness data were based on RCTs of varying quality, with some based on low to moderate quality retrospective comparative (non-randomised) studies. The authors in their conclusion stated that "We did not find convincing high-quality evidence supporting the use of five substantial, well known, and already implemented device innovations in orthopaedics." Hulstaert and colleagues report that based on a large subset of protocols that were reviewed in 2010, representatives of Belgian Ethics Committees reported that nearly all trials with high-risk devices in Belgium were conducted after the CE mark had been obtained, and that the identified pre-CE mark trials were not randomised.(3) Single arm clinical trials are a practical study design for novel joint replacement devices with controlled data derived from an alternative procedure, such as osteotomy or arthrodesis, or an alternative type of arthroplasty, such as total replacement vs partial. Such studies need to be explicit as to selection criteria and the demographics of the intervention group compared to any historical comparator data. For certain outcomes, such as mortality, underlying secular trends and the effects of differing co-interventions such as type of anaesthesia are extremely important. Where relevant, migration analysis techniques such as RSA or CTSA are informative and they can act as surrogates for longer term revision rates (24). The size of this migration analysis study may be small, confirming device stability thus answering the question whether the instrumentation allows the surgeon to perform the procedure adequately. A pre-CE study by definition studies a novel device, and thus may have no preliminary clinical data at all. This places a significant burden on the clinicians and the hospital, as every data point is important, and a small number lost to follow up may invalidate the findings. The pre-CE study should be powered to demonstrate 'non-inferiority' with a margin of non-inferiority that is appropriate to the device and its clinical use. Typically this will have an initial safety study, and a larger performance group, and may include migration analysis. The size of the larger group will be determined by the 'margin of non-inferiority'. This will be class specific, but has to be practical (25). In the cardiac world, in a trial of transluminal versus open aortic valve surgery, a margin of non inferiority of 10% was considered appropriate based upon the relatively high mortality rate of untreated patients (26). For procedures with high success rates such as hip replacement, the margin may be smaller, perhaps 5% than for those with higher rates of clinical failure such as total knee replacement may require a margin as large as 20% (27).

4. Approval by the FDA of novel joint replacements now requires a prospective clinical study with 24 months follow up of 400 cases. This clinical data should be acceptable for submission under MDR. The role of MDR should be to enable citizens of the EU to gain access to improved technologies in a timely fashion. The EU MDR allows the use of clinical data obtained outside the EU to be used for the initial CE-mark if the study were conducted under accepted standards (ISO 14155, GCP). Insisting on separate trials in the EU is not necessary and will delay access. In the US, the regulatory pathways used by the FDA include the 510(k) premarket notification and the PMA(5). A 510(k) is a premarket submission made to FDA to demonstrate that the device to be marketed is as safe and effective, that is, substantially equivalent, to a legally marketed device. The submitter may market the device immediately after 510(k) clearance is granted. It is by far the one most commonly used. In most instances, this proof of substantial equivalence allows manufacturers of medical devices to not use clinical trials. Several high-risk class III devices have been cleared via this less-rigorous 510(k) pathway, owing to a legal provision in the Medical Device Amendments Act,(28) allowing certain classes of high-risk devices to be exempt from the PMA process. As a result, most medical devices are approved without demonstrating safety or effectiveness.(5) In the field of orthopaedics this includes metal-on-metal hip prostheses, cement spacers, and spinal pedicle fixation systems.(5) Although the 510(k) premarket notification allows medical device innovations to reach patients promptly, safety issues have emerged. Devices approved through the 510(k) route have been shown to result in an 11.5-fold increased risk for recall when compared with PMA-authorized devices. Premarket trials of arthroplasty devices have average follow-ups of 24 months,(11, 12, 15, 16) which is a relatively short duration to adequately evaluate the benefits or disadvantages of arthroplasty. (11) National registries are used to report information on implant survival. They are not intended to estimate device benefits but function as a post-market safety surveillance mechanism. Hence, they are considered only in the systematic appraisal of harms.(1)

9. Acknowledgements

We wish to thank Setor Kunutsor, Andrew Beswick and Richard Donovan who undertook systematic review of the published literature to underpin this document.

10. References

1. Nieuwenhuijse MJ, Nelissen RGHH, Schoones JW, Sedrakyan A. **Appraisal of evidence base for introduction of new implants in hip and knee replacement: a systematic review of five widely used device technologies. BMJ 2014; 349:g5133.**

2. Sedrakyan A, Normand S-LT, Dabic S, Jacobs S, Graves S, Marinac-Dabic D. **Comparative assessment of implantable hip devices with different bearing surfaces: systematic appraisal of evidence. BMJ 2011; 343(nov29 2):d7434.**

3. Hulstaert F, Neyt M, Vinck I, Stordeur S, Huić M, Sauerland S et al. **Pre-market clinical evaluations of innovative high-risk medical devices in Europe. Int J Technol Assess Health Care 2012; 28(3):278–84.**

4. Goldberg VM, Buckwalter J, Halpin M, Jiranek W, Mihalko W, Pinzur M et al. **Recommendations of the OARSI FDA Osteoarthritis Devices Working Group. Osteoarthritis and Cartilage 2011; 19(5):509–14.**

5. Samuel AM, Rathi VK, Grauer JN, Ross JS. **How do Orthopaedic Devices Change After Their Initial FDA Premarket Approval? Clin Orthop Relat Res 2016; 474(4):1053–68.**

6. Kramer DB, Cutlip DE. **Regulatory science: Trust and transparency in clinical trials of medical devices. Nat Rev Cardiol 2015; 12(9):503–4.**

7. Chang L, Dhruva SS, Chu J, Bero LA, Redberg RF. **Selective reporting in trials of high risk cardiovascular devices: cross sectional comparison between premarket approval summaries and published reports. BMJ 2015; 350(may27 7):h2613-h2613.**

8. Weszl M, Rencz F, Brodszky V. **Is the trend of increasing use of patient-reported outcome measures in medical device studies the sign of shift towards value-based purchasing in Europe? Eur J Health Econ 2019; 20(Suppl 1):133–40.**

9. Sim I. **Trial registration for public trust: making the case for medical devices. J GEN INTERN MED 2008; 23 Suppl 1(S1):64–8.**

10. Cohen D, Billingsley M. **Europeans are left to their own devices. BMJ 2011; 342(may13 2):d2748.**

11. Delamarter R, Zigler JE, Balderston RA, Cammisa FP, Goldstein JA, Spivak JM. **Prospective, randomized, multicenter Food and Drug Administration investigational device exemption study of the ProDisc-L total disc replacement compared with circumferential arthrodesis for the treatment of two-level lumbar degenerative disc disease: results at twenty-four months. J Bone Joint Surg Am 2011; 93(8):705–15.**

12. Zigler J, Delamarter R, Spivak JM, Linovitz RJ, Danielson GO, Haider TT et al. **Results of the prospective, randomized, multicenter Food and Drug Administration investigational device exemption study of the ProDisc-L total disc replacement versus circumferential fusion for the treatment of 1-level degenerative disc disease. Spine 2007; 32(11):1155-62; discussion 1163.**

13. Yang BW, Iorio ML, Day CS. **Orthopaedic Device Approval Through the Premarket Approval Process: A Financial Feasibility Analysis for a Single Center. J Bone Joint Surg Am 2017; 99(6):e26.**

14. Kynaston-Pearson F, Ashmore AM, Malak TT, Rombach I, Taylor A, Beard D et al. **Primary hip replacement prostheses and their evidence base: systematic review of literature. BMJ 2013; 347(dec19 2):f6956.**

15. Blumenthal S, McAfee PC, Guyer RD, Hochschuler SH, Geisler FH, Holt RT et al. **A prospective, randomized, multicenter Food and Drug Administration investigational device exemptions study of lumbar total disc replacement with the CHARITE artificial disc versus lumbar fusion: part I: evaluation of clinical outcomes. Spine 2005; 30(14):1565-75; discussion E387-91.**

16. Guyer RD, Geisler FH, Blumenthal SL, McAfee PC, Mullin BB. **Effect of age on clinical and radiographic outcomes and adverse events following 1-level lumbar arthroplasty after a minimum 2-year follow-up. SPI 2008; 8(2):101–7.**

17. Institute CH. **Competitiveness and regulation: the FDA and the future of America's biomedical industry; 2011.**

18. Union OJotE. REGULATION (EU) 2017/745 OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL on medical devices, amending Directive 2001/83/EC, Regulation (EC) No 178/2002 and Regulation (EC) No 1223/2009 and repealing Council Directives 90/385/EEC and 93/42/EEC. Official Journal of the European Union 2017.

19. Atrey A, Heylen S, Gosling O, Porteous MJL, Haddad FS. **The manufacture of generic replicas of implants for arthroplasty of the hip and knee: is it regulated and will it save money? The Bone & Joint Journal 2016; 98-B(7):892–900.**

20. Malchau H, Bragdon CR, Muratoglu OK. **The stepwise introduction of innovation into orthopedic surgery: the next level of dilemmas. The Journal of Arthroplasty 2011; 26(6):825–31.**

21. Lombardi AV, Berend KR, Seng BE, Clarke IC, Adams JB. **Delta ceramic-on-alumina ceramic articulation in primary THA: prospective, randomized FDA-IDE study and retrieval analysis. Clin Orthop Relat Res 2010; 468(2):367–74.**

22. Sauerland S, Fujita-Rohwerder N, Zens Y, Molnar S. **Premarket evaluation of medical devices: a cross-sectional analysis of clinical studies submitted to a German ethics committee. BMJ Open 2019; 9(2):e027041.**

23. Freemantle N. **Commentary: Evaluating and regulating device therapy. BMJ 2011; 342(may13 2):d2839.**

24. Hasan S, Marang-van de Mheen PJ, Kaptein BL, Nelissen RGHH, Pijls BG. **RSA-tested TKA Implants on Average Have Lower Mean 10-year Revision Rates Than Non-RSA-tested Designs. Clin Orthop Relat Res 2020; 478(6):1232–41.**

25. Falk V, Friede T. Noninferiority trials: **What's clinically (ir)relevant? The Journal of Thoracic and Cardiovascular Surgery 2021; 161(6):2119–23.**

26. Seiffert M, Walther T, Hamm C, Falk V, Frey N, Thiele H et al. **The DEDICATE Trial: An independent all-comers trial of transcatheter aortic valve implantation vs. surgical aortic valve replacement in patients at low to intermediate operative risk is recruiting patients. European Heart Journal 2019; 40(4):331–3.**

27. Deere KC, Whitehouse MR, Porter M, Blom AW, Sayers A. **Assessing the non-inferiority of prosthesis constructs used in total and unicondylar knee replacements using data from the National Joint Registry of England, Wales, Northern Ireland and the Isle of Man: a benchmarking study. BMJ Open 2019; 9(4):e026736.**

28. U.S. Food and Drug Administration. **Federal Food, Drug, and Cosmetic Act. Federal Food, Drug, and Cosmetic Act. Available from: URL: http://uscode.house.gov/view.xhtml?path=/prelim@title21/chapter9&edition=prelim.**

RESEARCH TOPIC 25

Requirements to Study Design of Pre-CE Studies / Safety Studies

Authors

Ashley Blom[1,2], Dario Bergadano[3], Ina Wüstefeld[4], Justin Cobb[5], Faris Haddad[6], Marcus Jäger[7,8], Hassan Achakri[9], Matthias Fink[10], Arndt-Peter Schulz[11,12]

[1]National Institute for Health Research Bristol Biomedical Research Centre, University Hospitals Bristol and Weston NHS Foundation Trust and the University of Bristol, Bristol, UK

[2]Translational Health Sciences, Bristol Medical School, Musculoskeletal Research Unit, University of Bristol, Learning & Research Building, Southmead Hospital, Bristol, UK

[3]Medacta International SA, Castel San Pietro, Switzerland

[4]Aesculap AG Medical Scientific Affairs, Tuttlingen, Germany

[5]Imperial College London, MSk Lab, Sir Michael Uren Biomedical Engineering Research Hub, Imperial College London, White City Campus, London, UK

[6]University College London Hospitals, The Princess Grace Hospital, and The NIHR Biomedical Research Centre at UCLH, London, UK

[7]Department of Orthopaedics, Trauma & Reconstructive Surgery, St. Marien Hospital Mülheim an der Ruhr, Mülheim, Germany

[8]Orthopaedics and Trauma Surgery, University of Duisburg-Essen, Essen, Germany

[9]Zimmer Biomet, Clinical Affairs, Zug - Switzerland

[10]TÜV Süd group, Munich, Germany

[11]Department of Trauma Surgery, Orthopedic and Sports Traumatology, BG Klinikum Hamburg, Hamburg, Germany

[12]Fraunhofer Research Institution for Individualized and Cell-Based Medical Engineering, Lübeck, Germany

1. Abstract

Question: Pre-CE studies / safety studies

Which requirements to the study design of pre-CE studies/safety studies exist?

1. Which study set up is suitable to investigate relevant parameters and provide reliable results (required case number, necessary documentation, imaging modalities etc.)?
2. Which requirements must be demanded from users and centres to be able to perform such studies in a reasonable way and with a high qualitative standard, e.g. should protocols require previous solid clinical research from centres/surgeons involved?
3. How does the learning curve as well as the choice of study centres/surgeons influence the result of a tightly planned study?

Summary / Recommendation:

1. Consider use of the IDEAL framework for novel devices
   - safety study – small numbers, single centre, including migration analysis
   - multi-centre single arm study to confirm non-inferiority with comparator devices or procedures based upon clinical outcomes. This is essential to confirm that the results are generalisable.
2. Successful clinical trials require evidence from more than one centre, so input should be encouraged from all surgeons and centres. However lack of follow-up data, or poorly completed records may invalidate a trial, so new investigators and centres will need training and support by an appropriately experienced innovation team (surgeons, clinicians and researchers with specific technical expertise). They will benefit from access to a wider level of expertise including systematic reviewer, qualitative researcher, trialist/ statistician, health economist, and ethics support.
3. The influence of learning curves on the results of a tightly planned study may not be consistent. Pre-CE studies often aim to integrate new implants that are based on existing technologies, therefore the learning curve may not be as pronounced as, for example, learning a new approach for the first time. Therefore, the effect of the learning curve for experienced arthroplasty surgeons is expected to be limited.

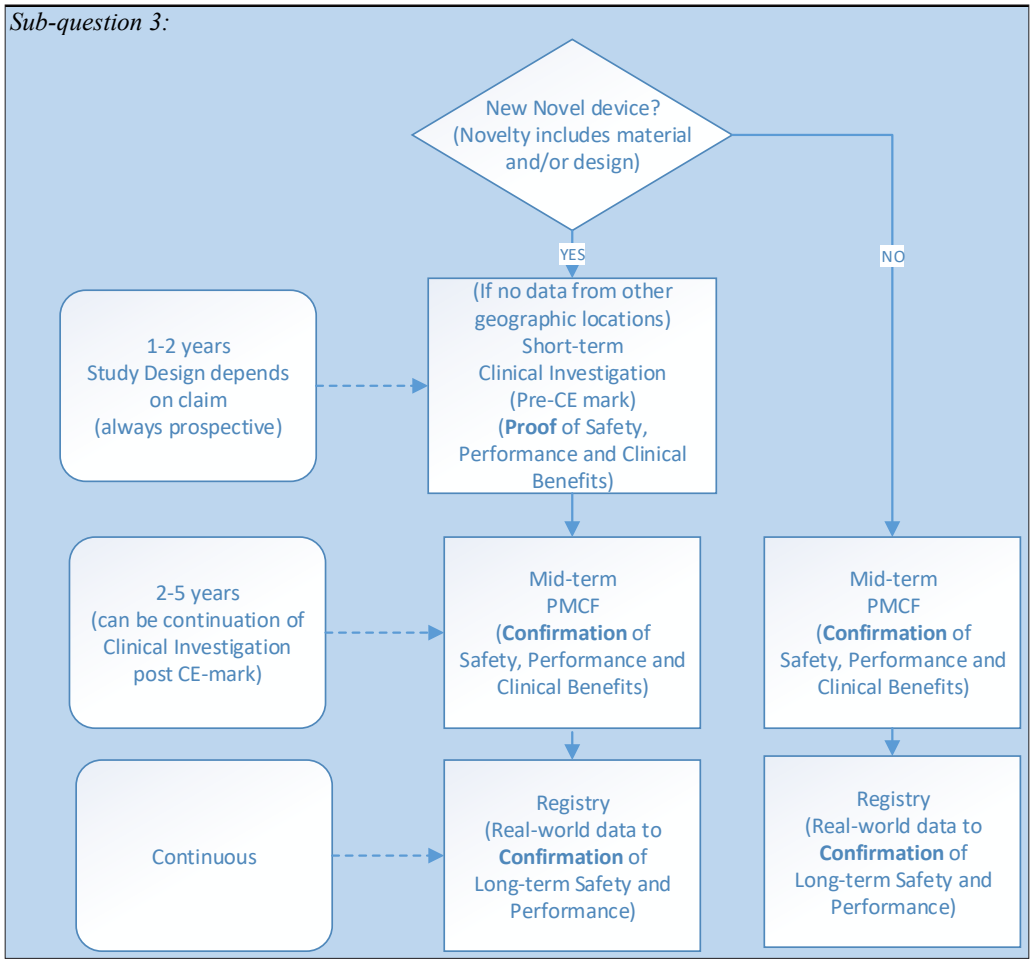
2. Level of Evidence

Strong

3. Consensus Delegate Vote

92% - super majority, strong consensus (92% agree / 3% disagree / 5% abstain)

4. Graphical Abstract



5. Search Strategy

Data Sources: MEDLINE; clinicaltrials.gov; systematic reviews of relevant studies in arthroplasty that have worked with the FDA to identify summaries and effectiveness of relevant premarket application trials;(1, 2) and reference lists of relevant studies.

6. Context and Background

Please refer to Consensus Statement 24/Pre-CE studies / safety studies.

7. Rationale

1. Traditionally, joint arthroplasty success was judged by the length of time the implant remained in situ.(3) NICE criticised this approach in 2020, pointing out that this metric is 'flawed' as it fails to recognise the differing thresholds to revision, not recognising the difference between minor and major reoperations(4). In a national registry based study, poorly functioning partial knee replacements are 6 times more likely to be revised than total knee replacements(5). The failure to recognise this difference in thresholds for revision leads to an excessively conservative viewpoint, with more easily revisable procedures appearing less successful(6). The key issues for surgeons and prosthesis manufacturers were the design and fixation of a prosthesis that would last for at least ten years, without the need for revision surgery. However, with advances in surgical technique and prosthesis design, joint replacement now has good survivorship.(7, 8) While uncommon, complications have serious implications for patients and include dislocation, infection, peri-prosthetic fracture, thromboembolism and neurovascular damage. There is also a small mortality risk related to class of prosthesis. For example unicompartmental knee replacemet is associated with markedly lower mortality than TKR.(6, 9–11) Prevention of poor long-term patient outcomes are now recognised as a key emphasis of patient care and research.(12)

Stages of innovation for a new or modified prosthesis follow the pathway shown in Figure 1 with intervention development described according to PI(C)OS. This is based on the IDEAL model described by McCulloch and colleagues and Hirst and colleagues.(13, 14) The IDEAL Collaboration advocate a "total product life cycle approach, whereby market access is granted in stages as clinical evidence accumulates."

Please note that although the IDEAL Framework states that long-term surveillance does not require ethical approval, since GDPR was introduced, most registries do obtain patient consent and ethical approval.

IDEAL in relation to orthopaedic surgery

A search of MEDLINE, Embase and Web of Science on 26th April 2021 noted that the IDEAL framework has not been used formally in the development and evaluation of arthroplasty and orthopaedic procedures. In Table 1, the requirements at each stage of the IDEAL pathway as required in studies evaluating new methods in arthroplasty and orthopaedic surgery are summarised. In Table 2, specific examples for each stage of surgical innovation are described in relation to arthroplasty and orthopaedic surgery.

Table 1. IDEAL stages of surgical innovation

	Method	Patient number. Surgeon/ innovator number	Documentation	Example measures
Pre-clinical	Simulator, cadaver, animal, modelling, cost-effectiveness	Patients do not receive procedure. Few innovators	Conducted to general standards of research integrity. Replace animal models where possible. Explicit reporting	Physical changes, mechanical properties, technical feasibility, level of difficulty, safety
Idea - proof of concept	First in human studies – planned or emergency	Single or a few highly selected patients. Very few innovator surgeons	Conduct according to local hospital and research ethics (IRB). Structured case reports (SCARE guidelines). Video recording and sharing	Details of need, and inception and evolution of procedure. Full technical description. Core outcome set to include short term physiological and clinical measures, adverse events
Proceed to development dependent on proof of concept, technical achievement, safety, and potential efficacy				
Development	Prospective development cohort. Iterative modification to achieve final, stable version.	Small number of patients. Few surgeons.	Institutional research ethics review.	Technical description of stable procedure with modifications explained. Reasons for ineligibility. Short term clinical, technical, safety outcomes.
Exploration - learning	Feasibility RCT. Research database. Propensity matching of observational data. Consider learning curves.	Larger less selective population (ideally 100+ patients), many surgeons, early adopters	Ethical approval. Patient consent. Consort for pilot/ feasibility studies. STROBE. IDEAL. Published protocol. Trial registered.	Safety, short term clinical outcomes. Trial feasibility and data for sample size calculation. Evolution and refinement of procedure, community learning. Confirmation of intervention, patient population, outcomes. Output reporting empirical comparison with other alternative care. Is equipoise feasible and randomisation acceptable to patients and surgeons? Consensus on the design and conduct of a future fully powered RCT.
Assessment	Fully powered RCTs. Comparison with usual care or sham.	Substantial number of patients and surgeons/ healthcare professionals at multiple centres	Trial registered. Well-defined outcome measures. Informed consent. Ethical approval. Appropriate guidelines	Clinical, PROM, health economic outcomes.
Long-term study - surveillance	Registry studies. Meta-analyses of RCTs and cohorts.	Clinical community	Consent at time of surgery	Rare events, long-term outcomes

Table 2. Examples of orthopaedic intervention development in the context of the IDEAL framework

	Method	Patient number. Surgeon/ innovator number	Documentation	Example measures
Pre-clinical				
Novel reverse hip replacement(15)	Biomechanical testing – extended fatigue test	6 hip stem assemblies, 50,000,000 cycles	Published article. Conflicts of interest described.	Taper corrosion Polyethylene wear "The results of this testing provide a firm foundation for further clinical investigation of the reverse total hip arthroplasty"
Idea- proof of concept				
Meniscal cartilage repair with undifferentiated autologous mesenchymal stem cells(16)	Single centre, prospective, open-label first-in-human safety study	5 patients recruited over 8 month period and followed up for 2 years	"All patients gave their informed consent and the study was carried out according to local ethical guidelines." Published article. Conflicts of interest described	Implant survival, clinical scores, MRI scans. "The preliminary clinical data are encouraging and suggest for the first time that repair of avascular meniscal tears is possible"
Development				
Treatment of hip prosthetic joint infection with custom made articulating spacer(17)	Consecutive group of patients from introduction of the technique. Single centre, experienced arthroplasty surgeons. Prospectively collected case series.	75 patients (76 hips) over 10.5 years followed up for a mean of 6.7 years (range 2.1, 12.1)	No ethics or consent described. Conflicts of interest described	Evolving technique and "learning curve". Eradication of infection, complications, unplanned revisions. "Longer term follow-up is necessary to fully evaluate this technique but early results are promising." Procedure developed from technique to deliver antibiotics through a prosthesis that preserves leg length and function between stages of a two-stage revision procedure. Value as a long-lasting alternative to second-stage reimplantation described
Exploration – learning				
Avoidance of tourniquet use in knee replacement(18)	Feasibility RCT. Single centre with 13 surgeons.	53 patients, 12 months follow up	NREC approval. Written informed consent	"A full trial is feasible, but using MRI as a primary outcome is unlikely to be appropriate or feasible. Suitable primary outcomes would be cognition measured using MoCA, pain and AEs, all of which warrant investigation in a large multicentre trial."
Assessment				
For displaced femoral neck fractures, the effect of a total hip arthroplasty as compared with hemiarthroplasty(19)	Fully powered RCT. 80 sites in 10 countries. Expertise-based trial design	1495 patients recruited over 8.5 years and followed up for 2 years	Full ethical approval and consent processes described in protocol.	Outcomes of secondary hip procedure, death, SAE, hip-related complications, HRQoL, function, general health. No difference in incidence of primary outcome at 2 years. Further follow up required
Long-term study – surveillance				
Failure rates of stemmed metal-on-metal hip replacements(20)	National Joint Registry.	402,051 hip replacements (31,171 stemmed metal on metal) performed at 447 units by 2578 surgeons	Patients consent to data collection at time of surgery.	Outcome of implant survival. "Metal-on-metal stemmed articulations give poor implant survival compared with other options and should not be implanted."

2. Input into the design and conduct is required from multidisciplinary researchers with appropriate expertise.
   - Innovators (surgeons, clinicians, researchers with specific technical expertise). Well-defined question or hypothesis. Procedure and protocol development.
   - Ethicist input. To advise on research design and conduct, and patient consent.
   - Systematic reviewer to contextualise intervention: scale of problem addressed, use of intervention in other conditions, effectiveness of alternatives.
   - Qualitative researcher and patient representatives. Acceptability of interventions assessed in qualitative studies with patients and surgeons. Also, opinions on equipoise and acceptability of randomisation.
   - Trialist/ statistician. Data analysis. Sample size calculations
   - Health economist

In a feasibility RCT, key considerations to be addressed by the multidisciplinary group(50) are:
   - Acceptability of the intervention to surgeon/ clinician and patients
   - Adherence to the intervention
   - Representative recruitment and engagement
   - Willingness of patients to be randomised
   - Willingness of clinicians to randomise patients
   - Choice of primary outcomes
   - Choice of a comparator
   - Follow-up rates, response rates to questionnaires, adherence/compliance rates
   - Data management and analysis
   - Practicality of intervention delivery in the proposed setting
   - Variation in use or delivery of the intervention in each setting

3. Adoption of new techniques and technologies follow a sigmoidal learning curve. Learning curves in the orthopaedic literature include learning a new surgical approach or the adoption of robot-assisted total knee replacement. Surgeons performing the direct anterior approach needed to perform this 50 times to achieve cumulative revision rates equivocal to those who had performed this >100 times. Performance typically improves with experience. Proxies for measuring the learning curve can be assessed by looking at surgical processes (operative time, blood loss, economy of movement) and patient outcomes (PROMs, complications, mortality).(21) Assessing learning curves of individual surgeons can be difficult, and they are rarely given formal consideration in health technology assessment.(22) The implementation of new techniques or technologies may increase the risk of complications or adverse events to the patient. There tends to be an early phase of learning and confidence-building, followed by a longer proficiency phase. Learning curves in arthroplasty can be a barrier to adoption of new techniques or technologies. New technologies often lack the medium/long-term outcome data to support their clinical use, which continues to be a challenge due to the time and costs associated with this, and the losses to prolonger follow-up over years. New and relatively untested implants may have unforeseen consequences (as seen in MoM hips).(23)

The literature remains mixed regarding the effect of learning curves. A large registry-based study in Finland found there seemed to be a learning curve which was implant-specific when the proxy for the learning curve was early revision.(24) Another study suggested there is very little additional learning curve when an experienced surgeon transitions from one hip resurfacing implant to a different but similar one.(25)

Pre-CE studies often aim to integrate new implants that are based on existing technologies, therefore the learning curve may not be as pronounced as, for example, learning a new approach for the first time. Therefore, the effect of the learning curve for experienced arthroplasty surgeons is expected to be limited.

Novel technologies such as virtual reality training suites have been shown to reduce the learning curve(26), so may help in the roll out from the first centre to others. Typically such learning tools require a Delphi type consensus, so inevitably come after a novel device or technique has been adopted for some time.

8. Acknowledgements

We wish to thank Setor Kunutsor, Andrew Beswick and Richard Donovan who undertook systematic review of the published literature to underpin this document.

9. References

1. Nieuwenhuijse MJ, Nelissen RGHH, Schoones JW, Sedrakyan A. **Appraisal of evidence base for introduction of new implants in hip and knee replacement: a systematic review of five widely used device technologies. BMJ 2014; 349:g5133.**

2. Sedrakyan A, Normand S-LT, Dabic S, Jacobs S, Graves S, Marinac-Dabic D. **Comparative assessment of implantable hip devices with different bearing surfaces: systematic appraisal of evidence. BMJ 2011; 343(nov29 2):d7434.**

3. Aamodt A, Nordsletten L, Havelin LI, Indrekvam K, Utvåg SE, Hviding K. **Documentation of hip prostheses used in Norway: a critical review of the literature from 1996--2000. Acta Orthopaedica Scandinavica 2004; 75(6):663–76.**

4. NICE. Joint replacement (primary): hip, knee and shoulder NICE guideline [NG157]; 2020. Available from: URL: https://www.nice.org.uk/guidance/ng157.

5. Goodfellow JW, O'Connor JJ, Murray DW. **A critique of revision rate as an outcome measure: re-interpretation of knee joint registry data. The Journal of Bone and Joint Surgery. British volume 2010; 92(12):1628–31.**

6. Liddle AD, Judge A, Pandit H, Murray DW. **Adverse outcomes after total and unicompartmental knee replacement in 101 330 matched patients: a study of data from the National Joint Registry for England and Wales. The Lancet 2014; 384(9952):1437–45.**

7. Evans JT, Walker RW, Evans JP, Blom AW, Sayers A, Whitehouse MR. **How long does a knee replacement last? A systematic review and meta-analysis of case series and national registry reports with more than 15 years of follow-up. The Lancet 2019; 393(10172):655–63.**

8. Evans JT, Evans JP, Walker RW, Blom AW, Whitehouse MR, Sayers A. **How long does a hip replacement last? A systematic review and meta-analysis of case series and national registry reports with more than 15 years of follow-up. The Lancet 2019; 393(10172):647–54.**

9. Berstock JR, Beswick AD, López-López JA, Whitehouse MR, Blom AW. **Mortality After Total Knee Arthroplasty: A Systematic Review of Incidence, Temporal Trends, and Risk Factors. J Bone Joint Surg Am 2018; 100(12):1064–70**.

10. Berstock JR, Beswick AD, Lenguerrand E, Whitehouse MR, Blom AW. **Mortality after total hip replacement surgery: A systematic review. Bone & Joint Research 2014; 3(6):175–82.**

11. Kendal AR, Prieto-Alhambra D, Arden NK, Carr A, Judge A. **Mortality rates at 10 years after metal-on-metal hip resurfacing compared with total hip replacement in England: retrospective cohort analysis of hospital episode statistics. BMJ 2013; 347(nov27 1):f6549.**

12. Wylde V, Blom AW. **The failure of survivorship. The Journal of Bone and Joint Surgery. British volume 2011; 93(5):569–70.**

13. McCulloch P, Altman DG, Campbell WB, Flum DR, Glasziou P, Marshall JC et al. **No surgical innovation without evaluation: the IDEAL recommendations. The Lancet 2009; 374(9695):1105–12.**

14. Hirst A, Philippou Y, Blazeby J, Campbell B, Campbell M, Feinberg J et al. **No Surgical Innovation Without Evaluation: Evolution and Further Development of the IDEAL Framework and Recommendations. Annals of Surgery 2019; 269(2):211–20.**

15. Braddon L, Termanini Z, MacDonald S, Parvizi J, Lieberman J, Frankel V et al. **Corrosion and Tribology of Materials Used in a Novel Reverse Hip Replacement. Materials (Basel) 2017; 10(7):751.**

16. Whitehouse MR, Howells NR, Parry MC, Austin E, Kafienah W, Brady K et al. **Repair of Torn Avascular Meniscal Cartilage Using Undifferentiated Autologous Mesenchymal Stem Cells: From In Vitro Optimization to a First-in-Human Study. STEM CELLS Translational Medicine 2017; 6(4):1237–48.**

17. Tsung JD, Rohrsheim JAL, Whitehouse SL, Wilson MJ, Howell JR. **Management of periprosthetic joint infection after total hip arthroplasty using a custom made articulating spacer (CUMARS); the Exeter experience. The Journal of Arthroplasty 2014; 29(9):1813–8.**

18. Wall PDH, Ahmed I, Edwin C, Farhan-Alanie MM, Parsons H, Price AJ et al. **Tourniquet use in total knee replacement surgery: a feasibility study and pilot randomised controlled trial (SAFE-TKR study). BMJ Open 2021; 11(1):e043564.**

19. Bhandari M, Einhorn TA, Guyatt G, Schemitsch EH, Zura RD, Sprague S et al. **Total Hip Arthroplasty or Hemiarthroplasty for Hip Fracture. N Engl J Med 2019; 381(23):2199–208.**

20. Smith AJ, Dieppe P, Vernon K, Porter M, Blom AW. **Failure rates of stemmed metal-on-metal hip replacements: analysis of data from the National Joint Registry of England and Wales. The Lancet 2012; 379(9822):1199–204.**

21. Valsamis EM, Sukeik M. **Evaluating learning and change in orthopaedics: What is the evidence-base? WJO 2019; 10(11):378–86.**

22. Ramsay CR, Grant AM, Wallace SA, Garthwaite PH, Monk AF, Russell IT. **Assessment of the learning curve in health technologies. A systematic review. Int J Technol Assess Health Care 2000; 16(4):1095–108.**

23. Sarpong NO, Herndon CL, Held MB, Neuwirth AL, Hickernell TR, Geller JA et al. **What Is the Learning Curve for New Technologies in Total Joint Arthroplasty? A Review. Curr Rev Musculoskelet Med 2020; 13(6):675–9.**

24. Peltola M, Malmivaara A, Paavola M. **Learning curve for new technology?: a nationwide register-based study of 46,363 total knee arthroplasties. J Bone Joint Surg Am 2013; 95(23):2097–103.**

25. Johnson AJ, Costa CR, Naziri Q, Mont MA. **Is there a new learning curve with transition to a new resurfacing system? Bull NYU Hosp Jt Dis 2011; 69 Suppl 1:S16-9.**

26. Logishetty K, Gofton WT, Rudran B, Beaulé PE, Cobb JP. **Fully Immersive Virtual Reality for Total Hip Arthroplasty: Objective Measurement of Skills and Transfer of Visuospatial Performance After a Competency-Based Simulation Curriculum. J Bone Joint Surg Am 2020; 102(6):e27.**

RESEARCH TOPIC

PERIOPERATIVE AND SHORT-TERM POSTOPERATIVE (SERIOUS) ADVERSE EVENTS

RESEARCH TOPIC 26

Approach for the Pre-Clinical Investigation of Adverse Events or Complications Related to the Clinical Application of Total Joint Arthroplasty Devices and their Implantation Procedure

Authors

Francesco Siccardi[1], Sabine Rusch[2], Søren Overgaard[3,4,5,6], Alexander Giurea[7], Thomas M. Grupp[8,9], Arndt-Peter Schulz[10,11]

[1]Medacta International SA, Castel San Pietro, Switzerland
[2]Aesculap AG Medical Scientific Affairs, Tuttlingen, Germany
[3]Department of Clinical Medicine, Copenhagen University, Copenhagen, Denmark
[4]Department of Orthopaedic Surgery and Traumatology, Bispebjerg and Frederiksberg Hospital, University of Copenhagen, Copenhagen, Denmark
[5]Orthopaedic Research Unit, Department of Orthopaedic Surgery and Traumatology, Odense University Hospital, Odense, Denmark
[6]Department of Clinical Research, University of Southern Denmark, Odense, Denmark
[7]Department of Orthopedics and Trauma Surgery, Medical University of Vienna, Vienna, Austria
[8]Aesculap AG Research & Development, Tuttlingen, Germany
[9]Ludwig Maximilians University Munich, Department of Orthopaedic and Trauma Surgery, Musculoskeletal University Center Munich (MUM), Campus Grosshadern, Munich, Germany
[10]Department of Trauma Surgery, Orthopedic and Sports Traumatology, BG Klinikum Hamburg, Hamburg, Germany
[11]Fraunhofer Research Institution for Individualized and Cell-Based Medical Engineering, Lübeck, Germany

1. Abstract

Question: Perioperative and short-term postoperative (serious) adverse events

Is it possible to prove pre-clinically that the implantation procedure and the implants do not induce unreasonably high rates of adverse events or complications (directly implant-related, e.g. inter-operative periprosthetic fractures, substantial bleeding, substantial migration, and generally related to the surgery / procedure, e.g. fast track, modified surgical approaches)?

Summary / Recommendation:

The current approach to address the research question relies on the Implant and Procedure Risk Management Process (RMP) following ISO 14971. To address the RMP, several strategies, depending on the level of innovation introduced by the new implant/procedure under development, can be applied in order to prove pre-clinically that the implantation will not induce unreasonably high short term rates of adverse events or complications. Strategies might include and combine

- Clinical feedback from proven substantially equivalent devices and procedures,
- Standardized and non-Standardized testing,
- In silico and cadaveric testing,
- Guidelines and Expert Opinions.

After having listed all possible risk, each risk should be carefully described in the Implant and Procedure Risk Analysis together with the relevant documentation to justify the risk/benefit ratio.

Based on the modern Total Hip and Total Knee performance and on the Literature Systematic Review we can confirm the current approach can be considered an effective way to assess pre-clinically the Implant/Procedure performance and avoid unreasonably high rates of Implants and Procedure short term adverse events and complications when appropriate substantially equivalent device/procedure are available and the Risk Management Process is duly implemented.

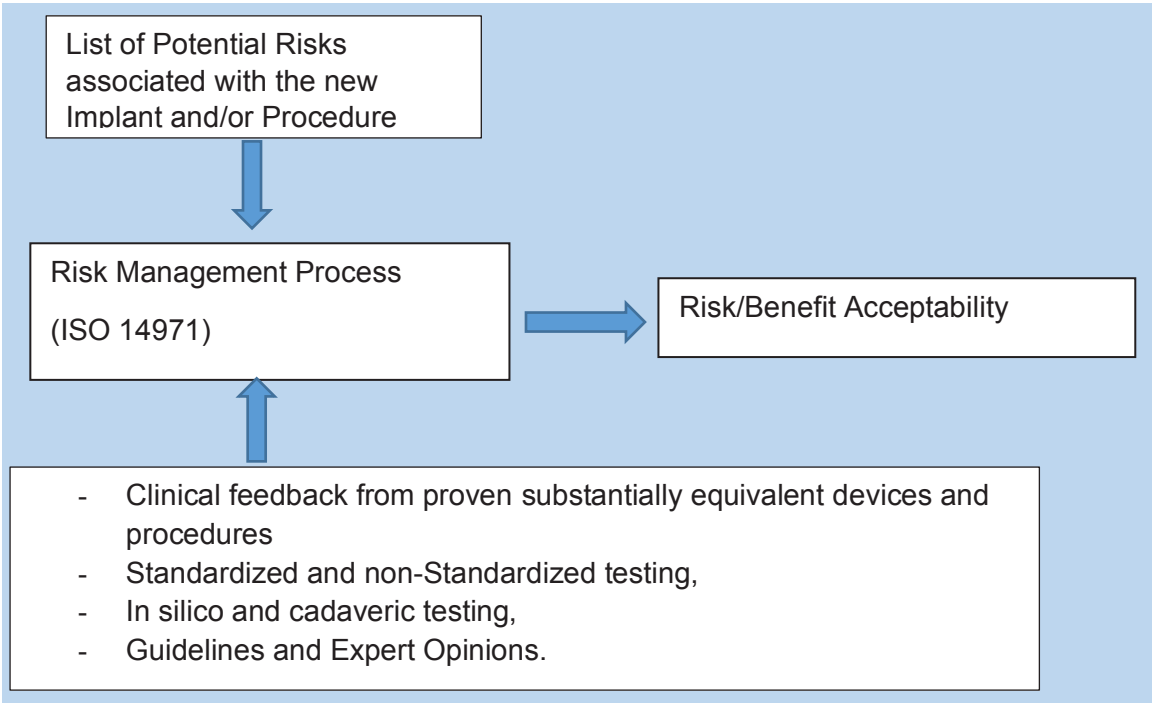
2. Level of Evidence

High

3. Consensus Delegate Vote

95% – unanimous, strongest consensus (95% agree / 0% disagree / 5% abstain)

4. Graphical Abstract



5. Search Strategy

A systematic literature search was conducted in March 2021 to identify literature with the potential to contribute to the answer of the above mentioned research question.

The electronic databases PubMed and EMBASE were explored using the combination of the following search-terms combined with the Boolean operator "AND":

- TKA/THA, knee/hip prosthesis, knee/hip replacement
- intra-/perioperative/short-term complications, adverse (device) effect, (serious) adverse event, complication rate
- cadaver/cadaveric/in vitro study/workshop/model, simulation, simulator, preclinical, biomechanical study, predicate device, in silico

Search words were adapted to the specific search terms of the database and were searched including different spelling and synonyms. Article language was restricted to English, French, German or Italian language. Because of the aim of the process only studies from 2001 to 2021 were included to provide insight into current practice and state of the art procedures. Publications citing 'periprosthetic joint infection' or 'prosthesis infection' as major topics were excluded as those papers mainly reported in vitro studies regarding different bacterial strains, antibiotic therapy etc.

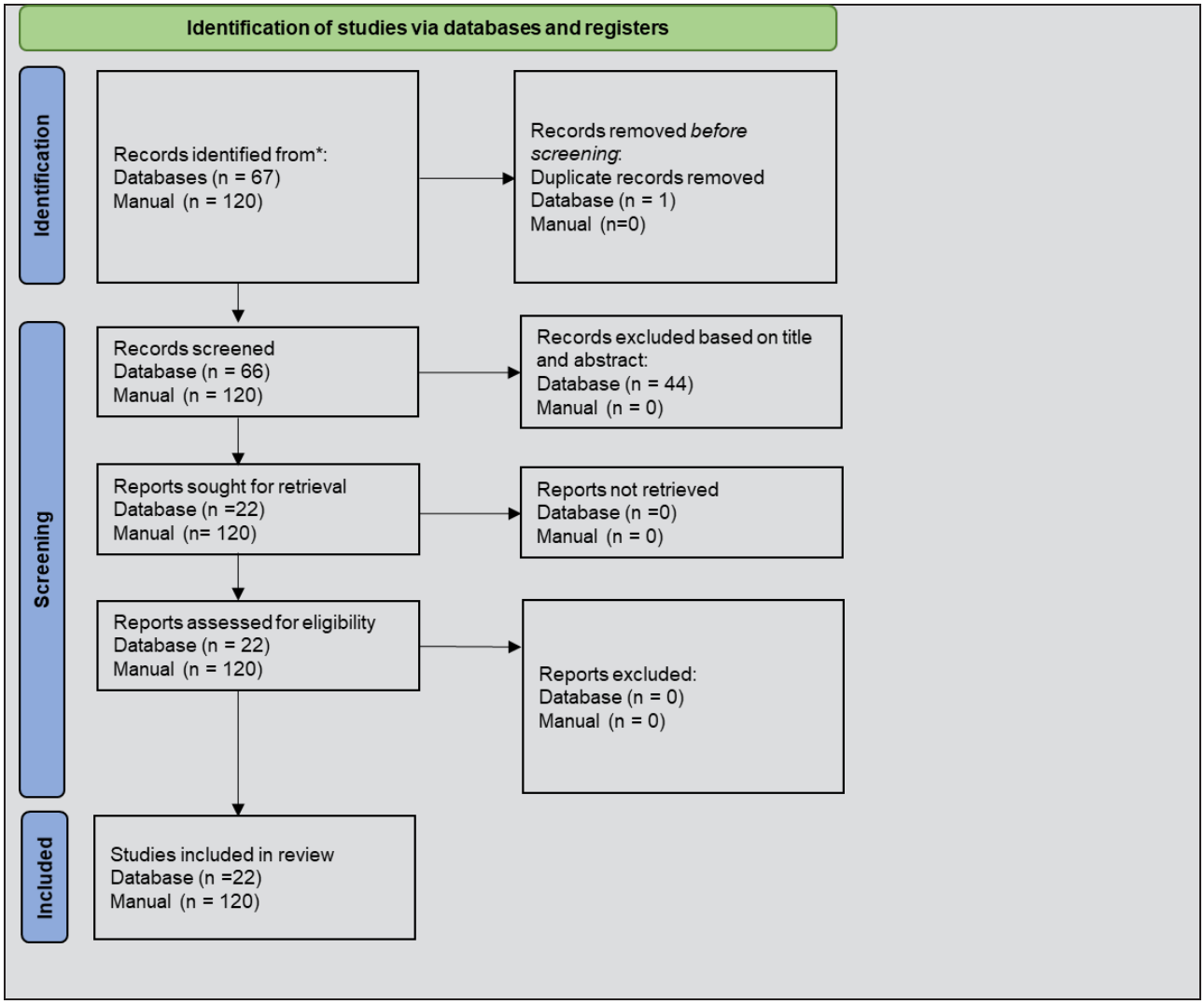
The title and abstract of all articles were screened to select articles contributing to the topic.

Potentially eligible articles were obtained for detailed assessment and were in detail screened for applicability. The selection was not further restricted regarding basic study design. Duplicate articles were removed.

Abstracts from scientific meetings were not included in the current review as sufficient data cannot be extracted.

67 publications met the search criteria. After the selection process, 22 articles were included in the review.

Manual search and inclusion of all relevant Medical Device Safety Standard contributed with additional 120 references.



6. Rationale

To design new Implants and Procedures that do not induce unreasonably high rates of adverse events or complications is a minimum goal of any new Implant or Procedure development.

To reliably establish in pre-clinical assessment that a (new) implant and the implant specific instruments are not associated with unreasonably high rates of complications it is initially necessary to define the complications to be taken into consideration.

Definition of short-term complications:

For the evaluation the primary focus are product-related complications. Procedure-related complications which are not primarily dependent on the implant but on the specific procedure or approach should also be considered. Complications or readmissions generally related to TJA surgery (e.g. general postsurgical pain, sepsis, gastrointestinal/genitourinary complications) should not be in the focus. The definition of early complications depends on the selected timeframe. In accordance with the findings of the literature, complications up to 6 months will be defined as short-term (see also EFORT European Consensus Statement 27).

We consider it necessary to establish a list of complications. Healy et al. developed standardised lists of complications associated with TKA in 2013 and with THA in 2016. TKA and THA Complications Workgroups proposed lists of complications based on orthopaedic literature to be validated by expert surveys (1,2). Complications were then further stratified for severity using a validated system based on a modification of the Sink classification (3,4). 22 TKA and 19 THA complications were identified. However, there is no distinction regarding timeframe or origin of the complications.

In accordance with this list we consider the following complications relevant for both TKA and THA in the short-term and related to the product or the procedure (1,2):

- Neurologic deficit (i.e. Postoperative nerval deficit (sensory or motor) related to the index TKA/THA)
- Periprosthetic fracture: fracture of distal femur, proximal tibia, or patella in TKA, or proximal femur or the acetabulum in THA
- Early implant loosening: Implant loosening confirmed intraoperatively or identified radiographically as a change in implant position (i.e. acetabular cup) or progressive, radiolucent lines at the bone-cement or bone-implant interface

In addition specifically for TKA we consider:

- Collateral ligament injury (medial and lateral): Intraoperative or early postoperative injury requiring repair, reconstruction, a change in prosthetic constraint, or TKA protocol
- Instability: Symptomatic instability reported by the patient and confirmed by laxity on physical examination as defined by The Knee Society Knee Score
- Malalignment: Symptomatic malalignment reported by the patient and confirmed radiographically with angular deformity in the coronal plane >5° from the mechanical axis or the intended intraoperative alignment (kinematic, anatomic etc.)
- Stiffness: Limited ROM reported by the patient and demonstrated in physical examination with extension limited to 15° short of full extension or flexion <90° (n/a if preoperative arc of motion <75°)
- Patellofemoral dislocation: Dislocation of patella from the femoral trochlea
- Tibiofemoral dislocation: Dislocation of the tibiofemoral joint

In addition specifically for THA we consider :

- Dislocation/Instability: Dislocation of the femoral head out of the acetabulum or recurrent symptomatic subluxation of the hip
- Cup/liner dissociation: Dissociation of the cup/liner from the acetabular shell

From literature review no further complications could be identified to be included. Nevertheless, any pre-clinical evaluation must always expect and allow the detection of previously unknown or new complications which might be associated with a device. Within the pre-clinical manufacturer risk assessment it is the goal to identify such further device-specific complications and consequently include them in the clinical assessment.

When developing new Implants and Procedures the industry is required to follow a so-called Risk Management Process (RMP) as defined in the ISO 14971 (5). Specifically, ISO 14971 establishes a framework for risk analysis, evaluation, control and review, and also specifies a procedure for review and monitoring during production and post-production.

As the results of the Systematic Literature Review did not provide sufficient support, the following methodology relies on the orthopedic sector expertise of the working group and it describes the current industry approach to RMP. RMP requires to list all possible risks related to the new Implant or Procedure and classify them according to their frequency and severity to define the risk/benefit ratio for each of those risks. To assess pre-clinically the expected performance of a new Implant or Procedure and address the possible related risks identified in the RMP different strategies are utilized including:

- Clinical feedback from proven and substantially equivalent devices and procedures,
- Standardized testing
- Non-Standardized biomechanical, in-silico and cadaveric testing,
- Guidelines and Expert Opinions.

Clinical feedback from proven substantially equivalent devices and procedures:

Addressing potential area of improvements based on the performance of a substantially equivalent device or procedure is often the goal of incremental innovation in THA and TKA (6,7). The definition of a substantially equivalent device is of paramount importance and the methodology to make sure it is done correctly will not be discussed here. In fact without a proper substantially equivalent device it is often difficult to identify pre-clinically all potential risks and failure modes associated with the new devices even when very experienced people are involved (8).

Standardized testing:

In the last decades a significant number of standard tests have been developed to assess the safety of new THA (9-26) and TKA Implants (27-41). Those standardized tests cover both overall design process, material selection (42-62) and the key manufacturing processes including coating (63-76), cleaning, passivation and biocompatibility (77-97), labelling (98), packaging (99-109) and sterilization (110-118). More recently additional standard tests have been defined to assess the MRI compatibility of new Implants addressing the risk associated to heat generation around Implants while performing an MRI (119-121).

Non-Standardized biomechanical, in-silico and cadaveric testing:

Although Standard tests are covering most of the design and manufacturing aspects of new Implants and Procedures, it is often necessary to develop non-standardized biomechanical, in silico or cadaveric testing to look at specific risk/benefit aspects of new devices. Specifically, to peri-op and short-term complications the testing of periprosthetic fracture related risks is often covered by those type of tests particularly when comparing different design options or minor modification versus a predicate device (122-127). Another important peri-op and short term complication addressed by non-standardized in-silico and cadaveric testing is instability, dislocation or pain caused by primarily by Implant malpositioning both in THA (128-130) and TKA (131-136). The correct Implant positioning and the overall safety of the Procedure is often tested on cadaver with the relevant contribution and feedback provided by expert surgeons (137-140)

Guidelines and Expert Opinions:

Strong collaboration between Industry and experts (biomechanical engineers, surgeons, ...) is often required to address a specific risk/benefit ratio during the development of a new Implant or Procedure (137-140). Especially in non-standardized testing areas this contribution is paramount to create non-standardized testing protocols and interpret the related results correctly.

Conclusion:

The current approach can be considered an effective way to assess pre-clinically the Implant/Procedure performance and avoid unreasonably high rates of Implants and Procedure short term adverse events and complications when appropriate substantially equivalent device/procedure are available and the above RMP methodology is duly implemented.

This conclusion is confirmed by the latest AOANJRR and UKNJR registry results (141,142) and most relevant literature on modern Total Hip and Total Knee short term performance (1,2,3,4).

7. References

1. Healy, William L.; Della Valle, Craig J.; Iorio, Richard; Berend, Keith R.; Cushner, Fred D.; Dalury, David F.; Lonner, Jess H. (2013): **Complications of total knee arthroplasty: standardized list and definitions of the Knee Society. In: Clinical Orthopaedics and Related Research 471 (1), S. 215–220. DOI: 10.1007/s11999-012-2489-y.**

2. Healy, William L.; Iorio, Richard; Clair, Andrew J.; Pellegrini, Vincent D.; Della Valle, Craig J.; Berend, Keith R. (2016): **Complications of Total Hip Arthroplasty: Standardized List, Definitions, and Stratification Developed by The Hip Society. In: Clinical Orthopaedics and Related Research 474 (2), S. 357–364. DOI: 10.1007/s11999-015-4341-7.**

3. Iorio, Richard; Della Valle, Craig J.; Healy, William L.; Berend, Keith R.; Cushner, Fred D.; Dalury, David F.; Lonner, Jess H. (2014): **Stratification of standardized TKA complications and adverse events: a brief communication. In: Clinical Orthopaedics and Related Research 472 (1), S. 194–205. DOI: 10.1007/s11999-013-2980-0.**

4. Sink, Ernest L.; Leunig, Michael; Zaltz, Ira; Gilbert, Jennifer Claire; Clohisy, John (2012): Reliability of a complication classification system for orthopaedic surgery. In: Clinical Orthopaedics and Related Research 470 (8), S. 2220–2226. DOI: 10.1007/s11999-012-2343-2.

5. ISO 14971, **"Application of risk management to medical devices"**

6. Latypova A., Taghizadeh E., Becce F., Büchler P., Jolles B.M., Pioletti D.P., Terrier A. **Patellar bone strain after total knee arthroplasty is correlated with bone mineral density and body mass index .Medical Engineering and Physics (2019) 68 (17-24).**

7. Willmann G. **Investigation of retrieved femoral heads and cups, Untersuchung von explantierten kugelköpfen und pfannen von hüftendoprothesen.Biomedizinische Technik (2001) 46:12 (343-350).**

8. Van Susante J.L.C., Verdonschot N., Bom L.P.A., Tomaszewski P., Campbell P., Ebramzadeh E., Schreurs B.W. **Lessons learnt from early failure of a patient trial with a polymer-on-polymer resurfacing hip arthroplasty. Acta Orthopaedica (2018) 89:1 (59-65).**

9. ISO 7206-1 **Partial and total hip joint prostheses – Part 1: Classification and designation of dimension.**

10. ISO 7206-2 **Partial and total hip joint prostheses – Part 2: Articulating surfaces made of metallic, ceramic and plastics materials.**

11. ISO 7206-4 **Partial and total hip joint prostheses – Part 4: Determination of endurance properties of stemmed femoral component.**

12. ISO 7206-6 **Partial and total hip joint prostheses – Part 6: Determination of endurance properties of head and neck region of stemmed femoral component**

13. ISO 7206-10 **Partial and total hip-joint prostheses -– Part 10: Determination of resistance to static load of modular femoral heads**

14. ISO 14242-1 **Wear of total hip-joint prostheses - Part 1: Loading and displacement parameters for wear-testing machines and corresponding environmental conditions for test**

15. ISO 14242-2 **Wear of total hip-joint prostheses - Part 2: Methods of measurement**

16. ASTM F543 **Standard Specification and Test methods for Metallic Medical Bone screw**

17. ASTM F732 **Standard Test Method for Wear Testing of Polymeric Materials Used in Total Joint Prostheses.**

18. ASTM F897 **Standard Test Method measuring Fretting Corrosion of Osteosynthesis Plates and Screw**

19. ASTM F1814 **Standard guide for Evaluating Modular Hip and knee joint components**

20. ASTM F1820 **Standard Test Method for Determining the Axial Disassembly Force of a Modular Acetabular Device.**

21. ASTM F1875 **Standard Practice for Fretting Corrosion Testing of Modular Implant Interfaces: Hip Femoral Head-Bore and Cone Taper Interface**

22. ASTM F2009 **Standard test method for determining the axial disassembly force of taper connections of modular prostheses**

23. ASTM F2068 **Standard for Femoral Prostheses - Metallic Implants**

24. ASTM F2091 **Standard Specification for Acetabular Prostheses**

25. FDA Guidance **Hip Joint Metal/Polymer Constrained Cemented or Uncemented Prosthesis**

26. FDA Guidance **Non-clinical Information for Femoral Stem Prostheses**

27. ISO 7207-1 **Femoral and Tibial components for partial and total knee joint prostheses – Part 1: Classification, definitions and designation of dimension**

28. ISO 7207-2 **Implants for surgery – Components for partial and total knee joint prostheses – Part 2: Articulating surfaces made of metal, ceramic and plastics materials.**

29. ISO 14243-1 **Wear of total knee-joint prostheses - Part 1: Loading and displacement parameters for wear-testing machines with load control and corresponding environmental condition for test**

30. ISO 14243-2 **Wear of total knee-joint prostheses - Part 2: methods of measurement**

31. ISO 14243-3 **Wear of total knee-joint prostheses - Part 3: Loading and displacement parameters for wear-testing machines with displacement control and corresponding environmental conditions for test**

32. ASTM F1800 **Standard Test Method for Cyclic Fatigue Testing of Metal Tibial Tray Components of Total Knee Joint Replacements**

33. ASTM F2083 **Standard Specification for Total Knee Prosthesis**

34. ASTM F1814 **Standard Guide for Evaluating Modular Hip and Knee Joint Components**

35. ASTM F1223 **Standard Test Method for Determination of Total Knee Replacement Constraint**

36. ASTM F1672 **Standard Specification for Resurfacing Patellar Prosthesis**

37. ASTM F543 **Standard Specification and Test methods for Metallic Medical Bone screw**

38. ASTM F897 **Standard Test Method measuring Fretting Corrosion of Osteosynthesis Plates and Screw**

39. ASTM F732 **Standard Test Method for Wear Testing of Polymeric Materials Used in Total Joint Prostheses.**

40. FDA Guidance **Testing Non-Articulating, "Mechanically Locked", Modular Implant Components**

41. FDA Guidance **Knee Joint Patellofemorotibial and Femorotibial Metal/Polymer Porous Coated Uncemented Prostheses**

42. ISO 5832-1 **Implants for surgery – Metallic materials – Part 1: Wrought stainless steel**

43. ISO 5832-2 **Implants for Surgery - Metallic Materials - Part 2: Unalloyed Titanium**

44. ISO 5832-3 **Implants for surgery – Metallic materials – Part 3: Wrought titanium 6- aluminium 4- vanadium alloy.**

45. ISO 5832-4 **Implants for surgery – Metallic materials – Part 4: Cobalt-chromium-molybdenum casting alloy.**

46. ISO 5832-9 **Implants for surgery – Metallic materials – Part 9: Wrought high nitrogen stainless steel.**

47. ISO 5832-11 **Implants for surgery – Metallic materials – Part 11: Wrought titanium 6–aluminium 7–niobium alloy.**

48. ISO 5832-12 **Implants for surgery – Metallic materials – Part 12: Wrought cobalt-chromium-molybdenum alloy.**

49. ISO 5833 **Implants for surgery – Acrylic resin cements**

50. ISO 5834-1 **Implants for surgery – Ultra-high molecular weight polyethylene – Part 1: Powder form.**

51. ISO 5834-2 **Implants for surgery – Ultra-high molecular weight polyethylene – Part 2: Moulded form.**

52. ISO 6474-2 **Implants for surgery – Ceramic materials – Part 2: Composite materials based on a high-purity alumina matrix with zirconia reinforcement**

53. ISO 13782 **Implants for surgery -- Metallic materials -- Unalloyed tantalum for surgical implant applications**

54. ASTM F67 **Standard Specification for Unalloyed Titanium, for Surgical Implant Applications.**

55. ASTM F75 **Standard specification for Cobalt-28 Chromium-6 Molybdenum Alloy Castings and casting Alloy for Surgical Implants (UNS R30075).**

56. ASTM F136 **Standard specification for Wrought Titanium-6Aluminium-4Vanadium ELI (Extra Low Interstitial) Alloy for Surgical Implant Applications.**

57. ASTM F560 **Standard Specification for Unalloyed Tantalum for Surgical Implant Applications (UNS R05200, UNS R05400)**

58. ASTM F746 **Standard Test Methods for Pitting or Crevice Corrosion of Metallic Surgical Implant Materials**

59. ASTM F2026 **Standard Specification for Polyetheretherketone (PEEK) Polymers for Surgical Implant Applications**

60. ASTM F2063 **Standard Specification for Wrought Nickel-Titanium Shape Memory Alloys for Medical Devices and Surgical Implants**

61. ASTM F2924 **Standard Specification for Additive Manufacturing Titanium-6 Aluminum-4 Vanadium with Powder Bed Fusion**

62. ASTM F1044 **Standard Test Method for Shear Testing of Calcium Phosphate Coating and Metallic Coatings**

63. ASTM F1147 **Standard Test Method for Tensile Testing of Calcium Phosphate Coatings and Metallic Coatings**

64. ASTM F1160 **Standard Test Method for Shear and Bending Fatigue Testing of Calcium Phosphate and Metallic Medical Coatings**

65. ASTM F1185 **Standard Specification for Composition of Ceramic Hydroxyapatite for Surgical Implants**

66. ASTM F1580 **Standard Specification for Titanium and Titanium-6 Aluminium-4 Vanadium Alloy Powders for Coatings of Surgical Implants**

67. ASTM F1854 **Standard Test Method for Stereological Evaluation of Porous Coating on Medical Implants**

68. ASTM F1978 **Standard Test Method for Measuring Abrasion Resistance of Metallic Thermal Spray Coating by Using the Taber Abraser**

69. ISO 4288 **Geometrical Product Specification (GPS) – Surface texture: Profile method – Rules and procedures for the assessment of surface texture**

70. ISO 13779-2 **Implants for surgery – Hydroxyapatite – Coatings of hydroxyapatite.**

71. ISO 13779-3 **Implant for surgery – Hydroxyapatite – Part 3: Chemical analysis and characterization of crystallinity and phase purity**

72. ISO 13779-4 **Implants for surgery – Hydroxyapatite – Determination of coating adhesion strength**

73. FDA Guidance **Testing of Metallic Plasma Sprayed Coatings on Orthopaedic Implants to Support Reconsideration of Post-market Surveillance Requirements**

74. VDI 3198 **Coating (CVD, PVD) of cold forging tools**

75. VDI 3824 **Quality assurance of PVD and CVD hard coating - Inspection planning of hard coatings**

76. ASTM A 380 **Standard Practice for Cleaning, Descaling, and Passivation of Stainless Steel Parts, Equipment, and Systems**

77. ASTM F3127 **Standard guide for validating cleaning processes used during the manufacture of medical devices**

78. ASTM F 2459 **Standard Test Method for Extracting Residue from Metallic Medical Components and Quantifying via Gravimetric Analysis**

79. USP <788> **Particulate Matter in Injections test for a Large Volume Parenteral (LVP)**

80. ASTM A 967 **Standard Specification for Chemical Passivation Treatments for Stainless Steel Parts**

81. ISO 10993-1 **Biological evaluation of medical devices -- Part 1: Evaluation and testing within a risk management system**

82. ISO 10993-3 **Biological evaluation of medical devices -- Part 3: Tests for genotoxicity, carcinogenicity and reproductive toxicity.**

83. ISO 10993-5 **Biological evaluation of medical devices - Part 5: Tests for in vitro cytotoxicity**

84. ISO 10993-6 **Biological evaluation of medical devices -- Part 6: Test for local effects after implantation**

85. ISO 10993-7 **Biological evaluation of medical devices – Part 7: Ethylene oxide sterilization residuals**

86. ISO 10993-10 **Biological evaluation of medical devices -- Part 10: Tests for irritation and delayed-type hypersensitivity**

87. ISO 10993-11 Biological evaluation of medical devices -- Part 11: Test for systemic toxicity

88. ISO 10993-12 **Biological evaluation of medical devices -- Part 12: Sample preparation and reference materials**

89. ISO 10993-18 **Biological evaluation of medical devices -- Part 18: Chemical characterisation of materials**

90. USP <151> **Evaluation of pyrogens Pyrogen test procedure**

91. USP <87> **Biological reactivity test, in vitro**

92. USP <88> **Biological reactivity test, in vivo**

93. USP <161> **Transfusion and Infusion Assemblies and Similar Medical Devices**

94. ASTM F2102 **Standard guide for evaluating di extent of the oxidation in polyethylene fabricated forms intended for surgical implants**

95. ASTM F2381 **Standard test method for evaluating Trans-Vinylene Yield in Irradiate Ultra-High Molecular Weight polyethylene fabricated forms intended for surgical implants by Infrared Spectroscopy**

96. OECD 439 **In Vitro Skin Irritation: Reconstructed Human Epidermis Test Method**

97. ISO 15223-1 **Medical devices - Symbols to be used with medical device labels, labelling and information to be supplied**

98. ISO 11607-1 **Packaging for terminally sterilized medical devices -- Part 1: Requirements for materials, sterile barrier systems and packaging systems**

99. ISO 11607-2 **Packaging for terminally sterilized medical devices -- Part 2: Validation Requirements for forming, sealing and assembling processes**

100. ASTM F 88 / F88M **Standard test method for seal strength of flexible barrier materials**

101. ASTM F 1886 / F1886M **Standard Test Method for Determining Integrity of Seals for Medical Packaging by Visual Inspection**

102. ASTM F 1929 **Standard Test Method for Detecting Seal Leaks in Porous Medical Packaging by Dye Penetration**

103. ASTM F 3039 **Standard Test Method for Detecting Seal Leaks in Non Porous Medical Packaging by Dye Penetration**

104. ASTM F 2096 **Standard Test Method for Detecting Gross Leaks in Packaging by Internal Pressurization (Bubble Test)**

105. ASTM F 1980 **Standard guide for accelerated aging of sterile barrier systems for medical devices**

106. EN 868-5 **Packaging for terminally sterilized medical devices - Part 5: Heat-sealable bags and tubes made of one side of porous material and one side of plastic film - Requirements and test methods**

107. ASTM D 4169 **Standard practice for performance testing of shipping containers and systems**

108. DIN 58953-6 **Microbial barrier testing of packaging materials for medical devices which are to be sterilized**

109. ISO 11135 **Sterilization of health care products – Ethylene oxide -- Part 1: Requirements for development, validation and routine control of a sterilization process for medical devices.**

110. ISO 10993-7 **Biological evaluation of medical devices – Part 7: Ethylene oxide sterilization residuals**

111. ISO 11138-2 **Sterilization of health care products – Biological indicator – Part 2: Biological indicators for Ethylene oxide sterilization processes**

112. ISO 11137-1 **Sterilization of health care products -- Radiation -- Part 1: Requirements for development, validation and routine control of a sterilization process for medical devices.**

113. ISO 11137-2 **Sterilization of health care products -- Radiation -- Part 2: Establishing the sterilization dose.**

114. ISO 11137-3 **Sterilization of health care products -- Radiation -- Part 3: Guidance on dosimetric aspects.**

115. ISO 11737-1 **Sterilization of medical devices -- Microbiological methods -- Part 1: Determination of a population of micro-organisms on products**

116. ISO 11737-2 **Sterilization of medical devices -- Microbiological methods -- Part 2: Tests of sterility performed in the validation of a sterilization process.**

117. ISO 11138-1 **Sterilization of health care products – Biological indicator – Part 1: General requirements**

118. ASTM F2503-13 **Standard Practice for Marking Medical Devices and Other Items for Safety in the Magnetic Resonance Environment**

119. ASTM F2052 **Standard test method for measurement of magnetically induced displacement force on medical devices in the magnetic resonance environment**

120. F2182 **Standard test method for measurement of radio frequency induced heating on or near passive implants during magnetic resonance imaging**

121. F2213-06 **Standard test method for measurement of magnetically induced torque on medical devices in the magnetic resonance environment**

122. Johnson A.J., Desai S., Zhang C., Koh K., Zhang L.-Q., Costales T., O'Toole R.V., Manson T.T. **A Calcar Collar Is Protective Against Early Torsional/Spiral Periprosthetic Femoral Fracture: A Paired Cadaveric Biomechanical Analysis. The Journal of bone and joint surgery. American volume (2020) 102:16 (1427-1433).**

123. Quirynen T., Corten K., Segal O., Simon J.-P., Vander Sloten J., van Lenthe H.G. **Small interprosthetic gaps do not increase femoral peri-prosthetic fracture risk. An in vitro biomechanical analysis. Acta orthopaedica Belgica (2017) 83:2 (197-204).**

124. Liu D., Zhang L., Dong J., Li Z., Yu Q., Gong L., Tang P. **Clinical and biomechanical studies on cemented hemiarthroplasty treatment of advanced-age comminuted intertrochanteric fractures. International Journal of Clinical and Experimental Medicine (2016) 9:7 (13433-13444).**

125. Martin C.T., Heiner A.D., Baer T.E., Pugely A.J., Noiseux N.O. **Protrusio After Medial Acetabular Wall Breach in Total Hip Arthroplasty. The Iowa orthopaedic journal (2015) 35 (99-107).**

126. Hu K., Zhang X., Zhu J., Wang C., Ji W., Bai X. **Periprosthetic fractures may be more likely in cementless femoral stems with sharp edges. Irish Journal of Medical Science (2010) 179:3 (417-421).**

127. Lerch M., Olender G., von der Höh N., Thorey F., von Lewinski G., Meyer-Lindenberg A., Windhagen H., Hurschler C. **Ex vivo microCT analysis of possible microfractures of the femoral head during implantation of a cementless hip resurfacing femoral component. Ex-vivo-microCT-Untersuchungen zu möglichen Mikrofrakturierungen des Hüftkopfs bei der Implantation eines zementfreien Oberflächenersatzes am Hüftgelenk. Zeitschrift für Orthopädie und Unfallchirurgie (2009) 147:6 (716-720).**

128. A., Colwell C.W., D'Lima D.D. **Effect of head diameter on passive and active dynamic hip dislocation. Journal of Orthopaedic Research (2014).**

129. Kajino Y., Kabata T., Maeda T., Iwai S., Kuroda K., Fujita K., Tsuchiya H. **Strict component positioning is necessary in hip resurfacing. Journal of Orthopaedic Science (2013) 18:2 (290-297).**

130. Stewart K.J., Pedersen D.R., Callaghan J.J., Brown T.D. **Implementing capsule representation in a total hip dislocation finite element model. The Iowa orthopaedic journal (2004) 24 (1-8).**

131. Steinbrück A., Schröder C., Woiczinski M., Fottner A., Müller P.E., Jansson V. **The effect of trochlea tilting on patellofemoral contact patterns after total knee arthroplasty: An in vitro study. Archives of Orthopaedic and Trauma Surgery (2014) 134:6 (867-872).**

132. Daniilidis K., Höll S., Gosheger G., Dieckmann R., Martinelli N., Ostermeier S., Tibesku C.O. **Femoro-tibial kinematics after TKA in fixed- and mobile-bearing knees in the sagittal plane. Knee Surgery, Sports Traumatology, Arthroscopy (2013) 21:10 (2392-2397).**

133. Clark D.A., Upadhyay N., Gillespie G., Wakeley C., Eldridge J.D. **The correct rotation of the femoral component in patellofemoral replacement: A laboratory assessment of a surgical technique. Journal of Bone and Joint Surgery - Series B (2012) 94 B:12 (1637-1640).**

134. Thompson J.A., Hast M.W., Granger J.F., Piazza S.J., Siston R.A. **Biomechanical effects of total knee arthroplasty component malrotation: A computational simulation. Journal of Orthopaedic Research (2011) 29:7 (969-975).**

135. Moyad T.F., Hughes R.E., Urquhart A.**"Grand piano sign," a marker for proper femoral component rotation during total knee arthroplasty. American journal of orthopedics (Belle Mead, N.J.) (2011) 40:7 (348-352).**

136. Ghijselings S.G.M., Driesen R., Simon J.-P., Corten K. **Distal Extension of the Anterior Approach to the Hip Using the Femoral Interbundle Technique: Surgical Technique and Case Series. Journal of Arthroplasty (2017) 32:7 (2186-2190).**

137. Fu C., Wai J., Lee E., Myden C., Batuyong E., Hutchison C.R., Anglin. **Computer-assisted patellar resection for total knee arthroplasty. Computer aided surgery : official journal of the International Society for Computer Aided Surgery (2012) 17:1 (21-28).**

138. Kraaij G., Malan D.F., van der Heide H.J.L., Dankelman J., Nelissen R.G.H.H., Valstar E.R. **Comparison of Ho: YAG laser and coblation for interface tissue removal in minimally invasive hip refixation procedures. Medical Engineering and Physics (2012) 34:3 (370-377).**

139. Van Gorp C.C., Falk J.V., Kmiec Jr. S.J., Siston R.A. **The Reamer/Irrigator/Aspirator reduces femoral canal pressure in simulated TKA. Clinical Orthopaedics and Related Research (2009) 467:3 (805-809).**

140. Breusch S.J., Schneider U., Reitzel T., Kreutzer J., Ewerbeck V., Lukoschek M. **The significance of pulsatile lavage for improved in vitro and in vivo cement penetration. Die bedeutung der jet-lavage für das in-vitro- und in-vivo-zementierergebnis. Zeitschrift fur Orthopadie und Ihre Grenzgebiete (2001) 139:1 (52-63).**

141. AOANJRR, 2020 Annual Report

142. UKNJR, 17th Annual Report

RESEARCH TOPIC 27

Approach for the Clinical Investigation of Adverse Events or Complications Related to the Clinical Application of Total Joint Arthroplasty Devices and their Implantation Procedure

Authors

Francesco Siccardi[1], Sabine Rusch[2], Søren Overgaard[3,4,5,6], Alexander Giurea[7], Thomas M. Grupp[8,9], Arndt-Peter Schulz[10,11]

[1]Medacta International SA, Castel San Pietro, Switzerland

[2]Aesculap AG Medical Scientific Affairs, Tuttlingen, Germany

[3]Department of Clinical Medicine, Copenhagen University, Copenhagen, Denmark

[4]Department of Orthopaedic Surgery and Traumatology, Bispebjerg and Frederiksberg Hospital, University of Copenhagen, Copenhagen, Denmark

[5]Orthopaedic Research Unit, Department of Orthopaedic Surgery and Traumatology, Odense University Hospital, Odense, Denmark

[6]Department of Clinical Research, University of Southern Denmark, Odense, Denmark

[7]Department of Orthopedics and Trauma Surgery, Medical University of Vienna, Vienna, Austria

[8]Aesculap AG Research & Development, Tuttlingen, Germany

[9] Ludwig Maximilians University Munich, Department of Orthopaedic and Trauma Surgery, Musculoskeletal University Center Munich (MUM), Campus Grosshadern, Munich, Germany

[10]Department of Trauma Surgery, Orthopedic and Sports Traumatology, BG Klinikum Hamburg, Hamburg, Germany

[11]Fraunhofer Research Institution for Individualized and Cell-Based Medical Engineering, Lübeck, Germany

1. Abstract

Question: Perioperative and short-term postoperative (serious) adverse events

How can be proven clinically, that the implantation procedure and the implants do not induce unreasonably high rates of adverse events or complications (directly implant-related, e.g. inter-operative periprosthetic fractures, substantial bleeding, substantial migration, and generally related to the surgery / procedure, e.g. fast track, modified surgical approaches)?

Summary / Recommendation

To assess this parameter, clinical data from either clinical studies or other sources, e.g. electronic hospital records or registries, should be evaluated regarding pre-defined product-and procedure related complications occurring intraoperatively or within 6 months postoperatively.

Product- and procedure-related complications must be assessed primarily. According to clinical literature, the detection of short-term intra- and postoperative complications in clinical application can be safely established. Nevertheless, some aspects must be taken into consideration when planning their evaluation:

- selection of adequate study centres and patient cohorts
- different implant sizes must not be included specifically
- inclusion of different approaches is not mandatory in case of established approaches
- care should be taken to the recording of complications
- it is necessary to consider the use and safety of instruments.

Adherence to a very strict study design allows to optimally evaluate potential product- and procedure-related complications of a specific device which might be the most important goal in a pre CE safety study. After market launch data from post-market clinical follow-up studies or in-depth registry analysis can provide a realistic assessment of broad daily application of an implant when utilised by a larger number of surgeons.

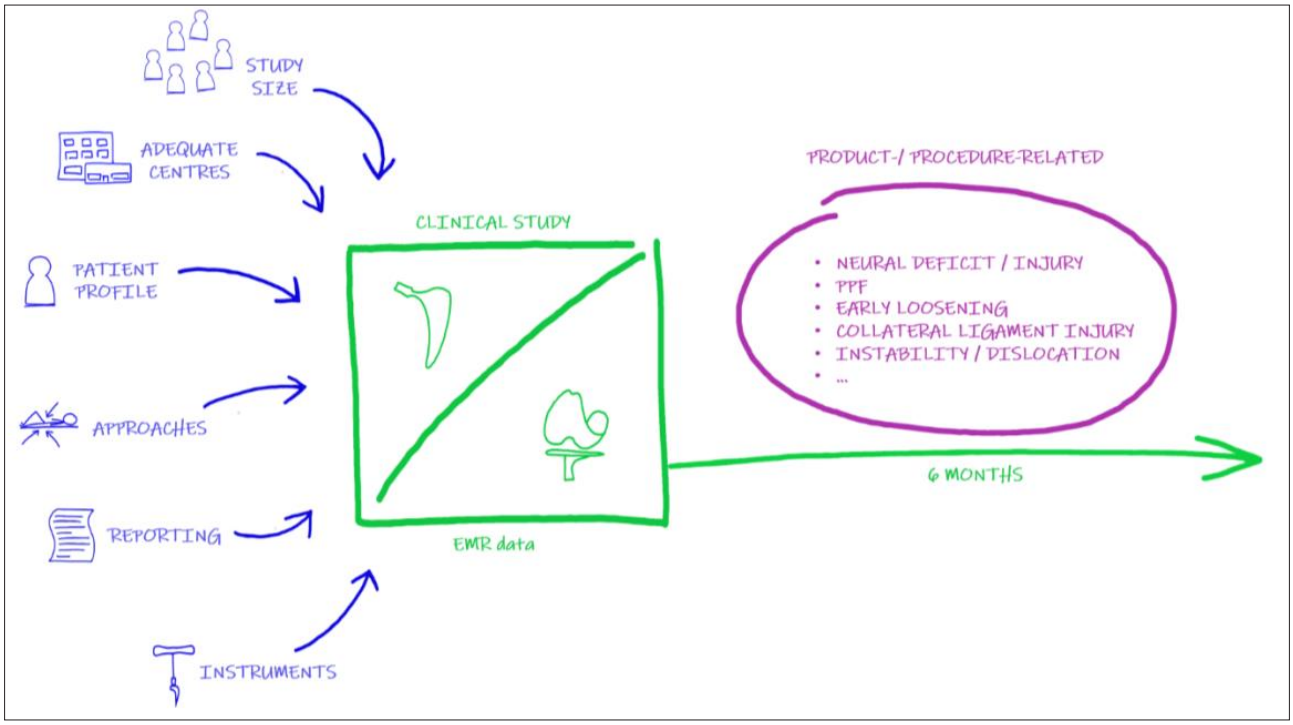
2. Level of Evidence

High

3. Consensus Delegate Vote

92% – super majority, strong consensus (92% agree / 3% disagree / 5% abstain)

4. Graphical Abstract



5. Search Strategy

A systematic literature search was conducted in March 2021 to identify literature with the potential to contribute to the answer of the above mentioned research question.

The electronic databases EMBASE including PubMed were explored using the combination of the following search-terms combined with the Boolean operator "AND":

- TKA/THA, knee/hip prosthesis, knee/hip replacement
- intra-/perioperative/short-term complications, adverse (device) effect, (serious) adverse event, complication rate
- (implantation procedure/process/workflow, surgical approach, alignment, surgical technique, safety study, classification) OR
  (automated/automatic, artificial intelligence, algorithm) OR
  (size/ implant/component/prosthesis possibilities/combinations/variations, implant/component/prosthesis sizes, combination possibilities, transfer of result/outcome/data) OR
  (orthopaedic equipment, surgical equipment, instrumentation, implantation instruments, instruments debris, instruments precision, reproducibility) OR
  (training surgery, surgical training, orthopaedic training, virtual reality/cyber training, training simulation, usability of instruments, ease of use, learning curve)

Search words were adapted to the specific search terms of the database and were searched including different spelling and synonyms. Article language was not restricted. Because of the aim of the process only studies from 2011 to 2021 were included to provide insight into current practice and state of the art procedures. Publication type was limited to papers for which a full text version was available. Abstracts from scientific meetings were not included in the current review as sufficient data cannot be extracted. Where necessary, the results were also restricted regarding study type to guarantee the inclusion of high quality literature.

The title and abstract of all articles were screened to select articles contributing to the topic. Potentially eligible articles were obtained for detailed assessment and were in detail screened for applicability. A manual search in the references of included articles yielded four more publications which were included. The selection was not further restricted. Duplicate articles were removed.

An article was found eligible when it concerned (1) information on product-/procedure-related short-term complications in total joint arthroplasty or (2) information about study design for the detection of product-/procedure-related complications in total joint arthroplasty.

116 publications met the search criteria. After the selection process, 31 articles were included in the review.

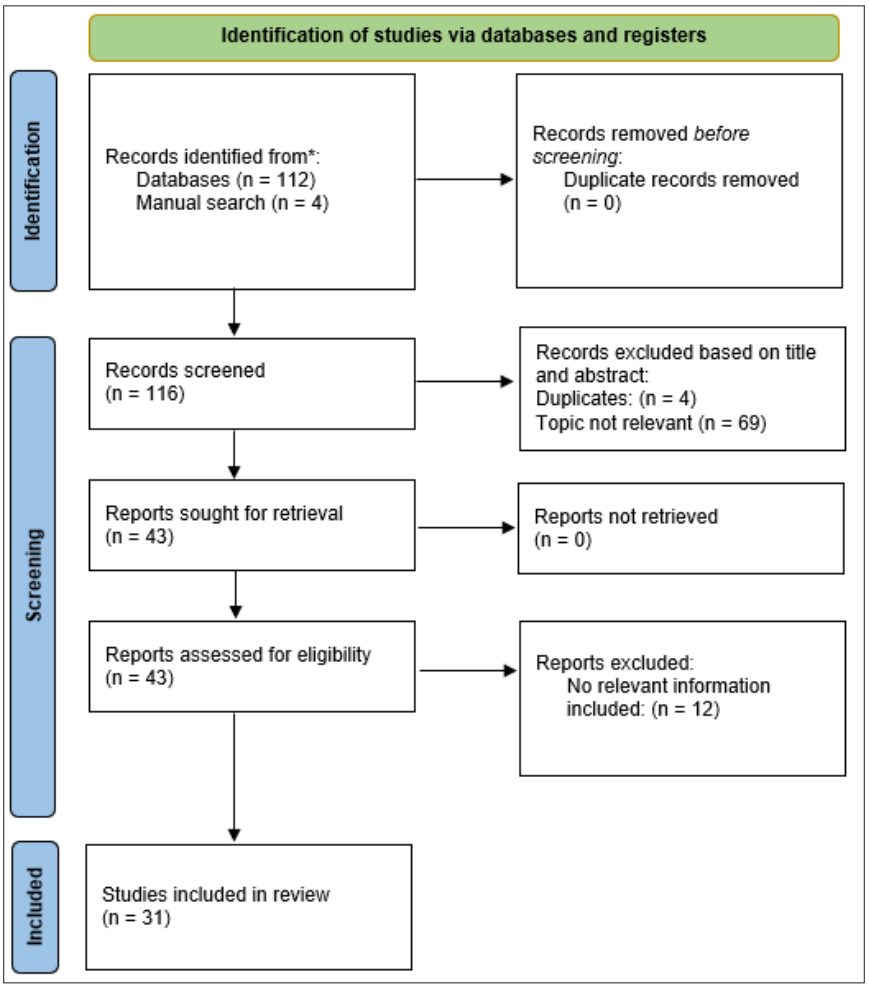


Figure 1: Flow chart literature selection process.

6. Rationale

Despite the importance of evaluating patient outcomes of total joint arthroplasty (TJA) clinical assessment of complications is equally important in evaluating orthopaedic implants. Short-term intra- or postoperative complications are closely related to implant safety and can occur for many reasons, including evolving disease processes, surgical, medical or nursing errors, patient noncompliance, or events beyond physician control such as falls or trauma and must be expected with surgical procedures at a small but finite incidence, despite reasonable and safe care (1).

To reliably establish in clinical assessment that a (new) implant and the implant specific instruments are not associated with unreasonably high rates of complications it is initially necessary to define the complications to be taken into consideration. A systematic search of the literature was undertaken to answer this question. The strength of evidence was graded as high, based on the identified literature pertaining to the topic as well as on the vast experience with TKA and THA implants, both published (registries, clinical results) and intangibly existing (manufacturer´s knowledge, predicate devices).

Definition of short-term complications

For the evaluation product-related complications must be considered primarily, although even those related to product design may be influenced by patient factors (2), e.g. unexpectedly high rates of periprosthetic fractures (PPF) which can clearly be associated with poor bone quality (3) or neurologic disorders and gait imbalance. Similarly, early hip instability/ dislocation also might be due to factors like biomechanics or neurologic impairment (4).

Procedure-related complications which are primarily dependent on the specific procedure or approach (5) should also be considered and might indicate shortcomings of instruments or implants within a specific approach, i.e. factors that can be controlled by manufacturers.

Complications or readmissions generally related to TJA surgery (e.g. general postsurgical pain, sepsis, gastrointestinal/genitourinary complications (6)) are usually not implant-related (7, 8). No information substantially opposing this assumption could be found in the literature and we assume that this is not of clinical relevance for assessment of implants.

The definition of early complications depends on the selected timeframe. In the assessed clinical literature complications are considered as early complications within 30 days after surgery (9–12), 90 days of surgery (13–17), or up to 4 months (3). In accordance with these findings, complications within 6 months are defined as short-term and a follow-up of 6 months is considered adequate for their detection as rates of those complications do not increase substantially also in longer follow-up (18).

We consider it necessary to establish a list of complications. Healy et al. developed standardised lists of complications associated with TKA in 2013 and with THA in 2016. TKA and THA Complications Workgroups proposed lists of complications based on orthopaedic literature to be validated by expert surveys (1, 6). Complications were then further stratified for severity using a validated system based on a modification of the Sink classification (19, 20). 22 TKA and 19 THA complications were identified. However, there is no distinction regarding timeframe or origin of the complications.

In accordance with this list we consider the following complications relevant for both TKA and THA in the short-term and related to the product or the procedure (1, 6):

- Neural deficit (i.e. Postoperative neural deficit (sensory or motor) related to the index TKA/THA)
- Periprosthetic fracture: fracture of distal femur, proximal tibia, or patella in TKA, or proximal femur or the acetabulum in THA
- Early implant loosening: Implant loosening confirmed intraoperatively or identified radiographically as a change in implant position (implant migration) or progressive, radiolucent lines at the bone-cement or bone-implant interface

In addition specifically for TKA we consider (1):

- Medial/lateral collateral ligament injury: Intraoperative or early postoperative injury requiring repair, reconstruction, a change in prosthetic constraint, or TKA protocol
- Instability: Symptomatic instability reported by the patient and confirmed by laxity on physical examination as defined by The Knee Society Knee Score
- Malalignment: Symptomatic malalignment reported by the patient and confirmed radiographically with angular deformity in the coronal plane >5° from the mechanical axis or the intended intraoperative alignment (kinematic, anatomic, etc.)
- Stiffness: Limited ROM reported by the patient and demonstrated in physical examination with extension limited to 15° short of full extension or flexion <90° (n/a if preoperative arc of motion <75°)
- Patellofemoral dislocation: Dislocation of patella from the femoral trochlea
- Tibiofemoral dislocation: Dislocation of the tibiofemoral joint

In addition specifically for THA we consider (6):

- Dislocation/Instability: Dislocation of the femoral head out of the acetabulum or recurrent symptomatic subluxation of the hip
- Cup/liner dissociation: Dissociation of the cup/liner from the acetabular shell

From literature review no further complications could be identified to be included. Nevertheless, any pre-clinical or clinical evaluation must always expect and allow the detection of previously unknown or new complications which might be associated with a device. If a device explicitly requires the surveillance of further device-specific complications, those will be identified within the pre-clinical manufacturer risk assessment and consequently have to be included in the clinical assessment.

Specific attention should be paid to those adverse events while they are not indicative of substandard care or suggest a TJA was not planned or performed correctly. As again emphasized, complications can be associated with medical treatments despite the delivery of reasonable and safe care. Overall complication rates of modern TJA are low and the number of truly product and procedure related complications is even lower (14, 9, 18).

Relevant aspects for the clinical collection of data

From clinical literature can be established that clinical detection and assessment of short-term intra- and postoperative complications is straightforward. Nevertheless, the following aspects must be considered when planning the evaluation of short-term complications in the assessment of orthopaedic implants. It has to be kept in mind, that these recommendations allow to optimally evaluate the occurrence of product- and procedure related complications, which is the most important goal in a pre CE safety study. Only data from post-market clinical follow-up studies or in-depth registry analysis after market launch can provide a realistic assessment of broad daily application of an implant as it includes a wider range of the following factors.

Adequate study centres must be selected as there are predictors for higher complication rates related to hospital factors which may confound results with the study products (17). The centres should provide adequate case volume and experience with the surgery as well as with Good Clinical Practice studies. If specific and justifiable requirements regarding type of user (surgeon/hospital) or included patient population exist for a device those must of course be considered.

It is necessary to carefully establish values for comparison and to select patients accordingly. As mentioned earlier, some complications e.g. femoral PPF in THA also strongly depend on patient factors (15, 2). Omari et al. showed that introduction of a new treatment algorithm selecting female patients >60 years to receive cemented instead of cementless femoral components could significantly reduce the number of PPF (15). Presurgical comorbidities also strongly influence surgical outcomes and should be considered in patient selection (21). Guofeng et al. established in a meta-analysis that patients with metabolic syndrome undergoing TJA are at increased risk of all-cause complications, and 30-day readmission (10). Sherman et al. compared medical records of TJA patients and found that patients with neuromuscular diseases exhibited a significantly higher risk for multiple systemic complications during hospital stay and post-discharge (14). So, as it may be important to include demographically different patients in a clinical study it must be considered that some patient groups have a higher risk for complications which could disguise truly implant-related complications.

Considering the inclusion of different implant sizes it is not necessary to specifically include all different implant sizes in a clinical study to identify implant-related short-term complications which may be size-related. From the literature no information could be identified that complication rates substantially differ between sizes of one specific implant (system). There are very few reports that specific complications may be related to implant size, but it is not feasible to detect rare complications in a conventional clinical study which can include only standard sizes in reasonable case numbers. Such reports are for example varus collapse after primary TKA which has been associated with obesity and generally smaller tibial baseplate sizes. Still, this is a long-term complication as average time to revision for varus collapse was 6.9 years (range 1-19) (22). Other complications, e.g. impingement after TKA induced by modified popliteal tracking, have been associated with inappropriate implant size related to surgical planning but not with specific implant sizes (23). A possible relationship between lower BMI and large prosthesis size and high intraoperative blood loss in TKA are explained anatomically by the larger bleeding surface from the bony cut and are also not considered truly implant related (8).

Inclusion of different approaches may be considered in clinical studies. Some approaches, which have promising advantages such as e.g. the direct anterior approach (DAA) in THA, are technically demanding procedures potentially prone to specific complications and pitfalls (24, 5, 25). Perioperative complications and component misalignment also seem to be more prevalent in MIS THA due to the constrained surgical field and difficult visualization of anatomical landmarks (26). Docter et al. compared complication rates of implant-related as well as surgery-related complications between common surgical approaches in THA in a systematic review and meta-analysis (13). They found that small differences, e.g., the posterior approach was associated with a higher risk of dislocation (compared to the anterior, lateral, and anterolateral) but lower risk of loosening (compared to the lateral and anterolateral approach). A systematic review by Geest et al. evaluated the cumulative risk incidence of intra- and postoperative complications with DAA in THA and also found that a risk for intra-operative fractures and postoperative transient lateral cutaneous femoral nerve impairment was evident, although the risk for other complications was comparable to other approaches (27). So, even though an implant may not be suitable for all approaches (5, 28), the inclusion of different approaches is only considered necessary if an implant is designed specifically for a new approach or requires the use of specialized instruments for surgical approaches (28).

Considering the study size and design, also a small group of patients allows the early detection of complications which are associated with a systematic failure mode if the study design is adequate (3). Clement et al. terminated their study early due to a significantly ($p = 0.004$) higher rate of intraoperative complications in uncemented compared to cemented THA for displaced intracapsular hip fractures while the overall patient number was only 50 (3). Controlled designs would of course be ideal but otherwise distinct comparative values should be determined. As especially truly implant-related complications are rare, a safety study with a small cohort the occurrence of any complications always requires qualitative analysis, taking into consideration possible influencing factors such as patient characteristics or surgical technique. A small study size is considered adequate to detect general safety issues with new implants but of course not to establish absolute occurrence rates. To determine small differences in early complication rates, large groups of patients are necessary which is not feasible before CE marking. To answer such specific questions or to rule out other source factors good post-market clinical follow-up (PMCF) measures must be established by manufacturers, e.g. registry surveillance or specific PMCF studies of hospital records (2, 9, 29). From such studies, even differences within one implant can be evaluated. Colacchio et al. showed that a tapered-wedge stem system resulted in over 4 times higher incidence of intraoperative femur fracture than its second-generation successor used for primary uncemented THA (2).

Special care should be taken to the recording of complications to ensure the quality of data. Additional information, e.g. intra- or postoperative occurrence of fractures, type of treatment, direction of any dislocation or instability, are essential to evaluate complications (1, 6). Valid reporting requires standardised data collection, consistent methods of reporting, and utilization of validated tools. Every complication should be reported and recorded to detect potential problems with patient selection, surgical procedures, follow-up care, or implants (1).

Only a few examples of a relationship between instruments and short-term complications could be found in the literature even though such a relationship clearly exists. One example is PPF in cementless stems during broaching which can be influenced by the geometry, offset or tooth pattern of the broach, or the operative technique employed. Specific offset broaches and reamers in DAA may alter force transmission or direction introduced into the bone and lead to specific complications (2, 28). Such instrument related complications are usually detected intraoperatively or early postoperatively within a typical clinical study setting or also from hospital records (16, 5). No specific study design is needed but a clinical study to detect early complications should specifically include questions and documentation about use, safety and performance of instruments, e.g. clear documentation which (generation of) instruments were used to establish a link between the results and defined instruments. To answer specific instrumentation questions, a multicentre design or at least inclusion of several surgeons is reasonable to eliminate confounders originating from different personal surgical techniques, habits and customs. The majority intraoperative complication related to instruments due to poor pre-clinical development should be detected pre-clinically or at the latest in a short-term clinical study, e.g. trial head disengagement from the femoral trial stem/rasp and migration into the pelvic cavity due to a misfit between trial and femoral broach (5, 30). If changes induced by long-term use of instrumentation, e.g. repeated sterilisation are the cause, this can also not be identified in short-term studies but requires an effective PMCF system of manufacturers (i.e. complaint surveillance and vigilance system).

Presence of a learning curve is common with new implants and one should remain cautious when utilizing a new system in a clinical study. Outcomes undoubtedly improve with experience as Padilla et al. demonstrate that for experienced surgeons, the majority of PPF (80%) following short stem THA with a new design occurs within the first 30 cases, representing the initial learning curve period (16). Several studies suggest the presence of a learning curve also when transitioning to a new approach on the basis of decreasing complication rates (5, 24, 28). Sendtner et al. found that MIS THA allows similar complication rates and proper implant positioning during the learning curve compared to standard THA (26). Kagan et al. on the other hand demonstrate no associated learning curve for an experienced surgeon when switching routine THA approach from mini posterior to anterior-based muscle-sparing regarding intraoperative and immediate postoperative complications (12). If a new approach is required precautions can be taken to minimise the learning curve or keep it short (26, 24). The surgeon should be familiar with the procedure and carefully trained for potentially unusual approaches.

Alternatives to traditional clinical studies

There may be alternatives to clinical studies for the detection of short-term complications. To facilitate data collection and allow the inclusion of large patient cohorts, especially for intraoperative and immediate postoperative complications follow-up can also be performed through hospital and primary care medical records (9).

The widespread use of electronic medical records (EMR) provides access to vast quantities of information. Artificial intelligence software allows to develop algorithms for extracting relevant complications from EMR to provide continuous feedback on outcomes with a performance level comparable to manual data extraction, but with greater speed. For data points with clear patterns, e.g. implant sizes, even with minimal training the algorithm outperforms manual processes (31). Semiautomated surveillance may substantially reduce workload and subjective data interpretation in the detection of procedure- and product-related complications after market launch of a product and may also allow to detect rare complications or complications associated with specific indications, patient groups or other factors. Verberk et al. assessed the validity of an algorithm for semiautomated surveillance of one specific complication (deep surgical site infection, not product-related) after TJA in Dutch hospitals (29). However, such possibilities depend on the adoption and form of electronic health records in a specific region of interest, validation of an algorithm in terms of sensitivity, positive predictive value, and workload reduction for a specific setting is challenging (29) and this is a promising option but not yet to be expected as a standard method.

National or regional databases if reported comprehensively also may provide a source for complication research after CE (14). Especially databases which record specific preoperative and postoperative patient characteristics may provide valuable information but may not be easily accessible or available (11). Reporting from registries is difficult as frequently complications are not listed exhaustively. Factors like thirty-day morbidity and mortality after surgery are reported but from those it is difficult to further conclude results. Extensive studies in high-quality registries, however, yield valuable results (17) as do in-depth analysis which can often be requested. Although, access to source data is frequently necessary to extract specific results which are not only overall outcomes.

Conclusion

Clinical assessment of product- and procedure-related complications occurring intraoperatively or within 6 months postoperatively can be safely established. While important aspects regarding study design must be taken into consideration when planning the evaluation of such complications this measure is always part of a stepwise process of product evaluation. Within the cascade of activities, the definition of specific complications from pre-clinical risk assessment is the basis. Pre CE studies then serve to exclude unknown short-term/intraoperative complications while additional data from post-market clinical follow-up studies or in-depth registry analysis later provide a realistic assessment of broad daily application of an implant when utilised by a larger number of surgeons.

7. References

1. Healy WL, Della Valle CJ, Iorio R, Berend KR, Cushner FD, Dalury DF et al. **Complications of total knee arthroplasty: standardized list and definitions of the Knee Society. Clin Orthop Relat Res 2013; 471(1):215–20.**

2. Colacchio ND, Robbins CE, Aghazadeh MS, Talmo CT, Bono JV. **Total Hip Intraoperative Femur Fracture: Do the Design Enhancements of a Second-Generation Tapered-Wedge Stem Reduce the Incidence? The Journal of Arthroplasty 2017; 32(10):3163–8.**

3. Clement ND, van der Linden M, Keating JF. **Higher rate of complications with uncemented compared to cemented total hip arthroplasty for displaced intracapsular hip fractures: A randomised controlled trial of 50 patients. Eur J Orthop Surg Traumatol 2021; 31(3):587–94.**

4. Sheth NP, Paprosky WG. **Chronic total hip arthroplasty instability—Causes and cures. Seminars in Arthroplasty 2016 [cited 2021 Apr 23]; 27(1):74–9.**

5. Melman WPR, Mollen BP, Kollen BJ, Verheyen CCPM. **First experiences with the direct anterior approach in lateral decubitus position: learning curve and 1 year complication rate. Hip Int 2015; 25(3):251–7.**

6. Healy WL, Iorio R, Clair AJ, Pellegrini VD, Della Valle CJ, Berend KR. **Complications of Total Hip Arthroplasty: Standardized List, Definitions, and Stratification Developed by The Hip Society. Clin Orthop Relat Res 2016 [cited 2021 Apr 23]; 474(2):357–64.**

7. Sephton BM, Bakhshayesh P, Edwards TC, Ali A, Kumar Singh V, Nathwani D. **Predictors of extended length of stay after unicompartmental knee arthroplasty. J Clin Orthop Trauma 2020; 11(Suppl 2):S239–S245.**

8. Sa-Ngasoongsong P, Wongsak S, Kulachote N, Chanplakorn P, Woratanarat P, Kawinwonggowit V. **Predicting Factors for Allogeneic Blood Transfusion and Excessive Postoperative Blood Loss after Single Low-Dosage Intra-Articular Tranexamic Acid Application in Total Knee** Replacement. Biomed Res Int 2017; 2017:2729487.

9. Ripollés-Melchor J, Abad-Motos A, Díez-Remesal Y, Aseguinolaza-Pagola M, Padin-Barreiro L, Sánchez-Martín R et al. **Association Between Use of Enhanced Recovery After Surgery Protocol and Postoperative Complications in Total Hip and Knee Arthroplasty in the Postoperative Outcomes Within Enhanced Recovery After Surgery Protocol in Elective Total Hip and Knee Arthroplasty Study (POWER2). JAMA Surg 2020; 155(4):e196024.**

10. Guofeng C, Chen Y, Rong W, Ruiyu L, Kunzheng W. **Patients with metabolic syndrome have a greater rate of complications after arthroplasty: A systematic review and meta-analysis. Bone & Joint Research 2020; 9(3):120–9.**

11. Duchman KR, Gao Y, Pugely AJ, Martin CT, Callaghan JJ. **Differences in short-term complications between unicompartmental and total knee arthroplasty: a propensity score matched analysis. J Bone Joint Surg Am 2014; 96(16):1387–94.**

12. Kagan RP, Greber EM, Richards SM, Erickson JA, Anderson MB, Peters CL. **Advantages of an Anterior-Based Muscle-Sparing Approach in Transitioning From a Posterior Approach for Total Hip Arthroplasty: Minimizing the Learning Curve. The Journal of Arthroplasty 2019; 34(12):2962–7.**

13. Docter S, Philpott HT, Godkin L, Bryant D, Somerville L, Jennings M et al. **Comparison of intra and post-operative complication rates among surgical approaches in Total Hip Arthroplasty: A systematic review and meta-analysis. Journal of Orthopaedics 2020; 20:310–25.**

14. Sherman WF, Wu VJ, Ofa SA, Ross BJ, Savage-Elliott ID, Sanchez FL. **Increased rate of complications in myasthenia gravis patients following hip and knee arthroplasty: a nationwide database study in the PearlDiver Database on 257,707 patients. Acta Orthopaedica 2021:1–6.**

15. Omari A, Nielsen CS, Husted H, Otte KS, Troelsen A, Gromov K. **Introduction of a New Treatment Algorithm Reduces the Number of Periprosthetic Femoral Fractures After Primary Total Hip Arthroplasty in Elderly Females. The Journal of Arthroplasty 2020; 35(12):3613–20.**

16. Padilla JA, Anoushiravani AA, Feng JE, Schwarzkopf R, Slover J, Marwin S. **The learning curve following adoption of a novel short-stem prosthesis in total hip arthroplasty: implications on short-term patient outcomes. Eur J Orthop Surg Traumatol 2019; 29(4):819–25.**

17. Jolbäck P, Rolfson O, Cnudde P, Odin D, Malchau H, Lindahl H et al. **High annual surgeon volume reduces the risk of adverse events following primary total hip arthroplasty: a registry-based study of 12,100 cases in Western Sweden. Acta Orthopaedica 2019; 90(2):153–8.**

18. Fuchs MCHW, Janssen RPA. **Clinical evaluation of 292 Genesis II posterior stabilized high-flexion total knee arthroplasty: range of motion and predictors. Eur J Orthop Surg Traumatol 2015; 25(1):161–6.**

19. Iorio R, Della Valle CJ, Healy WL, Berend KR, Cushner FD, Dalury DF et al. **Stratification of standardized TKA complications and adverse events: a brief communication. Clin Orthop Relat Res 2014; 472(1):194–205.**

20. Sink EL, Leunig M, Zaltz I, Gilbert JC, Clohisy J. **Reliability of a complication classification system for orthopaedic surgery. Clin Orthop Relat Res 2012; 470(8):2220–6.**

21. Kerkhoffs GMMJ, Servien E, Dunn W, Dahm D, Bramer JAM, Haverkamp D. **The influence of obesity on the complication rate and outcome of total knee arthroplasty: a meta-analysis and systematic literature review. J Bone Joint Surg Am 2012; 94(20):1839–44.**

22. Martin JR, Fehring KA, Watts CD, Springer BD, Fehring TK. **Radiographic Findings in Patients With Catastrophic Varus Collapse After Total Knee Arthroplasty. The Journal of Arthroplasty 2018; 33(1):241–4.**

23. Bonnin MP, Kok A de, Verstraete M, van Hoof T, van der Straten C, Saffarini M et al. **Popliteus impingement after TKA may occur with well-sized prostheses. Knee Surg Sports Traumatol Arthrosc 2017; 25(6):1720–30.**

24. Dall'Oca C, Ceccato A, Cresceri M, Scaglia M, Guglielmini M, Pelizzari G et al. **Facing complications of direct anterior approach in total hip arthroplasty during the learning curve. Acta Biomed 2020; 91(4-S):103–9.**

25. Edwards DS, Barbur SAR, Bull AMJ, Stranks GJ. **Posterior mini-incision total hip arthroplasty controls the extent of post-operative formation of heterotopic ossification. Eur J Orthop Surg Traumatol 2015; 25(6):1051–5.**

26. Sendtner E, Borowiak K, Schuster T, Woerner M, Grifka J, Renkawitz T. **Tackling the learning curve: comparison between the anterior, minimally invasive (Micro-hip®) and the lateral, transgluteal (Bauer) approach for primary total hip replacement. Arch Orthop Trauma Surg 2011; 131(5):597–602.**

27. Geest T de, Fennema P, Lenaerts G, Loore G de. **Adverse effects associated with the direct anterior approach for total hip arthroplasty: a Bayesian meta-analysis. Arch Orthop Trauma Surg 2015; 135(8):1183–92. Available from: URL: https://link.springer.com/article/10.1007/s00402-015-2258-y.**

28. Zlotnicki JP, O'Malley MJ. **Direct Anterior Approach for Total Hip Arthroplasty: Implant, Instrument, and Approach. Operative Techniques in Orthopaedics 2017; 27(3):192–7.**

29. Verberk JDM, van Rooden SM, Koek MBG, Hetem DJ, Smilde AE, Bril WS et al. **Validation of an algorithm for semiautomated surveillance to detect deep surgical site infections after primary total hip or knee arthroplasty-A multicenter study. Infect Control Hosp Epidemiol 2021; 42(1):69–74.**

30. Abouel-Enin S, Fraig H, Griffiths J, Latham J. **Intra-pelvic migration of femoral head trial in total hip arthroplasty, a rare intra-operative complication: a systematic literature review. Musculoskelet Surg 2016; 100(2):77–81.**

31. van de Meulebroucke C, Beckers J, Corten K. **What Can We Expect Following Anterior Total Hip Arthroplasty on a Regular Operating Table? A Validation Study of an Artificial Intelligence Algorithm to Monitor Adverse Events in a High-Volume, Nonacademic Setting. The Journal of Arthroplasty 2019; 34(10):2260–6.**

RESEARCH TOPIC

REVISION RATE/ SURVIVAL TIME

Research Topic 28

Assessing Revision Rates, Lifetime and Survival Time of Total Joint Arthroplasty Implants with Regard to Benchmark Values, Pre-CE Studies, Influencing Factors and Relevant Parameters

RESEARCH TOPIC 28

Assessing Revision Rates, Lifetime and Survival Time of Total Joint Arthroplasty Implants with Regard to Benchmark Values, Pre-CE Studies, Influencing Factors and Relevant Parameters

Authors

Volkmar Jansson[1], Ashley Blom[2,3], Barbara Bordini[4], Anne Lübbeke[5,6]

[1]Ludwig Maximilians University Munich, Department of Orthopaedic and Trauma Surgery, Musculoskeletal University Center Munich (MUM), Campus Grosshadern, Munich, Germany

[2]National Institute for Health Research Bristol Biomedical Research Centre, University Hospitals Bristol and Weston NHS Foundation Trust and the University of Bristol, Bristol, UK

[3]Translational Health Sciences, Bristol Medical School, Musculoskeletal Research Unit, University of Bristol, Learning & Research Building, Southmead Hospital, Bristol, UK

[4]Laboratorio di Tecnologia Medica, IRCCS Istituto Ortopedico Rizzoli, Bologna, Italy

[5]Division of Orthopaedics and Trauma Surgery, Geneva University Hospitals, Geneva, Switzerland

[6]Nuffield Department of Orthopaedics, Rheumatology and Musculoskeletal Sciences, University of Oxford, Oxford, UK

1. Abstract

Question: Revision rate / survival time

1. How can be verified that the lifetime of implants corresponds to the benchmark values of modern endoprosthesis?
2. It is possible to evaluate this pre-clinically or within a pre-CE study (study design, documentation of broad application including all indications ...)?
3. Which comparative values can be determined?
4. How can different factors regarding patients, at short or long-term, be considered in survival analysis?
5. Which rates can be analysed, e.g. reoperation, complications?
6. How can different competing risk factors be considered such as age, power calculation?

Summary / Recommendation:

There is agreement that the revision rate is the main outcome in the field of arthroplasty complications. It is an unambiguous indicator of the performance of the implant and the surgery, and relevant to the patient in terms of burden and society in terms of costs. All-cause revision should always be reported. Revision for specific causes such as infection, dislocation/instability, fracture and loosening may additionally be indicated. Reoperation rates and rates of major complications (specific causes) are other important outcomes although in need of data capture/quality improvement. Kaplan-Meier survival analysis is the recommended approach to determine revision rate estimates. Presentation of cumulative failure/survival estimates with 95% confidence intervals at standardized points in time by types of implant allows for comparison and aggregation. Competing risks analyses may be considered when follow-up times exceed 10 years. Indication, age and gender should be available at the least. The minimum number of time points for benchmarking should be 2, 5 and 10 years. An all-cause revision rate of 5% or less at 10 years is recommended/applied by NICE, ODEP rating 10A* and the ISAR benchmarking guidance document for the superiority approach. The minimum number at risk (10-year benchmark) should be 500 procedures at 10 years. Greater consistency in reporting of surgical and patient-reported outcomes among joint replacement registries will improve the interpretation and comparability of these data to monitor and benchmark outcomes accurately.

2. Level of Evidence

Low

3. Consensus Delegate Vote

97% – unanimous, strongest consensus (97% agree / 0% disagree / 3% abstain)

4. Graphical Abstract

n/a

5. Search Strategy

No formal literature search was conducted as registry data and methods are well described in the cited literature.

6. Rationale

1. How can be verified that the lifetime of implants corresponds to the benchmark values of modern endoprosthesis?

The European Medicines Agency (EMA) defines registries as organised systems that use observational methods to collect uniform data on a population defined by a particular disease, condition, or exposure that is followed over time with the aim to improve quality of patient care. The importance of registries is acknowledged within the MDR (Medical Device Regulations, article 108, Annex VII 4.11(g).

In principle it is difficult to verify the validity of benchmark values with respect to the lifetime of implants as the golden standard of lifetime evaluation of implants does not exists. The lifetime of an implant is currently verified in national or regional registries (population-based cohorts) with follow-up over decades (1). National or regional registries provide large amounts of standardized prospectively collected data. The quality and completeness of the data collection is paramount. Since the design is observational and not experimental several biases, especially due to selection and confounding, need to be taken into account. The collection of relevant baseline variables is important in this context. Randomized controlled trials (RCTs) - preferably nested RCTs within registries - will be needed for more detailed evaluation in particular when proxies (RSA, EBRA, CT based) (2–5) are used to compare implant safety and performance, when evaluating functional activity and habits of patients or when comparing surgical techniques. Nesting RCTs within registries makes their realisation faster at lower costs (6) and allows to better appreciate the generalisability of the RCTs findings since those patients who are not included in the RCT are still followed in the registry. A good overview of strengths and weaknesses of clinical data and registry data is given by the ISAR benchmarking workgroup (https://www.isarhome.org/publications).

According to ISAR benchmarking workgroup it is agreed that all-cause revision as failure of implant definition is the principal outcome measure of choice. Revision surgery is an indicator of the performance of the implant and the surgery, is relevant to the patient in terms of burden, to society in terms of costs, it is an unambiguous endpoint and there is longstanding international experience with its use

(ISAR document, 7, 8). There are also challenges and downsides associated with it (7), especially the fact that not all failures are revised. All-cause revision should always be reported. Revision for specific causes such as infection, dislocation/instability, fracture and loosening may additionally be indicated.

There might be the need to extend this definition by "additional operations" as used in the EPRD (EPRD Jahresbericht 2019/2020) or any other reoperation as in other registries (9). Secondary resurfacing of patella is an example in which such operation can be counted as failure of the whole implant including the unrevised femoral and tibial components or not. As indication and clinical standards do influence such additional operations it seems to be difficult to regard such additional operations as an implant specific problem alone.

Kaplan-Meier survival analysis is the statistical method of choice to determine the revision rate estimate (ISAR 2018,(10)). Some registries use PTIR (prosthesis time incident rate) as well or in addition (see EFORT European Consensus Statement 30 – Registry Studies). Presentation of cumulative failure/survival estimates with 95% confidence intervals at standardized points in time by implant/types of implants allows for comparison and aggregation and is reality in an increasing number of registry reports. Competing risks or other types of time-to-event analyses may be additionally considered in specific circumstances (in particular follow-up longer than 10 years, very old population).

2. It is possible to evaluate this pre-clinically or within a pre-CE study (study design, documentation of broad application including all indications ...)?

Pre-clinical or pre-CE studies can fill the gap between valid registry data and first clinical use. This is described in detail in the IDEAL framework and in the publications on stepwise introduction of new implants (11–13). These studies should be based on individual case evaluation rather than on statistical figures only. Study design and evaluation processes like NJR "Beyond compliance" are a good basis for such evaluation (www.beyondcompliance.org.uk). Registry data can be used but should be interpreted carefully with respect to very short evaluation time. More recently, IDEAL has advocated that registry monitoring of high-risk medical devices should start from "first-in-human experience onward" (14).

3. Which comparative values can be determined?

We understand that the question aims at adequate grouping of implants. This means e.g. that in TKR PS, CR, cemented etc. and in THR cemented stems, cementless cups etc. should be grouped for analysis. It is consensus that this grouping should be used mainly but adapted to the needs of future or special implants. In the ISAR benchmarking working group document (https://www.isarhome.org/publications) a well-designed grouping is suggested.

It is consensus that good quality registry data should be the preferred source for benchmarking implants ((15), ISAR document).

Revision for all reasons is the most straightforward and most easily assessed outcome measure for both early and later benchmarking (ISAR document).

However, a re-operation includes not only revisions for all reasons, but also includes any secondary procedure undertaken including added implant parts like a secondary retropatellar resurfacing. From a patient perspective they both represent forms of failure and ideally registries should endeavour to capture both revision and all reoperations.

An important point for benchmarking is evaluation time and statistical failure criteria. The ISAR benchmarking group document (ISAR 2018) states the following regarding evaluation times: "The minimum number of time points for benchmarking should be 2, 5 and 10 years with consideration of including a 15-year benchmark."

As stated by Kandala and coworkers (15) and according to the 2014 NICE Guidance (Total hip replacement and resurfacing arthroplasty for end-stage arthritis of the hip technology appraisal guidance 2014 www.nice.org.uk/guidance/ta304) it is appropriate to recommend that a prosthesis (for either conventional or resurfacing THR) should meet a revision rate of 5% or less at 10 years determined by Kaplan Meier survivorship. In this definition revision for any reason is included and revision is defined as "every re-operation of any parts of the prosthesis".

The ISAR benchmarking guidance document recommends the benchmark of 5% or less at 10 years for the superiority approach (please go to the document for details):

"When assigning a 10-year benchmark, both the non-inferiority and superiority approach should be used. Both approaches identify prostheses with quality outcomes. The superiority approach identifies a subset of benchmarked prostheses with the lowest rates of revision."

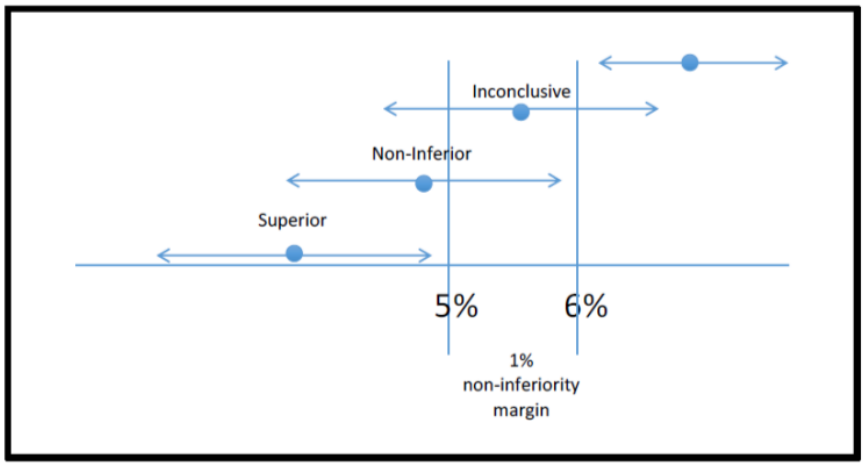
"For the 5-year benchmark this is a revision rate of 3.0% (3.6% for non-inferiority) and for 2-year benchmark 2% (2.4% for non-inferiority)." "All prostheses should receive an early benchmark if they have the appropriate number at risk at the specific benchmark time point and the revision rate does not exceed the benchmark standard for that early time point. Indication of not achieving the standard is if the lower CI exceeds the benchmark standard."

According to the www.nice.org.uk/guidance/ta304 the minimum number required for an assessment is 500 implants that have reached the nominated 5 or 10 year follow up. In the Netherlands LROI only reports on at least n=500. According to the benchmarking group of ISAR guidance 2018 the minimum number of a cohort is regarded 250 implants. As this number that contains the risk of too small sample size, it seems sensible and consensus to regard the number of 500 implants as a minimum sample size for assessment.

Other outcomes may be considered in the future for benchmarking, in particular "revision for specific reasons (e.g. dislocation, infection, fracture and early loosening), patient reported and functional outcomes as well as measures of surgical optimization, enhanced recovery, and prosthesis migration amongst others." However, "A standardized approach and an appropriate benchmark would need to be developed for each individual measure as these do not currently exist. There would also need to be a better understanding of the variation both within and between these different outcomes as well as determining how non-device and multiple extraneous factors may impact on their measurement." (ISAR document).

Additional information on benchmarking can be found in the publications by Deere (16) and Sayers (17) and on the ODEP website.

Figure 5.12.1: Application of Superiority, and Non-inferiority approaches to benchmarking at the 10-year benchmark assuming the benchmark level is 5%



4. How can different factors regarding patients, at short or long-term, be considered in survival analysis?

It is consensus that confounding factors (patient-related, surgeon-, surgery-, hospital specific factors can have an important influence on the outcome after arthroplasty.

Patient-related factors are taken into account most frequently either via restriction (e.g. inclusion of primary THAs for primary osteoarthritis only (18)), stratification (e.g. showing survival curves for men and women separately) or adjustment (hazard ratios adjusted for age and sex (18)).

For more detail, please see below.

5. Which rates can be analysed, e.g. reoperation, complications?

Reoperation rates and rates of major complications (specific causes) are other important outcomes although in need of data capture/ quality improvement. Main orthopaedic complications recorded in registries are infection, instability/dislocation, fracture, loosening. Several registries record any other reoperation in addition to revision (9). The recording of medical complications is most often not part of the registries' data collection. However, information on medical complications such as major bleeding, thromboembolic events, myocardial infarction or readmission rates is in some registries obtained via linkage to other databases (e.g. Sweden, Denmark). Complications can be related to the patient, the implant, the surgeon and the surgery (including perioperative management) or several of those factors together.

6. How can different competing risk factors be considered such as age, power calculation?

The term "competing risk factors" is unknown to us and might have been used here either instead of competing risks or competing events or instead of confounding factors. We have answered this question with respect to the latter (confounding factors). Patient-related factors assessed systematically and largely available in arthroplasty registries are laterality, diagnosis (or indication), age, sex, BMI and ASA score the latter as a measure of the patient's overall health status/comorbidity burden (19, 20). A good overview of patient related confounding factors is given in Podmore et al. (21) and Wright et al. (22) and in the ISAR recommendations (ISAR 2018).

The impact of patient related confounding factors is present in the short and the long term. A higher ASA score (>=3), a BMI >=35 and indication secondary OA were identified as important risk factors for short-term revision after THA (within 3 months postoperative) in multivariable analyses performed in two registries (23). Obesity is a risk factor for short-term revision also after knee arthroplasty (24).

Long-term revision rates are influenced by age, gender, indication, BMI and patient activity. Patient activity is a risk factor for loosening not routinely assessed in large cohorts/registries (25). However, age, gender and comorbidity burden partly account for its influence (26).

In observational studies confounders are taken into account either via restriction (e.g. inclusion of primary THAs for primary osteoarthritis only (18)), stratification (e.g. showing survival curves for men and women separately) or adjustment (hazard ratios adjusted for age and sex (18)). Considering several factors simultaneously requires a higher number of events (and larger sample size).

For benchmarking the following is recommended by ISAR (Benchmarking document): "Benchmarking assessment and setting of benchmarking standards should only be undertaken using data from primary hip and primary knee prostheses used for the management of osteoarthritis."

"Benchmark standards must be attained for a prosthesis regardless of the specific characteristics of the patients who received that prosthesis. The characteristics of the patients receiving the prosthesis should be summarized and the revision rates should be presented by (at least) age and gender. Direct adjustment is not applicable to benchmarking."

And "It is recommended that an individual prosthesis will be benchmarked against the benchmark standard of the clinical class it belongs to rather than subgroups based on patient and procedure characteristics. When applying for benchmark it will be possible to compare the characteristics of the prosthesis against the standard characteristics of patients in that clinical class. It is recommended that the benchmark application should provide revision rates for the prosthesis stratified by characteristics such as age group and gender. This will provide information about the performance of the implant under specific sets of circumstances and will help to understand why a particular prosthesis did not reach benchmark."

7. References

1. Malchau H, Garellick G, Berry D, Harris WH, Robertson O, Kärrlholm J et al. **Arthroplasty implant registries over the past five decades: Development, current, and future impact. Journal of Orthopaedic Research® 2018; 36(9):2319–30.**

2. Nab L, Groenwold RHH, Welsing PMJ, van Smeden M. **Measurement error in continuous endpoints in randomised trials: Problems and solutions. Statistics in Medicine 2019; 38(27):5182–96.**

3. Zuidgeest MGP, Goetz I, Groenwold RHH, Irving E, van Thiel GJMW, Grobbee DE. **Series: Pragmatic trials and real world evidence: Paper 1. Introduction. Journal of Clinical Epidemiology 2017; 88:7–13. Available from: URL: https://www.sciencedirect.com/science/article/pii/s0895435617305401.**

4. Hilton Boon M, Craig P, Thomson H, Campbell M, Moore L. **Regression Discontinuity Designs in Health: A Systematic Review. Epidemiology 2021; 32(1):87–93.**

5. Maciejewski ML, Basu A. **Regression Discontinuity Design. JAMA 2020; 324(4):381–2.**

6. Lauer MS, D'Agostino RB. **The randomized registry trial--the next disruptive technology in clinical research? N Engl J Med 2013; 369(17):1579–81.**

7. Liebs TR, Splietker F, Hassenpflug J. **Is a Revision a Revision? An Analysis of National Arthroplasty Registries' Definitions of Revision. Clin Orthop Relat Res 2015; 473(11):3421–30. Available from: URL: https://link.springer.com/article/10.1007/s11999-015-4255-4.**

8. Lübbeke A, Silman AJ, Prieto-Alhambra D, Adler AI, Barea C, Carr AJ. **The role of national registries in improving patient safety for hip and knee replacements. BMC Musculoskelet Disord 2017; 18(1):414. Available from: URL: https://bmcmusculoskeletdisord.biomedcentral.com/articles/10.1186/s12891-017-1773-0.**

9. Lacny S, Bohm E, Hawker G, Powell J, Marshall DA. **Assessing the comparability of hip arthroplasty registries in order to improve the recording and monitoring of outcome. The Bone & Joint Journal 2016; 98-B(4):442–51.**

10. Pocock SJ, Clayton TC, Altman DG. **Survival plots of time-to-event outcomes in clinical trials: good practice and pitfalls. The Lancet 2002; 359(9318):1686–9. Available from: URL: https://www.sciencedirect.com/science/article/pii/s014067360208594x.**

11. Malchau H. **On the importance of stepwise introduction of new hip implant technology : assessment of total hip replacement using clinical evaluation, radiostereometry, digitised radiography and a national hip registry 1995. Available from: URL: https://gupea.ub.gu.se/handle/2077/12932.**

12. Malchau H, Bragdon CR, Muratoglu OK. **The stepwise introduction of innovation into orthopedic surgery: the next level of dilemmas. The Journal of Arthroplasty 2011; 26(6):825–31. Available from: URL: https://www.sciencedirect.com/science/article/pii/s0883540310004730.**

13. McCulloch P, Altman DG, Campbell WB, Flum DR, Glasziou P, Marshall JC et al. **No surgical innovation without evaluation: the IDEAL recommendations. The Lancet 2009; 374(9695):1105–12. Available from: URL: https://www.sciencedirect.com/science/article/pii/s0140673609611168.**

14. Sedrakyan A, Campbell B, Merino JG, Kuntz R, Hirst A, McCulloch P. **IDEAL-D: a rational framework for evaluating and regulating the use of medical devices. BMJ 2016; 353:i2372.**

15. Kandala N-B, Connock M, Pulikottil-Jacob R, Sutcliffe P, Crowther MJ, Grove A et al. **Setting benchmark revision rates for total hip replacement: analysis of registry evidence. BMJ 2015; 350:h756. Available from: URL: https://www.bmj.com/content/350/bmj.h756.short.**

16. Deere KC, Whitehouse MR, Porter M, Blom AW, Sayers A. **Assessing the non-inferiority of prosthesis constructs used in hip replacement using data from the National Joint Registry of England, Wales, Northern Ireland and the Isle of Man: a benchmarking study. BMJ Open 2019; 9(4):e026685.**

17. Sayers A, Crowther MJ, Judge A, Whitehouse MR, Blom AW. **Determining the sample size required to establish whether a medical device is non-inferior to an external benchmark. BMJ Open 2017; 7(8):e015397**.

18. Australian Orthopaedic Association National Joint Replacement Registry (AOANJRR). **Hip, Knee & Shoulder Arthroplasty: 2020 Annual Report. Adelaide; AOA, 2020:; 2020:1-474. Available from: URL: [Accessed from: https://aoanjrr.sahmri.com/annual-reports-2020].**

19. Lübbeke A, Silman AJ, Barea C, Prieto-Alhambra D, Carr AJ. **Mapping existing hip and knee replacement registries in Europe. Health Policy 2018; 122(5):548–57. Available from: URL: https://www.sciencedirect.com/science/article/pii/s0168851018300629.**

20. Rolfson O, Rothwell A, Sedrakyan A, Chenok KE, Bohm E, Bozic KJ et al. **Use of patient-reported outcomes in the context of different levels of data. J Bone Joint Surg Am 2011; 93 Suppl 3(Supplement_3):66–71.**

21. Podmore B, Hutchings A, van der Meulen J, Aggarwal A, Konan S. **Impact of comorbid conditions on outcomes of hip and knee replacement surgery: a systematic review and meta-analysis. BMJ Open 2018; 8(7):e021784.**

22. Wright EA, Katz JN, Baron JA, Wright RJ, Malchau H, Mahomed N et al. **Risk factors for revision of primary total hip replacement: results from a national case-control study. Arthritis Care & Research 2012; 64(12):1879–85.**

23. Ferguson RJ, Silman AJ, Combescure C, Bulow E, Odin D, Hannouche D et al. **ASA class is associated with early revision and reoperation after total hip arthroplasty: an analysis of the Geneva and Swedish Hip Arthroplasty Registries. Acta Orthopaedica 2019; 90(4):324–30.**

24. Onggo JR, Onggo JD, Steiger R de, Hau R. **Greater risks of complications, infections, and revisions in the obese versus non-obese total hip arthroplasty population of 2,190,824 patients: a meta-analysis and systematic review. Osteoarthritis and Cartilage 2020; 28(1):31–44. Available from: URL: https://www.sciencedirect.com/science/article/pii/s1063458419312361.**

25. Münger P, Röder C, Ackermann-Liebrich U, Busato A. **Patient-related risk factors leading to aseptic stem loosening in total hip arthroplasty: a case-control study of 5,035 patients. Acta Orthopaedica 2006; 77(4):567–74.**

26. Flugsrud GB, Nordsletten L, Espehaug B, Havelin LI, Meyer HE. **The effect of middle-age body weight and physical activity on the risk of early revision hip arthroplasty: a cohort study of 1,535 individuals. Acta Orthopaedica 2007; 78(1):99–107. accessed 19 May 2021**

Endoprothesenregister Deutschland (EPRD) Jahresbericht 2019. https://www.eprd.de/de/ueber-uns/aktuelles/artikel/eprd-jahresbericht-2019/ accessed 19 May 2021

International Society of Arthroplasty Registries (ISAR). International Prosthesis Benchmarking Working Group guidance document. May 2018. https://www.isarhome.org/publications, accessed 19 May 2021

NICE Technology appraisal guidance. Total hip replacement and resurfacing arthroplasty for end-stage arthritis of the hip. Feb 2014. www.nice.org.uk/guidance/ta304. Accessed 19 May 2021

Orthopaedic Data Evaluation Panel (ODEP). https://www.odep.org.uk/ accessed 19 May 2021

REGULATION (EU) 2017/745 OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL of 5 April 2017 on medical devices, amending Directive 2001/83/EC, Regulation (EC) No 178/2002 and Regulation (EC) No 1223/2009 and repealing Council Directives 90/385/EEC and 93/42/EEC

RESEARCH TOPIC

PMCF STUDIES

| Research Topic 29

Potentials and Limitations of PMCF Studies in the Field of Total Joint Arthroplasty – Discussing Study Design, Parameters, and Alternatives

RESEARCH TOPIC 29

Potentials and Limitations of PMCF Studies in the Field of Total Joint Arthroplasty – Discussing Study Design, Parameters, and Alternatives

Authors

Anne Lübbeke[1,2], Hassan Achakri[3], Dario Bergadano[4], Joshua Bridgens[5], Ina Wüstefeld[6], Marcus Jäger[7,8], Ricardo Larrainzar-Garijo[9], Henning Windhagen[10], Philippe Massin[11], Eduardo Garcia-Rey[12,13]

[1]Division of Orthopaedics and Trauma Surgery, Geneva University Hospitals, Geneva, Switzerland
[2]Nuffield Department of Orthopaedics, Rheumatology and Musculoskeletal Sciences, University of Oxford, Oxford, UK
[3]Zimmer Biomet, Clinical Affairs, Zug - Switzerland
[4]Medacta International SA, Castel San Pietro, Switzerland
[5]DePuy Synthes, Leeds, United Kingdom
[6] Aesculap AG Medical Scientific Affairs, Tuttlingen, Germany
[7]Department of Orthopedics, Trauma & Reconstructive Surgery, St. Marien Hospital Mülheim an der Ruhr, Mülheim, Germany
[8]Orthopaedics and Trauma Surgery, University of Duisburg-Essen, Essen, Germany
[9]Orthopaedics and Trauma Department, Medicine School, Universidad Complutense, Hospital Universitario Infanta Leonor, Madrid, Spain
[10]Department of Orthopaedic Surgery, Medical School Hannover, Hannover, Germany
[11]CMC Ambroise-Paré-Hartmann, Neuilly-sur-Seine, France
[12]Hôpital Universitario La Paz-Idi Paz, Madrid, Spain
[13]Centro de Investigación Biomédica en Red de Bioingeniería, Biomateriales y Nanomedicina, CIBER-BBN, Madrid, Spain

1. Abstract

Question: PMCF Studies

1. What are potentials and limitations of PMCF studies in the field of arthroplasty?
2. Which study design specifications are adequate for a PMCF study?
3. Which aspects can actually be assessed and which must be assessed (documentation of all possible surgical approaches, indications, patient characteristics etc.)?
4. Are alternative methods to traditional clinical studies imaginable and is it somehow possible to generate additional data fast and with low effort?

Summary / Recommendation:

PMCF studies must resolve unanswered questions at time of regulatory approval regarding clinical benefit throughout expected lifetime, safety under wide-spread use, and generalisability of pre-market findings and must ensure continued acceptability of the benefit-risk ratio. Under the EU MDR post-market surveillance is proactive, continuous and involves comparison to a clinically meaningful comparator group as well as use of clinically relevant endpoints (risks and benefits). Main performance indicator after arthroplasty is all-cause revision. Measures of benefit include patient-reported outcomes. Current regulatory documents do not provide precise indications regarding design of PMCF studies in arthroplasty. Traditional follow-up studies may not always be sufficient to serve as PMCF because of the need for large sample sizes, long follow-up and real-world results. Registries/large observational population-based cohorts are unanimously identified and recommended in the literature as current best method to generate faster, more efficient high-quality post market clinical evidence for legacy and new devices. Most importantly registries are newly recognized as source for PMCF by the regulators. Randomized trials may advantageously be nested in registries. RSA, EBRA and other similar validated algorithms are good surrogate markers for later fixation failure used in early clinical evaluation, which would benefit from validation against clinically relevant long-term outcomes in PMCF.

2. Level of Evidence

N/A

3. Consensus Delegate Vote

95% – unanimous, strongest consensus (95% agree / 0% disagree / 5% abstain)

4. Graphical Abstract

<table>
<tr><th></th><th>Legacy Devices</th><th>New Devices
with Equivalence</th><th>New Novel Devices</th></tr>
<tr><td rowspan="2">Pre-Market</td><td rowspan="2"></td><td rowspan="2">Clinical Evaluation</td><td>Clinical Investigation
(RSA, EBRA, etc.)
(short-term)</td></tr>
<tr><td>Clinical Evaluation</td></tr>
<tr><td rowspan="3">Post-Market</td><td>Clinical Evaluation</td><td>PMCF study
(mid-term)</td><td>PMCF study
(mid-term)</td></tr>
<tr><td>Vigilance</td><td>Vigilance</td><td>Vigilance</td></tr>
<tr><td>Registries
&
Systematic Literature</td><td>Registries
&
Systematic Literature</td><td>Registries
&
Systematic Literature</td></tr>
</table>

Adapted from: H. Achakri. Surveillance and vigilance: What can manufacturers contribute? EFORT Implant & Patient Safety Initiative Inauguration Workshop, Brussels, 2020

5. Search Strategy

Published literature

Using key words: post market clinical studies AND arthroplasty; post market clinical studies AND medical device; post market studies AND arthroplasty; post market studies AND medical device; post market surveillance AND arthroplasty; post market surveillance AND medical device

Using the MeSH term: Product surveillance post market AND medical device; Product surveillance post market AND arthroplasty

Search was limited to reviews, to year of publication 2011-2021, and to publications in English. We did not include published literature on PMCF studies solely relevant to a particular regulatory environment (e.g. US, China).

Additional topic-relevant publications not identified through the key word/mesh term search were included during the review and writing process.

Regulatory documents

Current (April 2021) EU commission's medical device related regulatory documents, explanatory documentation from the Medical Device Coordination Group Document (MDCG) and documents from the International Medical Device Regulator Forum were reviewed.

6. Rationale

Post market clinical follow-up (PMCF) studies: Review of literature and regulatory documents

Pre-market clinical studies often lack power for rare outcomes, enrol relatively restricted populations and provide little to no information on long-term adverse effects (Nelson 2011). The role of post market clinical follow studies is to fill these and other gaps. However, post market surveillance of medical devices has not been up to this task (Resnic 2012). There is a paucity of completed well-conducted medical device PMCF studies in the US (Rathi 2015), and in Europe so far no public data are available on PMCF (Cipriano 2020). Hwang et al. reported a greater risk for post-marketing safety alerts and recalls for medical devices first approved in Europe compared to the US (Hwang 2016). Moreover, a survey of medical device manufacturers in Germany showed that the systematic conduct of PMCF studies was not an integral part of all manufacturers' product surveillance system yet (Zippel 2017).

As Nelson (2011) states "The lack of a single system … simultaneously prospective, population-based, and timely has led to gaps in post-licensure safety evidence, product recalls, and growing public concern about the health risks associated with licensed products." Consequences of the lack of PMCF were also reported by Nieuwenhuijse et al. (2014), who concluded: "We did not find convincing high quality evidence supporting the use of five substantial, well known, and already implemented device innovations in orthopaedics. Moreover, existing devices may be safer to use in total hip or knee replacement. Improved regulation and professional society oversight are necessary to prevent patients from being further exposed to these and future innovations introduced without proper evidence of improved clinical [performance] and safety."

To change this situation the European Medical Device Regulation enforces post market clinical follow-up (PMCF) under MEDDEV 2.12/2 rev.2 ANNEX XIV CLINICAL EVALUATION AND POST-MARKET CLINICAL FOLLOW-UP: "5. PMCF shall be understood to be a continuous process that updates the clinical evaluation referred to in Article 61 and Part A of this Annex and shall be addressed in the manufacturer's post-market surveillance plan. When conducting PMCF, the manufacturer shall proactively collect and evaluate clinical data from the use in or on humans of a device which bears the CE marking and is placed on the market or put into service within its intended purpose as referred to in the relevant conformity assessment procedure, with the aim of confirming the safety and performance throughout the expected lifetime of the device, of ensuring the continued acceptability of identified risks and of detecting emerging risks on the basis of factual evidence."

1. What are potentials and limitations of PMCF studies in the field of arthroplasty?

The purpose/potential of PMCF studies is (1) to resolve unanswered questions at time of regulatory approval regarding clinical benefit of the device throughout its expected lifetime, safety under wide-spread use, and generalisability of findings, (2) to ensure the continued acceptability of the benefit-risk ratio, and (3) to identify possible systematic misuse or off-label use of the device as well as emerging risks on the basis of factual evidence (REGULATION (EU) 2017/745; MDCG 2020-7).

Post-market surveillance under the new EU MDR will function as a continuous proactive safety evaluation of medical devices in contrast to the past passive-reactive system (Pane 2019). Annex X of the EU MDR states that "The clinical evaluation and its documentation must be actively updated with data obtained from the PMS. Where a postmarketing clinical follow-up as part of the PMS plan for the device is not deemed necessary, this may be duly justified and documented."

A major problem in post market surveillance in the past, the lack of precise device identification, is now going to be solved through the introduction of the unique device identifier (UDI), which will facilitate PMCF and improve its quality (Resnic 2012; Pane 2017). Finally, post market study conduct and methodology is much more developed in medication and vaccine evaluation. The medical device field can learn from their experience and adopt certain aspects (Pane 2017).

2. Which study design specifications are adequate for a PMCF study?

The current regulatory documents do not provide precise indications as to the design of PMCF studies in the field of arthroplasty. In the published literature there is agreement that in the field of arthroplasty conventional non-registry cohort studies alone are not sufficient to serve as PMCF because of the need for large sample sizes, very long follow-up and real-world results (Pabinger 2015, Lübbeke 2017). The post market clinical follow-up (PMCF) plan should include both post market studies and registries (MDCG 2020-6, IMDRF, Pane 2017). And "The MDR specifies that data derived from independent medical device registries can be submitted by a manufacturer to its notified body (MDR Annex VII, 4.11.h)", (Fraser 2020).

With respect to study design the International Medical Device Regulator Forum (IMDRF) document on PMCF studies states that "PMCF studies can be designed to collect data from routine use in clinical practice. Such study designs range from practical/pragmatic investigations to various types of observational studies, including cross sectional study, cohort study, case-control study." And "The study should be designed to address the objective(s) of the study. The PMCF study can take several forms, for example: the extended follow-up of patients enrolled in premarket investigations; a new post market clinical study; a review of data derived from a device registry; or a review of relevant retrospective data from patients previously exposed to the device."

The MDCG 2020-6 document indicates: "While controlled clinical investigations might be the preferred method for collecting clinical data as part of the PMCF studies for some products, there are other possibilities to gather relevant clinical data in the field" ...such as ... "systematic reviews of clinical data published in the literature, evaluation of results from PMCF studies such as clinically relevant scientifically sound questionnaires or registries". Scientifically sound studies include a "clearly stated research question(s), objective(s) and related endpoints; an evaluation of potential sources of bias or study distortion, and the impact of these factors on the potential validity of results; Design with an appropriate rationale and statistical analysis plan; plan for an analysis of the data and for drawing appropriate conclusion(s)."

The document also includes the following "suggested hierarchy of clinical evidence (full cycle):

- Results of high quality clinical investigations covering all device variants, indications, patient populations, duration of treatment effect, etc.
- Results of high quality clinical investigations with some gaps
- Outcomes from high quality clinical data collection systems such as registries
- Outcomes from studies with potential methodological flaws but where data can still be quantified and acceptability justified
- Equivalence data (reliable/quantifiable)
- Evaluation of state of the art, including evaluation of clinical data from similar devices
- Complaints and vigilance data; curated data
- Proactive PMS data, such as that derived from surveys
- Individual case reports on the subject device
- Compliance to non-clinical elements of common specifications considered relevant to device safety and performance
- Simulated use/animal/cadaveric testing involving healthcare professionals or other end users
- Pre-clinical and bench testing/compliance to standards"

The MDCG 2020-7 document insists on the appropriateness of the chosen methods, including the justification for (1) sample size, timescales and endpoints; (2) the comparator, on the basis of intended purpose and state of the art; and (3) study design on the basis of all of the above, and why it is sufficient to ensure representative patient populations. Moreover, are required an adequate control of sources of bias, a statistical justification for the expected quality of outcomes, and a justification for why this is satisfactory in light of the residual risks.

Regarding the published literature the Lancet paper "Generating comparative effectiveness on new drugs and devices after approval" (Cipriani 2020) indicates: "Post-marketing studies should be designed hierarchically: priority should be given to efforts aimed at evaluating a product's net clinical benefit in randomised trials compared with current known effective therapy, whenever possible. Use of non-randomised studies for the evaluation of clinical benefit in post-marketing period should be limited to instances of large magnitude of effect or when it is possible to reasonably infer the comparative benefits or risks in settings, in which RCT is not feasible. Non-randomised study designs have a clear role for the post-market evaluation of safety, especially for rare or uncommon adverse effects. Non-randomised studies could also be used to evaluate whether drugs and devices with an optimal evidence package can be extended to populations outside of those included in RCTs."

Similarly Fraser et al. (2020) request larger RCTs, where possible, with longer follow-up. RCTs should be nested in registries. This has the potential to combine the advantages of both study designs (observational and experimental) and to facilitate the conduct of multi-centre trials with reduced duration and costs (Lauer 2013, Sedrakyan 2016). Furthermore, independent medical device registries that are systematic and comprehensive should become the new standard in PMCF, in particular in safety assessment (Fraser 2020), and all stakeholders should contribute to and benefit from them (Lübbeke 2019). The new proactive safety evaluation of medical devices of the EU MDR needs large prospectively collected observational data for active adverse event signal detection, refinement and confirmation (Nelson 2011).

Comparative effectiveness evaluation is a crucial element of PMCF to continuously ensure high safety and performance levels of new and legacy implants. The EU MDR indicates (Article 71, 3d) that member states shall examine in particular ..."the reliability and robustness of the data generated in the clinical investigation, taking account of statistical approaches, design of the investigation and methodological aspects, including sample size, comparator and endpoints;". Cipriani et al. also highlight the importance of conducting comparative studies and of the choice of active clinically meaningful comparator group(s). In contrast to a traditional PMCF study, restricted in sample size, follow-up time and generalisability, a registry records data from multiple implants simultaneously and over a very long time. It contains thus easily available information regarding comparator groups both current and from the past as shown for knee prostheses by Lewis et al. (2020). A discussion of superiority/non-inferiority is beyond the scope of this task.

PMCF studies must use clinically relevant endpoints/outcomes, which are to be preferred over surrogate measures (Cipriani 2020). For the latter a simple correlation with clinical outcomes is not enough, the change in surrogate measure should also reliably predict change in clinical outcome at the individual and aggregate level. A systematic review of surrogate markers for long-term outcomes after hip

arthroplasty (Malak 2016) identified two validated surrogate markers of long-term primary THA outcome: RSA and EBRA, each measuring implant migration and wear. In the field of arthroplasty surrogate measures have been advocated for use in the pre-market clinical phase (Nelissen 2011, Malchau 2011). In instances where possible (traditional/registry-nested studies) they should be included in PMCF since it is there where their usefulness can be validated against clinically relevant mid- and long-term outcomes.

3. Which aspects can actually be assessed and which must be assessed (documentation of all possible surgical approaches, indications, patient characteristics etc.)?

First, unambiguous implant identification is a prerequisite for PMCF, which will be much facilitated through the unique device identifier. Second, under the new EU MDR both risks and benefits of medical devices must be assessed in PMCF studies.

Risks defined as a "combination of the probability of occurrence of harm and the severity of that harm" include: "3.3 Serious Adverse Event (SAE): Any adverse event that led to any of the following: a) death, b) serious deterioration in the health of the subject, that resulted in any of the following: i. life-threatening illness or injury, ii. permanent impairment of a body structure or a body function, iii. hospitalisation or prolongation of patient hospitalisation, iv. medical or surgical intervention to prevent life-threatening illness or injury or permanent impairment to a body structure or a body function, v. chronic disease, c) foetal distress, foetal death or a congenital physical or mental impairment or birth defect (MDR Article 2(58))

3.4 Device deficiency: Any inadequacy in the identity, quality, durability, reliability, safety or performance of an investigational device, including malfunction, use errors or inadequacy in information supplied by the manufacturer." (MDCG-2020 10/1)

After arthroplasty, thromboembolic events and myocardial infarction are serious rare medical complications. Important orthopaedic complications treated with or without further surgery are instability/dislocation, infection, loosening, and fracture.

Revision surgery is the most widely reported adverse outcome and performance indicator of an implant for the following reasons: "(1) it is the major indicator of an implant's quality; (2) it constitutes a substantial burden to the patient and the society (costs); and (3) it is a "hard endpoint" (reproducible, comparable)." (Lübbeke 2017). Revision should first of all be reported for each implant for all-causes combined. Reporting by main specific causes of revision is desirable in a second step (e.g. dislocation, deep infection, periprosthetic fracture, aseptic loosening for the hip). Categories should be standardized and harmonized as is already mostly the case in arthroplasty registries. A reoperation not requiring removal or total /partial exchange of implant is another relevant endpoint (Rolfson 2011).

A regulator-led initiative to develop a product specific risk based approach guidance to recommend the inclusion of only important risks that have an impact on the benefit-risk balance (based on ISO 14971) has been suggested (Pane 2019).

Regarding benefits according to the regulation (MEDDEV 2.7/1 revision 4) "Positive impacts of a device on the health of an individual should be meaningful (relevant for the patient) and measurable. The nature, extent, probability and duration of benefits should be considered. Benefits may include: positive impact on clinical outcome (such as reduced probability of adverse outcomes, e.g. mortality, morbidity; or improvement of impaired body function); the patient's quality of life (significant improvements, including by simplifying care or improving the clinical management of patients, improving body functions, providing relief from symptoms),..."

For arthroplasty benefits to be assessed are reduction of revision/reoperation/complications as well as indicators of symptoms, function, activity and quality of life through joint specific and generic patient-reported outcomes. There is detailed literature regarding use of patient-reported outcomes in the field of arthroplasty (Rolfson 2011, Rolfson 2016 Part I and II, Wilson 2019, OECD). Additional objective measures of body function and activity in arthroplasty are the topic of another task.

Patient-related factors which should be assessed systematically (and are largely available in arthroplasty registries) are laterality, age, sex, BMI and ASA score the latter as a measure of the patient's overall health status/comorbidity burden (Rolfson 2011, Lübbeke 2018). Moreover, patient activity is a risk factor for loosening not routinely assessed in large cohorts/registries (Münger 2006). However, age and comorbidity burden partly account for it (Flugsrud GB 2007). The indication for the surgery needs to be assessed systematically and in a harmonized way (see categories used in arthroplasty registries). And surgery-related factors to be assessed are surgical approach, fixation method, length of surgery, and prophylactic measures (Rolfson 2011). It should be indicated if the population studied was restricted (specialized facility) or general (national level). Implant-related factors/implant choice are crucial (Evans 2020), and are discussed in other tasks.

4. Are alternative methods to traditional clinical studies imaginable and is it somehow possible to generate additional data fast and with low effort?

Yes, there are alternatives to the conduct of traditional clinical studies.

First, the literature unanimously identifies registries/large observational population-based cohorts as the current best method to generate faster, more efficient high-quality post market clinical evidence for high-risk devices (Berger 2012, Cipriani 2020, Cnudde 2016, Franklin J 2019, Franklin P 2020, Fraser 2020, Lübbeke 2017, Nelissen 2011, Nelson 2011, Pabinger 2015, Pane 2017). Moreover, the concept of stepwise introduction of new prostheses (Malchau 1995) as well as the IDEAL framework for surgical innovations (McCulloch 2009) both identify registries as main tool in post-marketing surveillance. More recently, IDEAL has advocated that registry monitoring of high-risk medical devices should start from "first-in-human experience onward" (Sedrakyan 2016): "If prospective registries were started from the outset of clinical use they could be used continuously for safety surveillance of higher risk devices, picking up weak signals from rare events and outlier device reports as technology proliferates." Stepwise introduction of new implants has been called for repeatedly since the nineties (Malchau 2011, Nelissen 2011) but was largely ignored with well-known consequences (Reito 2014, Butler 2020). Finally, as a major step forward, registries are recognized in the European legislation for the first time as source for post market surveillance (EU MDR, Melvin 2019).

Registries monitor real-world treatment on a national level with a focus on long-term surveillance of implant performance (universally used endpoint revision) and increasingly patient-reported outcomes (Cnudde 2016). "Combining the data from the existing registries will increase the numbers of an implant or surgical technique possible to evaluate within a fixed time span. Combining data will also

allow testing the consistency of the findings by validating them in different populations and settings" (Lübbeke 2017). An increasing number of existing national/regional arthroplasty registries collect the main exposure, outcome and baseline variables prospectively (or enrich via linkage) and additional inter-registry harmonization efforts are under way (Lübbeke 2019). They also fulfil the requirement of comparative effectiveness evaluation (Berger 2012). High-quality registries provide detailed annual reports publicly available for all stakeholders. Finally, registries constitute both a health information system and an evaluation infrastructure that can host many types of study designs. Especially, they have a high potential to host nested RCTs and more data-rich multi-centre observational studies. This needs to be recognised, supported and further developed (Cipriani 2020, Franklin J 2020, Lauer 2013).

It is necessary to optimize registries for PMCF. Two IMDRF's Patient registry working group documents (IMDRF 2016 and 2018) provide in-depth information. Registries with a high coverage (over 90 or mandatory participation (IMDRF 2016)) should only be considered to avoid errors due to lost patients. Ideally, registers should be linked to civilian registers that may include patients who died from other causes. Further harmonization in implant identification (integration of UDIs), baseline and outcome variable collection and reporting of revision rates is absolutely required. Consistent reporting of cumulative failure (or survival) rates (and 95% CIs) at specific points in time as done in Australia, UK, Finland, Germany, Switzerland and others largely facilitates comparison (Foster 2019) and aggregation of results in the future. Due to insufficient utilization, the number of implantations in some countries may be not sufficient to detect rare side effects or identify heterogeneity of treatment effects. Thus, results of different national registries should be compared and/or aggregated to increase statistical power (Lübbeke 2017).

Second, real-world data other than registries/prospective cohorts can be a source for PMCF. The IMDRF states "Data generated from real-world clinical experience should be considered for PMCF studies. Examples of such data sources include: Patient generated health data; Device Registry; Health Record / Medical Record: Clinical data that are generated from routine clinical and medical practice and are maintained by professionals over time; Administrative data: Administrative data can include claims, health insurance data, and other sources; Survey Data collected by means of surveying healthcare professionals, customers and patients" (IMDRF 2020). Real-world data already contribute evidence to regulatory decisions on medication safety after approval (Franklin J 2020).

Data quality from real-world data other than registries is often inferior. Reasons are that they are collected for other purposes than post market implant surveillance, that data collection is not prospectively standardized, among others. "Relying on RWE for regulatory decisions also requires measures to ensure that analyses are completed and reported accurately. There have been several recent initiatives aimed at improving transparency in reporting of RWE database studies by encouraging standardized and comprehensive reporting of design and analytic choices." (Franklin J 2019). Better integration of data from electronic health care records into registries is highly desirable and under way (Franklin P 2020, Giori 2021).

Cipriani et al. (2020) state "Governments should directly support and facilitate the production of comparative post-marketing data by investing in the development of collaborative research networks and data systems that reduce the complexity, cost, and waste of rigorous post-marketing research efforts. Independent organisations should have a greater role in designing and running post-marketing trials, ideally leveraging funding from industry."

Finally, "Careful clinical follow-up and full reporting of any adverse events or complications related to high-risk medical devices are responsibilities that all physicians owe to their patients. With the development of independent registries for high-risk medical devices that are comprehensive and systematic, this should now become standard practice" (Fraser 2020).

7. References

1. Berger ML, Dreyer N, Anderson F, Towse A, Sedrakyan A, Normand S. **Prospective Observational Studies to Assess Comparative Effectiveness: The ISPOR Good Research Practices Task Force Report. Value in Health 2012;15(2):217–30**

2. Butler P, Gorgis J, Viberg B Overgaard S. **Low evidence for implementation of well documented implants regarding risk of early revision: a systematic review on total hip arthroplasty. EFORT Open Reviews 2021;6(1):3–8**

3. Callahan A, Fries JA, Re C, Huddleston 3rd JI, Giori N, Delp S, Shah NH. **Medical device surveillance with electronic health records. NPJ Digital Med 2019;2:94**

4. Cipriano A, Ioannidis JPA, Rothwell PM, Glasziou P, Tianjing L, Hernandez AF, Tomlinson A, Simes J, Naci H. **Generating comparative evidence on new drugs and devices after approval. Lancet 2020;395(10228):998–1010**

5. Cnudde P, Rolfson O, Nemes S, Kärrholm J, Rehnberg C, Rogmark C, Timperly J, Garellick G. **Linking Swedish health data registers to establish a research database and a shared decision-making tool in hip replacement. BMC Musculoskeletal Disord 2016;17(1):414**

6. Drozda JP, Roach J, Forsyth T, Helmering P, Dummitt B, Tcheng JE. **Constructing the informatics and information technology foundations of a medical device evaluation system: a report from the FDA unique device identifier demonstration. J Am Med Inform Assoc 2018;25(2):111–120**

7. Evans JT, Blom AW, Timperley AJ, Dieppe P, Wilson MJ, Sayers A, Whitehouse MR. **Factors associated with implant survival following total hip replacement surgery: A registry study of data from the National Joint Registry of England, Wales, Northern Ireland and the Isle of Man. PLoS Med 2020; 31;17(8):e1003291**

8. Flugsrud GB, Nordsletten L, Espehaug B, Havelin LI, Meyer HE. **The effect of middle-age body weight and physical activity on the risk of early revision hip arthroplasty: a cohort study of 1,535 individuals. Acta Orthop 2007;78(1):99–107**

9. Foster C, Posada C, Pack B, Hallstrom BR, Hughes RE. **Summary of knee implant one, three, five, and 10-year revision risk reported by national and regional arthroplasty registries: a valuable source of evidence for clinical decision-making. EFORT Open Reviews 2019; 5(5):268-272**

10. Franklin JM, Glynn RJ, Martin D, Schneeweiss S. **Evaluating the Use of Nonrandomized Real-World Data Analyses for Regulatory Decision Making. Clinical Pharmacology and Therapeutics 2019;105(4):867-877**

11. Franklin JM, Pawar A, Martin D, Glynn RJ, Levenson M, Temple R, Schneeweiss S. **Nonrandomized Real-World Evidence to Support Regulatory Decision Making: Process for a Randomized Trial Replication Project. Clin Pharmacol Ther 2020;107(4):817-826**

12. Franklin PD, Lurie J, Tosteson TD, Tosteson AN, **Task Force on Musculoskeletal Registries. Integration of Registries with EHRs to Accelerate Generation of Real-World Evidence for Clinical Practice and Learning Health Systems Research: Recommendations from a Workshop on Registry Best Practices. JBJS Am 2020; 102(19):e110**

13. Fraser AG, Butchart EG, Szymanski P, Caiani EG, Crosby S, Kearney P, Van de Werf F. **The need for transparency of clinical evidence for medical devices in Europe. Lancet 2018;392(10146):521-530**

14. Fraser AG, Byrne RA, Kautzner J, Butchart EG, Szymanski P, Leggeri I, de Boer RA, Caiani EG, Van de Werf F, Vardas PE, Badimon L. **Implementing the new European Regulations on medical devices—clinical responsibilities for evidence-based practice. European Heart Journal 2020;41(27):2589-2596**

15. Gagliardi AR, Umoquit M, Lehoux P, Ross S, Ducey A, Urbach DR. **Identifying optimal postmarket surveillance strategies for medical and surgical devices: implications for policy, practice and research. BMJ Qual Saf 2013;22:210–8**

16. Giori NJ, Radin J, Callahan A, Fries JA, Halilaj E, Ré C, et al. **Assessment of Extractability and Accuracy of Electronic Health Record Data for Joint Implant Registries. JAMA Netw Open 2021;4(3):e211728**

17. Hasan S, Marang-van de Mheen PJ, Kaptein BL, Nelissen RGHH, Pijls BG. **RSA-tested TKA Implants on Average Have Lower Mean 10-year Revision Rates Than Non-RSA-tested Designs. Clin Orthop 2020;478(6):1232-41**

18. Hwang TJ, Sokolov E, Franklin JM, Kesselheim AS. **Comparison of rates of safety issues and reporting of trial outcomes for medical devices approved in the European Union and United States: cohort study. BMJ. 2016;353:I3323**

19. Kramer DB, Tan YT, Sato C, Kesselheim AS. **Postmarket surveillance of medical devices: a comparison of strategies in the US, EU, Japan, and China. PLoS Med 2013;10:e1001519**

20. Kesselheim AS, Rajan PV. **Regulating incremental innovation in medical devices. BMJ 2014;349(sep09 5):g5303-g5303**

21. Kynaston-Pearson F, Ashmore AM; Malak TT, RombachI, Taylor A, Beard D, Arden NK, Price A, Prieto-Alhambra D, Judge A, Carr AJ, Glyn-Jones S. **Primary hip replacement prostheses and their evidence base: systematic review of literature. BMJ 2013;347:f6956**

22. Lauer MS, D'Agostino RB, Sr. **The randomized registry trial–the next disruptive technology in clinical research? N Engl J Med 2013;369(17):1579- 81**

23. Lewis PL, Graves SE, de Steiger RN, Campbell DG, Peng Y, Hatton A, et al. **Does Knee Prosthesis Survivorship Improve When Implant Designs Change? Findings from the Australian Orthopaedic Association National Joint Replacement Registry. Clin Orthop 2020;478(6):1156-72**

24. Lübbeke A, Silman AJ, Prieto-Alhambra D, Adler AI, Barea C, Carr AJ. **The role of national registries in improving patient safety for hip and knee replacements. BMC Musculoskeletal Disord 2017;18(1):414**

25. Lübbeke A, Silman AJ, Barea C, Prieto-Alhambra D, Carr AJ. **Mapping existing hip and knee replacement registries in Europe. Health Policy. 2018;122(5):548-557**

26. Lübbeke A, Carr AJ, Hoffmeyer P. **Registry stakeholders. EFORT Open Rev. 2019 3;4(6):330-336**

27. Malak TT, Broomfield JA, Palmer AJ, Hopewell S, Carr A, Brown C, Prieto-Alhambra D, Glyn-Jones S. **Surrogate markers of long-term outcome in primary total hip arthroplasty: A systematic review. Bone Joint Res 2016;5(6):206-14**

28. Malchau H. **On the importance of stepwise introduction of new hip implant technology. Assessment of total hip replacement using clinical evaluation, radiostereometry, digitised radiography and a national hip registry, in Department of Orthopaedics. Sweden: Goteborg: Goteborg University; 1995**

29. Malchau H, Bragdon CR, Muratoglu OK. **The stepwise introduction of innovation into orthopedic surgery: The next level of dilemmas. J Arthroplasty 2011;26(6):825-31**

30. McCulloch P, Altman D G, Campbell W B, Flum D R, Glasziou P, Marshall JC, Nicholl J, Balliol Collaboration. **No surgical innovation without evaluation: The IDEAL recommendations. Lancet 2009;374 (9695): 1105-12**

31. Melvin T, Torre M. **New medical device regulations: the regulator's view. EFORT Open Reviews 2019;4(6):351-356**

32. Münger P, Röder C, Ackermann-Liebrich U, Busato A. **Patient-related risk factors leading to aseptic stem loosening in total hip arthroplasty: A case-control study of 5,035 patients. Acta Orthop 2006; 77(4):567-74**

33. Nelissen RGHH, Pijls BG, Kärrholm J, Malchau H, Nieuwenhuijse MJ, Valstar ER. **RSA and Registries: The Quest for Phased Introduction of New Implants. JBJS Am 2011;93 Suppl 3:62-5**

34. Nelson JC, Cook AJ, YU O, Zhao S, Jackson LA, Psaty BM. **Methods for observational post-licensure medical product safety surveillance. Stat Methods Med Res 2015;24(2):177-93**

35. NICE Guidance. **Total hip replacement and resurfacing arthroplasty for end-stage arthritis of the hip. February 2014**

36. Nieuwenhuijse MJ, Nelissen RGHH, Schoones JW, Sedrakyan A. **Appraisal of evidence base for introduction of new implants in hip and knee replacement: a systematic review of five widely used device technologies. BMJ 2014;349:g5133**

37. OECD Patient-Reported Indicator Surveys. https://www.oecd.org/health/health-systems/OECD-PaRIS-hip-knee-data-collection-guidelines-en-web.pdf; accessed May 8, 2021

38. Pabinger C, Bridgens A, Berghold A, Wurzer P, Boehler N, Labek G. **Quality of outcome data in total hip arthroplasty: comparison of registry data and worldwide non-registry studies from 5 decades. Hip Int 2015;25(5):394-401**

39. Pabinger C, Lumenta DB, Cupak D, Berghold A, Boehler N, Labek G. **Quality of outcome data in knee arthroplasty. Acta Orthop 2015;86(1):58-62**

40. Pane J, Coloma PM, Verhamme KMC, Sturkenboom MCJM, Rebollo I. **Evaluating the Safety Profile of Non-Active Implantable Medical Devices Compared with Medicines. Drug Safety 2017;40(1):37-47**

41. Pane J, Francisca RDC Verhamme KMC, Orozco M, Viroux H, Rebollo I, Sturkenboom MCJM. **EU postmarket surveillance plans for medical devices. Pharmacoepidemiol Drug Saf 2019;28(9):1155-1165**

42. Pane J, Verhamme KMC, Shrum L, Rebollo I, Sturkenboom MCJM. **Blockchain technology applications to postmarket surveillance of medical devices. Expert Rev Med Devices 2020;17(10):1123-1132**

43. Rathi VK, Krumholz HM, Masoudi FA, Ross JS. **Characteristics of Clinical Studies Conducted Over the Total Product Life Cycle of High-Risk Therapeutic Medical Devices Receiving FDA Premarket Approval in 2010 and 2011. JAMA 2015**

44. Resnic FS, Normand S-LT. **Postmarketing Surveillance of Medical Devices — Filling in the Gaps. N Engl J Med 2012;366(10):875-7**

45. Reito A, Lehtovirta L, Lainiala O, Mäkelä K, Eskelinen A. **Lack of evidence—the anti-stepwise introduction of metal-on-metal hip replacements. Acta Orthop 2017; 88 (5): 478–483**

46. Rolfson O, Rothwell A, Sedrakyan A, Chenok KE, Bohm E, Bozic K, Garellick G. **Use of patient-reported outcomes in the context of different levels of data. JBJS Am 2011;93 Suppl 3(Suppl 3):66-71**

47. Rolfson O, Chenok K, Bohm E, Lübbeke A, Denissen G, Dunn J, Lyman S, Franklin P, Dunbar M, Overgaard S, Garellick G, Dawson J. **Patient-reported outcome measures in arthroplasty registers: Report of the Patient-Reported Outcome Measures Working Group of the International Society of Arthroplasty Registers. Part I. Overview and rationale for patient-reported outcome measures. Acta Orthop. 2016;87 Suppl 1:3-8**

48. Rolfson O, Bohm E, Franklin P, Lyman S, Denissen G, Dawson J, Dunn J, Chenok K, Dunbar M, Overgaard S, Garellick G, Lübbeke A. **Patient-reported outcome measures in arthroplasty registries: Report of the Patient- Reported Outcome Measures Working Group of the International Society of Arthroplasty Registries. Part II. Recommendations for selection, administration, and analysis. Acta Orthop. 2016;87 Suppl 1:9-23**

49. Sedrakyan A, Graves S, Bordini B, Pons M, Havelin L, Mehle S, et al. **Comparative Effectiveness of Ceramic-on-Ceramic Implants in Stemmed Hip Replacement: A Multinational Study of Six National and Regional Registries. J Bone Jt Surg 2014;96(Supplement_1):34-41**

50. Sedrakyan A, Campbell B, Merino JG, Kuntz R, Hirst A, McCulloch P. **IDEAL-D: a rational framework for evaluating and regulating the use of medical devices. BMJ 2016;353:i2372**

51. Wilson I, Franklin PD, Lübbeke A, Lyman S, Overgaard S, Rolfson O, W-Dahl A, Wilkinson M, Bohm E. **Orthopaedic Registries with Patient-Reported Outcome Measures. EFORT Open Reviews 2019;4(6):357-367**

52. Zannad F, Stough WG, Pina IL, Mehran R, Abraham WT, et al. **Current challenges for clinical trials of cardiovascular medical devices. Int J Cardiol 2014;175(1):30-7**

53. Zippel C, Bohnet-Joschko S. **Post market surveillance in the German medical device sector – current state and future perspectives. Health Policy 2017;121(8):880-886**

Regulatory documents

REGULATION (EU) 2017/745 OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL of 5 April 2017 on medical devices, amending Directive 2001/83/EC, Regulation (EC) No 178/2002 and Regulation (EC) No 1223/2009 and repealing Council Directives 90/385/EEC and 93/42/EEC

EU Commission MEDDEV 2.7/1 revision 4. CLINICAL EVALUATION: A GUIDE FOR MANUFACTURERS AND NOTIFIED BODIES UNDER DIRECTIVES 93/42/EEC and 90/385/EEC

IMDRF Patient Registries Working Group. Principles of International System of Registries Linked to Other Data Sources and Tools. 30 September 2016

IMDRF Patient Registries Working Group. Tools for Assessing the Usability of Registries in Support of Regulatory Decision-Making. 27 March 2018

IMDRF Medical Device Clinical Evaluation Working Group. Post-Market Clinical Follow-Up Studies. International Medical Device Regulator Forum (IMDRF). 7 August 2020

Medical Device Coordination Group Document MDCG 2020-6. Regulation (EU) 2017/745: Clinical evidence needed for medical devices previously CE marked under Directives 93/42/EEC or 90/385/EEC. A guide for manufacturers and notified bodies. April 2020

Medical Device Coordination Group Document MDCG 2020-7. Post-market clinical follow-up (PMCF) Plan Template. A guide for manufacturers and notified bodies. April 2020

Medical Device Coordination Group Document MDCG 2020-8. Post-market clinical follow-up (PMCF) Evaluation Report Template. A guide for manufacturers and notified bodies. April 2020

Medical Device Coordination Group Document MDCG 2020-10/1. Safety reporting in clinical investigations of medical devices under the Regulation (EU) 2017/745. May 2020

RESEARCH TOPIC

REGISTRY STUDIES

Research Topic 30

Potentials of Registry Studies in Total Joint Arthroplasty Considering Quality and Quantity of Data, Relevant Parameters and their Potential to Increase the Exploratory Power of Pre-CE Studies

RESEARCH TOPIC 30

Potentials of Registry Studies in Total Joint Arthroplasty Considering Quality and Quantity of Data, Relevant Parameters and their Potential to Increase the Exploratory Power of Pre-CE Studies

Authors

Volkmar Jansson[1], Ashley Blom[2,3], Barbara Bordini[4], Søren Overgaard[5,6,7,8], Rob G.H.H. Nelissen[9]

[1]Ludwig Maximilians University Munich, Department of Orthopaedic and Trauma Surgery, Musculoskeletal University Center Munich (MUM), Campus Grosshadern, Munich, Germany

[2]National Institute for Health Research Bristol Biomedical Research Centre, University Hospitals Bristol and Weston NHS Foundation Trust and the University of Bristol, Bristol, UK

[3]Translational Health Sciences, Bristol Medical School, Musculoskeletal Research Unit, University of Bristol, Learning & Research Building, Southmead Hospital, Bristol, UK

[4]Laboratorio di Tecnologia Medica, IRCCS Istituto Ortopedico Rizzoli, Bologna, Italy

[5]Department of Clinical Medicine, Copenhagen University, Copenhagen, Denmark

[6]Department of Orthopaedic Surgery and Traumatology, Bispebjerg and Frederiksberg Hospital, University of Copenhagen, Copenhagen, Denmark

[7]Orthopaedic Research Unit, Department of Orthopaedic Surgery and Traumatology, Odense University Hospital, Odense, Denmark

[8]Department of Clinical Research, University of Southern Denmark, Odense, Denmark

[9]Department of Orthopaedics, Leids Universitair Medisch Centrum, Leiden, Netherlands

1. Abstract

Question: Registry studies

1. What are the potentials of registry studies in the first field of arthroplasty?
2. What quality and what quantitative amounts of registry data is required to allow reliable answers to specific questions?
3. Which parameters can be assessed, also going beyond sheer product-related performance (e.g. "Mix&Match" compatibility between different manufacturers)?
4. Is it possible to derive tangible/practice-related information from registry data and benchmark values, with the potential to increase the exploratory power of pre-CE studies (e.g. for a reduction of case number)?

Summary / Recommendation:

Statement Question 1.:

- Registries contribute to evaluate implant safety like early revision due to a specific cause (eg. dislocation, fracture) and late failure (e.g. implant loosening) and thus post-market clinical follow-up (PMCF).
- Detection of outliers from the benchmark (implant and/or surgeons) with feedback to stakeholders.

Statement Questions 2. & 4.:

- Registries should report on their validity: coverage of patients, completeness (e.g. primary and revision surgery), implant classification and methodology.
- Randomised controlled trials, nested within registries, should be used in the early phase of new implant introduction, with focus on adverse effects and implant migration (e.g. RSA, EBRA, CT, see also EFORT Consensus Statements 17 and 18).

Statement Question 3.:

- Implants should preferably be evaluated as constructs due to their intricate relationship (e.g. hipstem and cups -and heads- are dependent variables).

Statement Question 4.:

- High Validity Registries have a clear description of coverage and completeness which should be 80-100% for primary and revision surgery in order to be used for benchmarking of joint replacing implants. But, outliers (i.e. high revision) from registries with lower validity are important signals to be analysed.

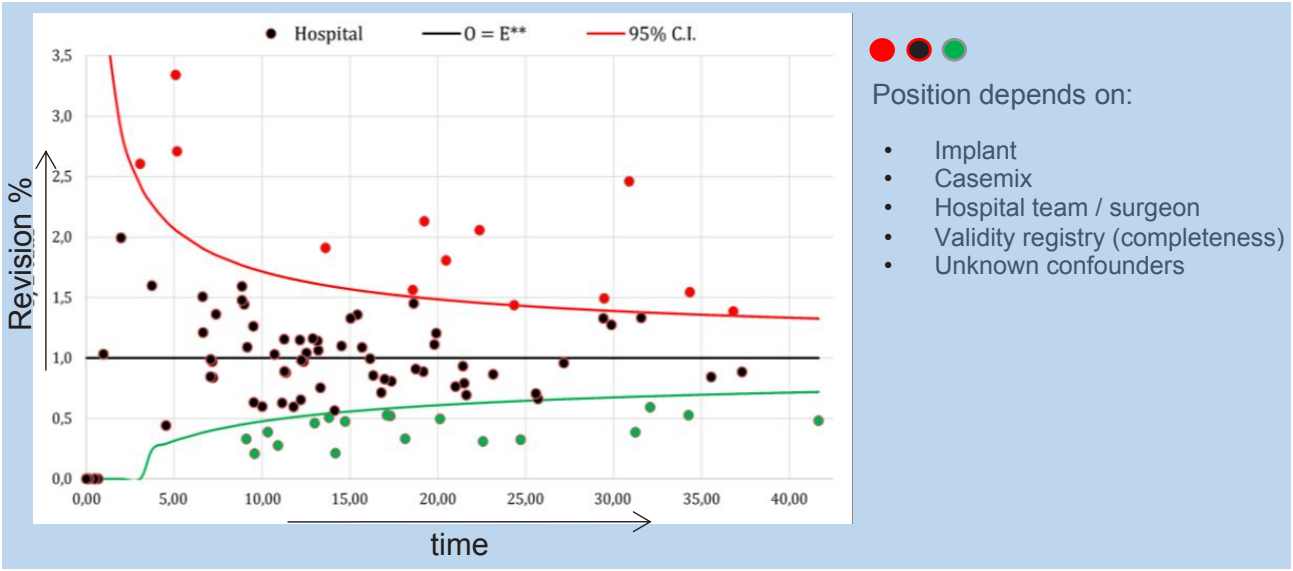
2. Level of Evidence

High

3. Consensus Delegate Vote

95% – unanimous, strongest consensus (95% agree / 2.5% disagree / 2.5% abstain)

4. Graphical Abstract



Benchmarking and effect confounders (green is excellent performance, red bad)

Legend Funnelplot 1.0 = benchmark with 95% confidence interval, higher revision than upper limit 95% CI = red, less than lower limit 95% CI, green dots

5. Search Strategy

Methodology of registry data and methods are well described, see reference list.

6. Rationale

The European Medicines Agency (EMA) defines registries as organised systems that use observational methods to collect uniform data on a population defined by a particular disease, condition, or exposure that is followed over time, with the aim to improve quality of patient care. The importance of registries is acknowledged within the MDR (Medical Device Regulations, article 108, Annex VII 4.11(g).

Studies and registries

Registries are observational cohorts and as such it is difficult to ascertain causal inference from them, however, due to the large amount of data they contain, they are extremely useful in showing association. This lack of causal inference underscores the importance of randomized controlled trials including nested RCT within registries or the use of surrogate randomization methods [1, 2, 3, 4, 6]. In studies and registry likewise, input data are crucial and should follow some general principles. FAIR principles of data: Findable, Accessible, Interoperable, Reusable. In order to compare across registries. Different registries use different input data.
It seems sensible to define minimal standards for collected data:

- Data collection
  - Medical Device: article/lot number or unique device identification (UDI).
  - Patient demographics age, gender, diagnosis, comorbidities (ASA), BMI
  - Outcome: mortality, revision, adverse event.
- Coverage and representativeness of collected data should be published.
- Completeness of primary and revision surgery data should be published [5].
- Data on completeness of all data fields should be readily accessible [5].
- Succinct definition of end-point of analysis: usually revision (all revisions or septic revision or non-septic revision).
- Definition of revision and reason for revision: any revision or only revision if a component (i.e. liner or stem) has been exchanged.

Additional information could be:

- Additional Data collection
  - Medical Device: feature characteristics of implant like material, surface characteristics, sizing and (Mix&Match) compatibility of components used.
  - Patient characteristics and habits like smoking, activity, social economic status (SES).
  - Details of the surgical procedure such as anaesthetic, surgical approach, auxiliary procedures, surgical team, robotics, navigation, bone graft, air flow condition.
  - Different end-point definitions: Function of implant like patient reported outcome measurements (PROMs) or VAS (visual analog scale) Pain etc.
  - Addition of radiographs both preoperative and postoperative (info bone adaptation to implants analysed with AI algorithms).

Number of required data

According to the HTA report [10] the minimum number required for an assessment is 500 implants that have reached the nominated 5 or 10 year follow up. In the Netherlands LROI only reports on at least n=500. According to the benchmarking group of ISAR guidance 2014 [13] the minimum number of a cohort is regarded 250 implants. As this number contains the risk of too small sample size, it is more sensible to regard the number of 500 implants as a minimum sample size for assessment of performance at a certain point in time for a non-inferiority analysis of an implant.

Definition across implants

Analysis of failure should be done preferably by construct components, since the "isolated" components are not independent variables if failure occurs. But poorly performing individual parts of the construct may be masked by other parts performing so well that the construct doesn't fail and vice versa. Analysis methods should be able to address this problem [7,8]. Registries should at least have a broad classification of the most important attributes such as fixation, constraint, bearing, modularity and sizing. Greater granularity of implant attributes is desirable as when the data set is large enough granularity allows interrogation of more implant variables.

Mix&match constructs should be analysed as such (see earlier).

Definitions of revision / failure analysis

A revision by definition is when an implant is removed and replaced by another implant, be it the entire construct or part of a construct. In the benchmarking group of ISAR guidance 2018 [13] it is suggested that revision for all reasons should be the standard

A re-operation includes revisions, but also includes any secondary procedure undertaken including added implant parts like a secondary retropatellar resufacing. From a patient perspective they both represent forms of failure and ideally registries should endeavour to capture both revision and all reoperations.

Continuous monitoring of real-time or at least frequent analysis of clinical outcome of the arthroplasty and alternative methods such as CUSUM (Cumulative SUM based surveillance system of clinical outcome) and VLAD (Variable Life Adjusted Display) scores may add value in early detection of failure if they are used in frequent feedback loops to surgical groups / surgeons.

Follow up time/end point evaluation

Follow-up time in terms of benchmarking, however, poses a more difficult problem in terms of referencing. The benchmarking group of ISAR guidance 2018 [13] suggests Kaplan Meier (KM) survivorship (1-KM = percentage revision) or Prosthesis Time Incidence Rate (PTIR) for referencing. Some registries use both Kaplan Meier (KM) survivorship as well as competing risk analyses in their report [9, 11, 12, 14-19]. Kaplan Meier (KM) methods are appropriate for describing implant failure (thus are used when comparing implants with each other), whereas crude survival estimation using competing risk methods estimates the risk of surgical revision as it depends on both implant failure and mortality. Competing risk is a good way of estimating how many implants will fail and is helpful in planning of healthcare delivery [27]. Both competing risk models and Kaplan-Meier methods are useful in arthroplasty, and both provide unbiased estimates of crude and net failure in the absence of any confounding or selection respectively. Lower estimates of failure from competing risk models may be misleading to surgeons who are attempting to select the best implant with the lowest failure rates for their patients. Although in older populations (80+) and at long-term follow-up (> 10 yrs.) mortality of the patient is competing with implant survival, thus comparison between implants may be misleading. In summary, take the premisses of the research question to be answered into account when using either KM or competing risk.

Benchmarking and number of data

As for benchmarking implants several methods have been proposed. It should be acknowledged that performance of the implant as such has an intricate relationship with the surgical procedure and the patient. Revision rates can be influenced by the surgeon's threshold for revision, the patient's desire for revision and the ease of revision. These inherent sources of bias need to be considered when interpreting benchmarking guidelines. Some sources of bias such as individual surgeon threshold for revision are unlikely to have an effect at a population level, but others such as ease of revision will. Several implant benchmarking guidelines have been suggested, using heuristic and somewhat arbitrary levels such as 95% implant survival at 10-year follow-up. However, it is not clear how or if implants should be compared to a current "best performing" or benchmark implant (see graphical abstract). The new implant should be at least non-inferior to the benchmark. Theoretical model simulation studies show that at least 6000 implants (90% study power) are needed to proof non-inferiority to the benchmark [20]. Thus, registry data can be used for benchmarking, although some discussion exists on the uncertainty margins around mean survival or 1-survival (i.e. revision), number of registries or even including cohort studies exist [21, 22, 23, 24, 25]. Although 10 years gives great certainty on non-inferiority of a "new" implant, earlier time intervals on benchmarking are necessary for new implants as deviation from acceptable outcomes may be evident at earlier time points [25, 13, 26]. Both lower patient numbers and shorter time intervals, but with detailed information on failure modes like dislocation, instability, implant malfunction, extended operation time caused by poor instruments and revision for surgical reasons like bleeding or infection have been proposed [26]. Such studies should be performed within a. National registry as nested trials. A prior condition to use registry data for this very short-time evaluation like 3 or 6 months is that the registry evaluation process is corrected for evaluation time. Such short-term evaluation of registry data can contribute to pre CE-studies.

Definition of outcome of analysis determines comparability between different registries. The majority of registries use the Kaplan-Meier (KM) survival analysis, some use competing risk, which is comparable to the KM analysis, with the aim to account for a larger number of deaths which could "compete" with presence of "revision" and thus effect the KM estimator [9].

Lost to follow up ratio

In the benchmarking group of ISAR guidance 2014 [13] the problem of lost to follow up is considered. Especially death rate can become a problem if not monitored precisely. This applies especially if not time referenced benchmarking is performed at short-term.

Missing data, either individual items or complete records can be problematic, especially if data is not missing at random. The problem of some surgeons failing to register their revisions may contribute to underestimation of failure [12]. Registries should thus have universal population coverage and be mandated by law

Regular audit of source data will deal with complete missingness (lost to follow-up) and there are statistical methods such as data imputation that deal with item missingness. Comparing registry data with other secondary source data such as insurance databases is not the ideal standard, but may be a less labour intensive form of audit. The ideal is to compare with primary source data in each hospital.

"Real-time" analysis

If product library allows for implant compatibility evaluation, mismatch of used implants can be reflected to hospitals by the time of data transfer to the registry.

7. References

1. Nab L, Groenwold RHH, Welsing PMJ, van Smeden M. **Measurement error in continuous endpoints in randomised trials: Problems and solutions. Stat Med. 2019 Nov 30;38(27):5182-5196. Doi: 10.1002/sim.8359. Epub 2019 Sep 2.**

2. Zuidgeest MGP, Goetz I, Groenwold RHH, Irving E, van Thiel GJMW, Grobbee DE; **GetReal Work Package 3 Series: Pragmatic trials and real world evidence: Paper 1. Introduction. J Clin Epidemiol. 2017 Aug;88:7-13. Doi: 0.1016/j.jclinepi.2016.12.023.**

3. Hilton Boon M, Craig P, Thomson H, Campbell M, Moore L.Hilton Boon M, et al. **Regression Discontinuity Designs in Health: A Systematic Review. Epidemiology. 2021 Jan;32(1):87-93. Doi: 10.1097/EDE.0000000000001274.Epidemiology. 2021. PMID: 33196561 Free PMC article.**

4. Maciejewski ML, Basu A.Maciejewski ML, et al. **Regression Discontinuity Design. JAMA. 2020 Jul 28;324(4):381-382. Doi: 10.1001/jama.2020.3822.**

5. van Steenbergen LN, Denissen GA, Spooren A, van Rooden SM, van Oosterhout FJ, Morrenhof JW, Nelissen RG.van Steenbergen LN, et al. **More than 95% completeness of reported procedures in the population-based Dutch Arthroplasty Register. Acta Orthop. 2015;86(4):498-505. Doi: 10.3109/17453674.2015.1028307. Epub 2015 Mar 11.**

6. Nelissen RG, Brand R, Rozing PM. **Survivorship analysis in total condylar knee arthroplasty. A statistical review. J Bone Joint Surg Am. 1992 Mar;74(3):383-9.**

7. Denissen GAW, van Steenbergen LN, Lollinga WT, Verdonschot NJJ, Schreurs BW, Nelissen RGHH. **Generic implant classification enables comparison across implant designs: the Dutch Arthroplasty Register implant library. EFORT Open Rev. 2019 Jun 3;4(6):344-350. Doi: 10.1302/2058-5241.4.180063. eCollection 2019 Jun.**

8. Spekenbrink-Spooren A, Van Steenbergen LN, Denissen GAW, Swierstra BA, Poolman RW, Nelissen RGHH. **Higher mid-term revision rates of posterior stabilized compared with cruciate retaining total knee arthroplasties: 133,841 cemented arthroplasties for osteoarthritis in the Netherlands in 2007-2016. Acta Orthop. 2018 Dec;89(6):640-645. Doi: 10.1080/17453674.2018.1518570. Epub 2018 Oct 23.**

9. Van Der Pas S, Nelissen R, Fiocco M. **Different competing risks models for different questions may give similar results in arthroplasty registers in the presence of few events. Acta Orthop. 2018 Apr;89(2):145-151. Doi: 10.1080/17453674.2018.1427314. Epub 2018 Feb 1.**

10. HTA report. www.nice.org.uk/guidance/ta304)

11. EPRD Jahresbericht 2019 www.eprd.de/de/ueber-uns/aktuelles/artikel/eprd-jahresbericht-2019/

12. LROI annual report 2020 www,LROI.nl/reportage

13. ISAR Report by International Prosthesis Benchmarking Working Group May 2018. https ://www.isarhome.org/publications

14. Kaplan EL, Meier P (1958) **Nonparametric estimation from incomplete observations. J Am Stat Assoc 53:457–481**

15. Fennema P, Lubsen J. **Survival analysis in total joint replace ment: an alternative method of accounting for the presence of com peting risk. J Bone Joint Surg Br 2010, 92:701–706**

16. Putter H, Fiocco M, Geskus RB. **Tutorial in biostatistics: competing risks and multi-state models. Stat Med. 2007, 26:2389–2430**

17. Biau DJ, Latouche A, Porcher R. **Competing events influence estimated survival probability: when is Kaplan-Meier analysis appropriate? Clin Orthop Relat Res. 2007; 462:229–233**

18. Gillam MH, Ryan P, Graves SE, Miller LN, de Steiger RN, Salter A (2010) **Competing risks survival analysis applied to data from the Australian orthopaedic association national joint replacement registry. Acta Orthop 81:548–555**

19. Rowden NJ, Harrison JA, Graves SE, Miller LN, de Steiger RN, Davidson DC. **Loss to follow-up after arthroplasty: a new use for registry data..J Bone Joint Surg Br. 2012 Apr;94(4):493-6. doi: 10.1302/0301-620X.94B4.27532.**

20. Sayers A, Crowther MJ, Judge A, Whitehouse MR, Blom AW. **Determining the sample size required to establish whether a medical device is non-inferior to an external benchmark. BMJ Open. 2017 Aug 28;7(8):e015397. doi: 10.1136/bmjopen-2016-015397.**

21. Keurentjes JC, Pijls BG, Van Tol FR, Mentink JF, Mes SD, Schoones JW, Fiocco M, Sedrakyan A, Nelissen RG. **Which implant should we use for primary total hip replacement? A systematic review and meta-analysis. J. Bone Joint Surg Am. 2014 Dec 17;96 Suppl 1:79-97. doi: 10.2106/JBJS.N.00397.**

22. Deere KC, Whitehouse MR, Porter M, Blom AW, Sayers A. **Assessing the non-inferiority of prosthesis constructs used in hip replacement using data from the National Joint Registry of England, Wales, Northern Ireland and the Isle of Man: a benchmarking study. BMJ Open. 2019 Apr 29;9(4):e026685. doi: 10.1136/bmjopen-2018-026685.**

23. Deere KC, Whitehouse MR, Porter M, Blom AW, Sayers A. **Assessing the non-inferiority of prosthesis constructs used in total and unicondylar knee replacements using data from the National Joint Registry of England, Wales, Northern Ireland and the Isle of Man: a benchmarking study. BMJ Open. 2019 Apr 29;9(4):e026736. doi: 10.1136/bmjopen-2018-026736.**

24. Nelissen RGHH, Tucker K, Kjaersgaard-Andersen P. **Benchmarking Implants, A NORE report. EFORT.org/NORE**

25. ODEP (orthopaedic data evalation panel). ODEP.org.uk

26. Beyond compliance. beyondcompliance.org.uk

27. Sayers A, Evans JT, Whitehouse MR, Blom AW. **Are competing risks models appropriate to describe implant failure? Acta Orthop. 2018 Jun;89(3):256-258. doi: 10.1080/17453674.2018.1444876. Epub 2018 Mar 9.PMID: 29521152**

RESEARCH TOPIC

FUNCTIONALIZED IMPLANTS/ BIOMATERIALS/ SURFACES/INNOVATIONS

| Research Topic 31

Functionalized Surfaces or Novel Aspects in Hip and Knee Arthroplasty – Review and Proposal of a Stepwise Analysis Approach

RESEARCH TOPIC 31

Functionalized Surfaces or Novel Aspects in Hip and Knee Arthroplasty – Review and Proposal of a Stepwise Analysis Approach

Authors

Georg N. Duda[1], Marcus Jäger[2,3], Bernard Masson[4], Eduardo Garcia-Rey[5,6], Maria Angeles Pérez Ansón[7], Gwendolen Reilly[8,9], Ronja A. Schierjott[10]

[1]Julius Wolff Institute and Center for Musculoskeletal Surgery, Berlin Institute of Health at Charité - Universitätsmedizin Berlin, Berlin, Germany

[2]Department of Orthopaedics, Trauma & Reconstructive Surgery, St. Marien Hospital Mülheim an der Ruhr, Mülheim, Germany

[3]Orthopaedics and Trauma Surgery, University of Duisburg-Essen, Essen, Germany

[4]Medical Division, BioConnect, Vieille-Toulouse, France

[5]Hôpital Universitario La Paz-Idi Paz, Madrid, Spain

[6]Centro de Investigación Biomédica en Red de Bioingeniería, Biomateriales y Nanomedicina, CIBER-BBN, Madrid, Spain

[7]Department of Mechanical Engineering, Aragon Institute for Engineering Research (I3A), University of Zaragoza, Zaragoza, Spain

[8]Department of Materials Science and Engineering, University of Sheffield, Sheffield, S1 3BJ, United Kingdom

[9]INSIGNEO Institute of in Silico Medicine, University of Sheffield, Sheffield, S1 3BJ, United Kingdom

[10]Aesculap AG Research & Development, Tuttlingen, Germany

1. Abstract

Question: Functionalized implants / biomaterials / surfaces / innovations

How is it possible to evaluate functionalized surfaces or novel aspects of implants for which no standardized / established test methods yet exist and where no proof of function is yet defined?

Summary / Recommendation:

Based on the intended use and effect(s), resulting requirements and potential negative effects/risks, a test cascade (with increasing complexity from standardized in vitro tests to clinical studies) should be developed individually for the functionalized surfaces or novel aspect that shall be evaluated.

A clear idea about which components of the new product are actually novel (novel component) and which components have already been applied, i.e. are already approved (existing component), should be obtained. Assessment results for specific aspects of the existing components may be transferred to the new product – provided that they are unaffected by the novel components – to reduce redundant test effort.

A detailed list of the requirements and potential risks of the product in the planned application should be prepared. Based on this, for each individual requirement and risk, it should be reflected upon the following: Can the specific intended use or aspect be tested pre-clinically or is a clinical study required? Are there standardized test methods which can be directly applied? Are there test methods which can be adapted? If no: Which type of study (pre-clinical or clinical) could be useful to assess the intended use or additionally relevant aspect of an innovation? Is it possible to develop appropriate but novel test methods sufficiently based on existing standards?

In areas where distinct regulations are not yet available, i.e. for real innovations and where specific aspects (requirements/risks) cannot be assessed by pre-clinical tests, a de-risked strategy of early clinical assessments could be used, i.e. a stepwise clinical application with a minimal number of patients (1. Single center, 2. Multiple centers, 3. Post-Market surveillance/Registry data).

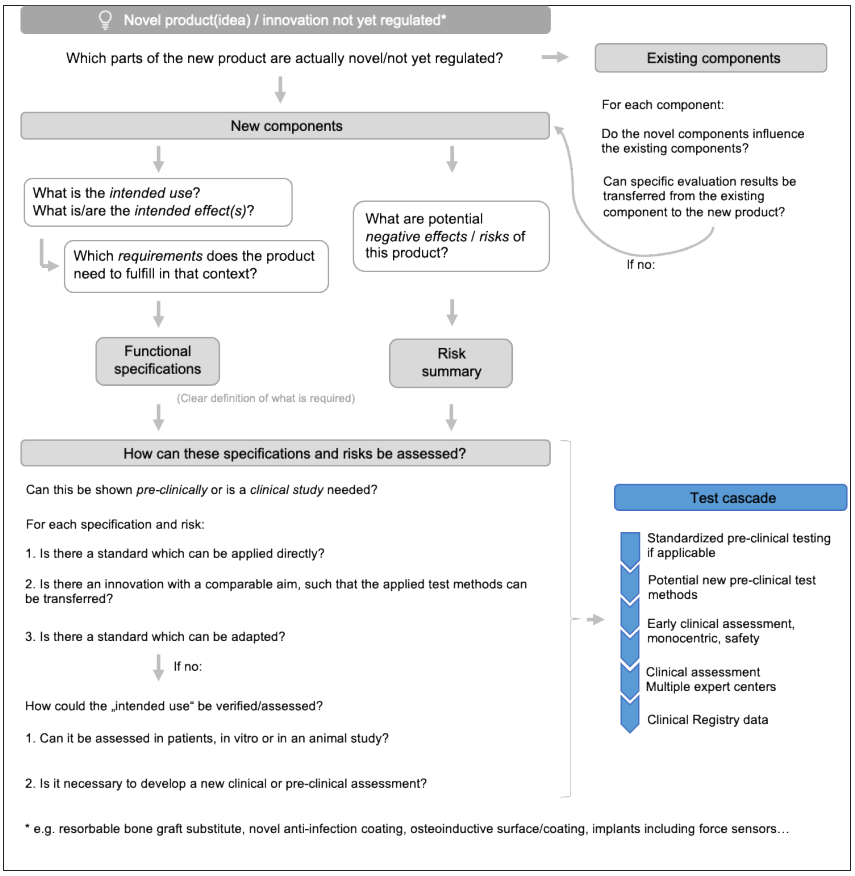
2. Level of Evidence

N/A

3. Consensus Delegate Vote

89% - super majority, strong consensus (89% agree / 0% disagree / 11% abstain)

4. Graphical Abstract



5. Search Strategy

A systematic literature review was not found to be effective for the assigned question. Hence, the here presented draft is mainly based on previous knowledge and experience of the group members and the level of evidence is not applicable. There is no classical evidence level existing for the proposed processing suggested herein, but a similar Intended-use oriented approach has been established in ATMP and other regulatory pathways and might also help here.

6. Rationale

This topic refers to currently not yet regulated new technologies or improvements. Therefore, we suggest in the following a sketch of a guide that could be used by regulatory bodies or basic / clinician researchers on their path to seek regulatory approval in translating novel therapies and strategies into regular clinical usage.

In the following, a first sketch of a stepwise approach to define suitable analysis methods for novel product(ideas) or innovations, including some examples, is presented.

Differentiation between novel components and already applied/existing components

In the very beginning, one should obtain a clear idea about which parts of the new product to be tested are actually novel (novel components) and which parts may have already been applied in existing products (existing components). Knowledge and assessment results for specific aspects or characteristics (e.g. biocompatibility, fatigue resistance…) may be transferred from existing components to the new product resulting in a reduction of redundant test efforts. However, this should only be done in cases where the specific aspect or characteristic of the existing component is not affected by novel component.

Example: A new bone graft substitute consisting of granules has been developed. The material (hydroxyapatite) has already been used clinically (existing component), but the tetrapod-shape of the single granules is new (novel component). In this case, the novel component, i.e. the new shape aims at influencing the durability of mechanical competence under cyclic loading (intended use: more durable), but not for example the biocompatibility. Hence, for this example, biocompatibility results can be transferred from previous products, i.e. biocompatibility does not need to be evaluated again, whereas fatigue tests need to be repeated for the new product's novel intended use.

Intended use and intended effect(s)

Based on proper pre-clinical work considering all available regulatory settings, a very distinct "intended use" definition should be formulated and thereby defined that reflects the potential and intention of the novel therapeutics or clinical strategy (i.e. new product(idea) or innovation) standing up for approval.

The "intended use" definition should be carefully formulated since it is the only reference that will later be assessed by regulatory authorities and that needs to be also demonstrated e.g. by first clinical data.

An important prerequisite is to define the expected and intended effect(s) of a distinct functionalized implant surface (i.e. new product(idea) or innovation in general) by the distributor prior to pre-clinical or clinical testing since this leads to its intended use. All intended effects should be listed.

Example: A novel synthetic bone graft substitute (BGS) has been developed. It is a composite material, consisting of granules made of a novel (not yet clinically applied) ceramic and a PEG-matrix.

Intended use: Regeneration of bone in a defect in a load-bearing situation.

Intended effects: Easy filling of defects of any shape (moldable material), load-bearing filling of defects, osseointegration, biological reconstruction by resorption of the biomaterial over time and replacement by new bone.

Requirements in the specific context of intended use

Once the intended use and intended effect(s) are defined, the requirements which need to be fulfilled by the new product(idea) / innovation should be summarized. This includes general and well-defined regulatory requirements (e.g. biocompatibility), as well as more specific requirements related to the intended use but where adequate regulations are missing (e.g. complete resorption) or where it will be hard to evaluate (e.g. mechanical integrity under dynamic loading, no failure of the void filling…).

Remark: Some of the here mentioned aspects are already regulated and are listed for exemplary purposes, i.e. to give an impression of the stepwise evaluation approach. For these regulated aspects, the corresponding standards should be applied, and they will not be discussed here in detail.

Example: Some exemplary requirements for the BGS include biocompatibility, moldability, ability to achieve primary stability, long-term stability under physiologic loading (application behind acetabular cup, around hip stem, behind tibial or femoral component), and osseointegration potential.

Potential negative effects / risks

Besides the requirements to fulfill a certain intended use and effect(s), it should also be assessed, which potential negative effects or risks are associated with the new product. This is especially important for products with pharmaceutical effects (regulated as combination devices) or materials of animal origin.

Example: Potential risks of the BGS include potential toxicity, the possibility of a granule getting between the articulation partners or subsidence of the implant component due to the BGS resolving too quickly.

The intended use formulation, intended effects and requirements should ideally result in functional specifications, i.e. a list of precisely and positively formulated requirements. Where possible, a clear parameter to assess the specific requirement should be defined and requirements should be quantified.

All identified risks or potential negative effects should be listed in a risk summary.

After summarizing and organizing all the information, the question "How can these specifications and risks be assessed?" can now be answered stepwise for each listed specification/risk. First, one could ask himself/herself whether the specification/risk could be assessed pre-clinically or if it has to be assessed clinically.

Example: For an antimicrobial coating, it can be proven pre-clinically that there is an effect (i.e. less bacterial growth on the coating than on specimens without coating). However, from this it cannot be concluded that the coating will achieve the intended effect in the clinical application (i.e. reduction of infection rate). As prevalence of infections is usually less than 3% in TJA procedures (1, 2) and the potential effect of the coating is hence quite low (i.e. max. reduction 3%) In order to really show a statistical superiority, a very large patient group would be required. Given the fact that at the point of CE-marking it has already been proven that the coating does not represent a risk for the patient, the positive clinical effect (i.e. reduced infection rate) of such a coating may be demonstrated post-marketing, but a preventive means would be impossible to declare as intended use up front.

When it comes to clinical assessments, a clear definition of what parameters would be used or assessed to document in a later clinical setting that the defined intended use has been reached/fulfilled would be useful. This formulation of intended use should be based on the magnitude and severity of the addressed problem, and, the type, potential success or risks and costs of the proposed solution. It may be helpful to define an expected effect size/expected effect that later serves as reference for benchmarking the innovation. This appears essential relevant in all cases where no current standards yet exist.

Once it is clear if the corresponding aspect can generally be assessed clinically or pre-clinically, it should be checked if there are existing standards which can be applied directly or adapted to test the specific aspect.

If neither of this is possible, new pre-clinical or clinical test methods may be developed (e.g. referring to literature or existing standards).

Finally, after repeating this stepwise process for each specification and risk, a test cascade from standardized in vitro tests to clinical studies may be defined.

Example for a new pre-clinical test method (very specific for each new product to be tested):

In order to test the primary stability achievable with the BGS, an in vitro test method was developed to assess relative motion between the implant and the surrounding setup (3), oriented towards previous primary stability tests (4, 5).

Early clinical assessment (example for new clinical assessment)

We assume that all available and required pre-clinical assessments will be properly done.

It must be emphasized that the pre-clinical evaluation must be investigated very carefully. The main target is patient safety. In order to reduce the potential risk of the introduction of a new technology or innovation, all existing tests (in vitro, biomechanical, toxicity, interaction with the host etc.) must be performed according the current regulations and should be within the current safety margin allowances, if available.

All such information should be obtained prior to any first clinical application, i.e. where possible, aspects of the innovation/functionalized material should be assessed pre-clinically (e.g. in vitro or in animal studies) to reduce risks for the patients.

In case of real innovations (i.e. products for which no regulations yet exist) and if the specific aspects of these innovations cannot be assessed pre-clinically,

a distinct de-risked strategy of early clinical assessments could be used

- 1. Single center study: The intended use will be analyzed in a small controlled group of patients that allow to assess an effect size, e.g. a group of 10-20 patients, in a clinical expert center such as an university hospital setting with proper clinical and scientific expertise to perform such first-in-human therapy. The effect of the intended use is verified according to the measures pre-defined by the inventors. After regulatory approval of the outcomes of this first phase clinical trial setting the second level may be applied for. Small prospective randomized clinical studies are ideal for this purpose. The introduction of innovative implant surface structures or coatings might lead to a scenario, where standard and approved tests did not meet the requirements to measure and evaluate the effects expected. In such cases, an early clinical application to a defined and small cohort of probands who have no clinical useful treatment options regarding standard treatment protocols or the application of approved but insufficient products (e.g. amputation if a limb).
- 2. Multicenter study: Under control of the regulatory authorities, we propose to assess in a second (or possibly third), independent clinical expert center the performance of the novel therapy in respect to the intended use. Additionally, multicenter prospective clinical studies, including larger number of patients, can also provide significant findings. The setting should be identical to the one of the initial phases in respect to outcome assessments of the intended use.
- 3. Post-Market surveillance/registry data/PMCF study: After regulatory approval of the outcome of the above clinical evaluations, the novel therapy could be approved and post-marketing surveillance is further to be conducted, e.g. in terms of registry data bases, as currently defined.

Although randomized controlled clinical trials are more difficult to perform (6), they may be the only and ultimate check of an intended use and should be performed in small and well controlled settings to avoid later failures. Previous, well documented, failures regarding a non-ideal stepwise introduction of a novel technology should be avoidable by such an iterative approach (7-9). Further, an independent post-marketing surveillance accompanied by further clinical analyses should be aimed for to allow a confirmation of the initial findings in both existing and emerging technologies (10, 11).

7. References

1. George J, Klika AK, Navale SM, Newman JM, Barsoum WK, Higuera CA. **Obesity Epidemic: Is Its Impact on Total Joint Arthroplasty Underestimated? An Analysis of National Trends. Clin Orthop Relat Res. 2017 Jul;475(7):1798–1806. doi: 10.1007/s11999-016-5222-4.**

2. Kong, L., Cao, J., Zhang, Y., Ding, W., & Shen, Y. (2017). **Risk factors for periprosthetic joint infection following primary total hip or knee arthroplasty: a meta-analysis. International wound journal, 14(3), 529–536.**

3. Schierjott RA, Hettich G, Ringkamp A, Baxmann M, Morosato F, Grupp TM. **A method to assess primary stability of acetabular components in association with bone defects. Journal of Orthopaedic Research 2020;38(8):1769–78, doi:10.1002/jor.24591.**

4. Beckmann NA, Bitsch RG, Gondan M, Schonhoff M, Jaeger S. **Comparison of the stability of three fixation techniques between porous metal acetabular components and augments. Bone & joint research 2018;7(4):282–88, doi:10.1302/2046-3758.74.BJr-2017-0198.r1.**

5. Huber WO, Noble PC. **Effect of design on the initial stability of press-fit cups in the presence of acetabular rim defects: experimental evaluation of the effect of adding circumferential fins. International orthopaedics 2014;38(4):725–31, doi:10.1007/s00264-013-2187-8.**

6. Griffin XL, Haddad FS. **Evidence-based decision making at the core of orthopaedic practice. Bone Joint J. 2014 Aug;96-B(8):1000-1. doi: 10.1302/0301-620X.96B8.34614. PMID: 25086112.**

7. Malchau H, Bragdon CR, Muratoglu OK. **The stepwise introduction of innovation into orthopedic surgery: the next level of dilemmas. J Arthroplasty. 2011 Sep;26(6):825-31. doi: 10.1016/j.arth.2010.08.007. Epub 2010 Oct 2. PMID: 20888183.**

8. Morlock MM. **The taper disaster--how could it happen? Hip Int. 2015 Jul-Aug;25(4):339-46. doi: 10.5301/hipint.5000269. Epub 2015 May 20. PMID: 26044535.**

9. Reito A, Lehtovirta L, Lainiala O, Mäkelä K, Eskelinen A. **Lack of evidence-the anti-stepwise introduction of metal-on-metal hip replacements. Acta Orthop. 2017 Oct;88(5):478-483. doi: 10.1080/17453674.2017.1353794. Epub 2017 Jul 18. PMID: 28715928; PMCID: PMC5560208.**

10. Charnley J. **The long-term results of low-friction arthroplasty of the hip performed as a primary intervention. J Bone Joint Surg Br. 1972 Feb;54(1):61-76. PMID: 5011747.**

11. Hauptfleisch J, Glyn-Jones S, Beard DJ, Gill HS, Murray DW. **The premature failure of the Charnley Elite-Plus stem: a confirmation of RSA predictions. J Bone Joint Surg Br. 2006 Feb;88(2):179-83. doi: 10.1302/0301-620X.88B2.17055. PMID: 16434520.**

RESEARCH TOPIC

IN SILICO TRIALS

Research Topic 32

In Silico Trials Methodologies within the Development, Pre-Clinical Assessment and Clinical Evaluation Process of Total Joint Arthroplasty Implants

RESEARCH TOPIC 32

In Silico Trials Methodologies within the Development, Pre-Clinical Assessment and Clinical Evaluation Process of Total Joint Arthroplasty Implants

Authors

Marco Viceconti[1], Bernd Grimm[2], Walter Van der Weegen[3], Francesco Traina[4,5], Ina Wüstefeld[6], Claudia Mazzà[7,8], Marcel Dreischarf[9], Christoph Lohmann[10]

[1]Department of Industrial Engineering, Alma Mater Studiorum – University of Bologna and Medical Technology Lab, IRCCS Istituto Ortopedico Rizzoli, Bologna, Italy
[2]Human Motion, Orthopaedics, Sports Medicine and Digital Methods Group, Luxembourg, Institute of Health, Transversal activities, Luxembourg, Luxembourg
[3]Sports and Orthopaedics Research Center, Anna Hospital, Geldrop, the Netherlands
[4]Ortopedia-Traumatologia e Chirurgia protesica e dei reimpianti d'anca e di ginocchio, IRCCS Istituto Ortopedico Rizzoli, Bologna, Italy
[5]DIBINEM, Università di Bologna, Bologna, Italy
[6]Aesculap AG Medical Scientific Affairs, Tuttlingen, Germany
[7]INSIGNEO Institute for in silico Medicine, The University of Sheffield, Sheffield, UK
[8]Department of Mechanical Engineering, The University of Sheffield, Sheffield, UK
[9]Berlin Institute of Health, Julius Wolff Institute for Biomechanics and Musculoskeletal Regeneration, Charité-Universitätsmedizin Berlin, Germany
[10]Department of Orthopaedic Surgery, Otto-Von-Guericke-University, Magdeburg, Germany

1. Abstract

Question: In silico Trials (Big Data Analytics, Machine Learning, System Biology models, system physiology models)

Part 1

In what ways can In Silico Trials methodologies (whether mechanistic like Finite Element Analysis or data-based like machine learning) contribute to the assessment and evaluation of implants?

- Is it possible to support planning, implementation and data analysis of safety studies in the pre-CE phase with In Silico Trials relying on adequate data (regarding quality and quantity/amount, e.g. registry data or HIS data)?
- Is it possible to increase the exploratory power of studies and tests in the pre-CE phase (e.g. of small cohorts or regarding surgical approaches, indications etc.) with In Silico Trials?

Part 2

How is it possible to use elements of in silico pre-clinical and clinical trials (i.e. FEA, multi-body simulations), AI/ML & Big Data as basis for implants, instruments, procedure (e.g. pre-op planning, pre-op positioning)?

- In what way would implant systems relying on AI/ML & Big Data (e.g. patient specific target tibial slope) require pre-clinical testing/assessment or clinical follow-up different to classic systems?
- Should it be required to show that the training and validation datasets for such a system a) represent the targeted patient population in its diversity, b) was large enough to achieve a specific threshold of predictive quality?

Summary / Recommendation:

Part 1

- The recent work done by the Avicenna Alliance GSP Task Force has made evident a wide range of possible Context of Use for In Silico Trials technologies, including their use in the pre-clinical evaluation of medical devices.
- With respect to subquestion a) the experts agree that In Silico Trials can support pre-clinical safety and efficacy studies. Mechanistic models can be used to reduce and refine, and for some simpler cases even replace, bench, in vitro, and ex vivo experiments. Phenomenological models are more difficult to use in this context.
- With respect to subquestion b), the experts agree that In Silico Trials can and should be used to increase the exploratory power of pre-clinical studies.

Part 2

- AI technologies could automate 2D templating and 3D planning in orthopaedic implantology; supervised AI methods will face easier regulatory pathways.
- Mechanistic patient-specific models (e.g. finite element, etc.) can predict the risk associated to a known failure mode in specific patient and for given surgical planning. However, their need for detailed patient-specific information (3D imaging, gait analysis, etc.) limit their clinical applicability as clinical decision-support systems.
- Mechanistic patient-specific models are ideal for In Silico Trials where, combined with Virtual Cohorts, can predict the safety and efficacy of new implant designs or new surgical techniques.
- Phenomenological (data-driven) models (e.g. Machine Learning) are more likely to develop into clinically applicable decision-support systems to plan the treatment of individual patients.
- However, the low failure rate of joint replacements requires huge data collections to accurately train such data-driven model; this calls for collaborative, data-sharing projects.

2. Level of Evidence

Moderate

3. Consensus Delegate Vote

Part 1

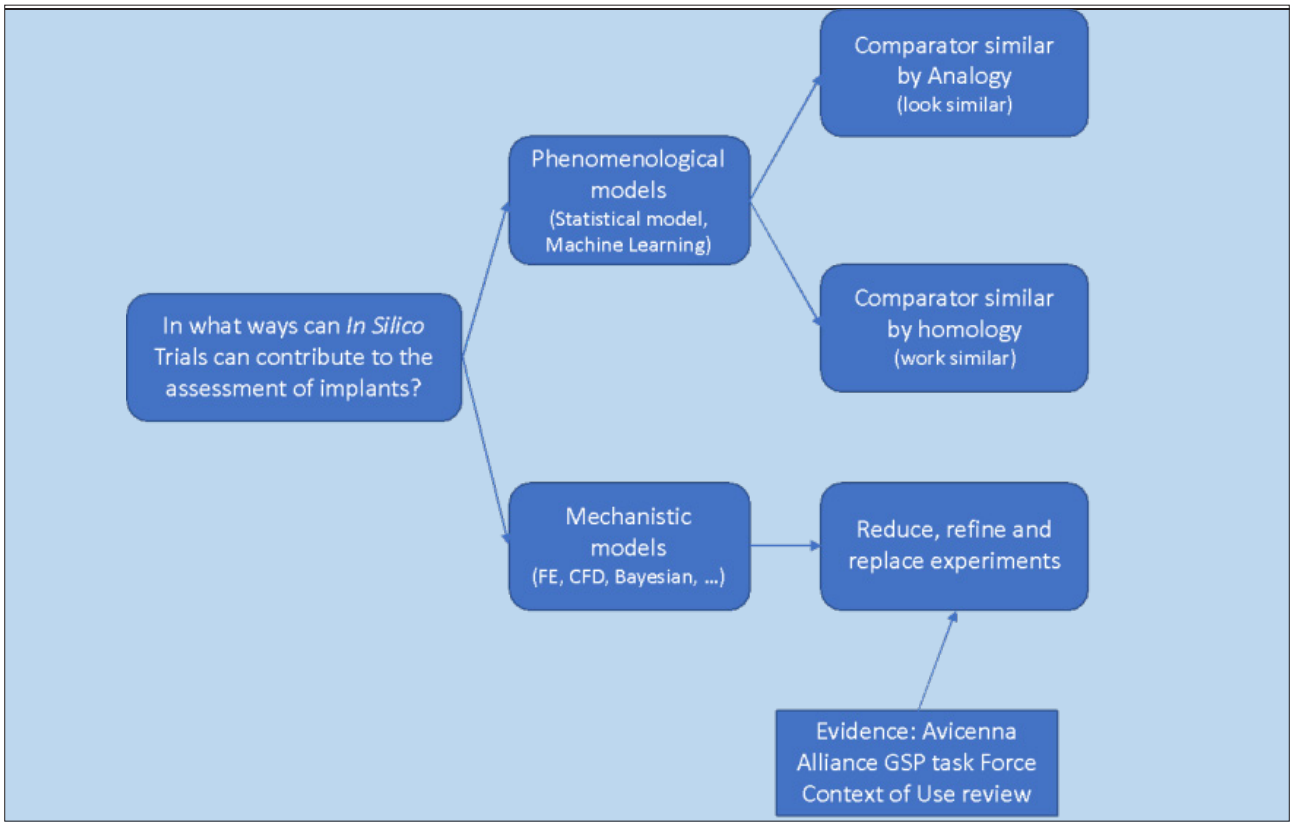
97% – unanimous, strongest consensus (97% agree / 0% disagree / 3% abstain)

Part 2

95% – unanimous, strongest consensus (95% agree / 0% disagree / 5% abstain)

4. Part 1

a. Graphical Abstract



b. Search Strategy

Consensus process

c. Rationale

The term In Silico Trials indicates the use of any kind of modelling & simulation method to evaluate the safety and/or efficacy of a new medical product, independently from the specific modelling methodology in use, or the product development phase. We can have In silico Trials that use machine learning to do pre-clinical testing, or system physiology models to do clinical testing. Hereinafter we will separate such methods in phenomenological and mechanistic:

- Phenomenological models are those exclusively based on data, and not of pre-existing knowledge. This category includes all classical frequentist statistics modelling methods, and machine learning methods.
- Mechanistic models are those build using data and pre-existing knowledge, in various degree. This includes Bayesian statistical models, non-linear system identification models, system biology and system physiology models based on differential equations, agent-based models build with mechanistic rules, etc.

This distinction is important for what follows.

The recent work done by the Avicenna Alliance GSP Task Force has made evident a wide range of possible Context of Use for In Silico Trials technologies. A recent consensus review (1) identified 45 possible contexts of use in which an In Silico Trials methodology can be used to reduce, refine, and replace experimental methods in the producing evidence to obtain the marketing authorization for a new product. Among them some aim to the pre-clinical evaluation of medical devices. For 31 of the 45, at least one reference could be found, while the others should for now be considered only speculative.

With respect to subquestion a) the experts agree that In Silico Trials can support pre-clinical safety and efficacy studies. Mechanistic models (e.g. biophysics models like FEM) can be used to reduce and refine, and for some simpler cases even replace, bench, in vitro, and ex vivo experiments.

Phenomenological models (such as machine learning) are more difficult to use in this context. Mechanistic models are largely based on prior knowledge we have on the anatomy, physiology, chemistry, and physics of the phenomenon of interest; the data required to inform such models are limited to specific features that describe the device, the inter-subject variability, and the surgical variability.

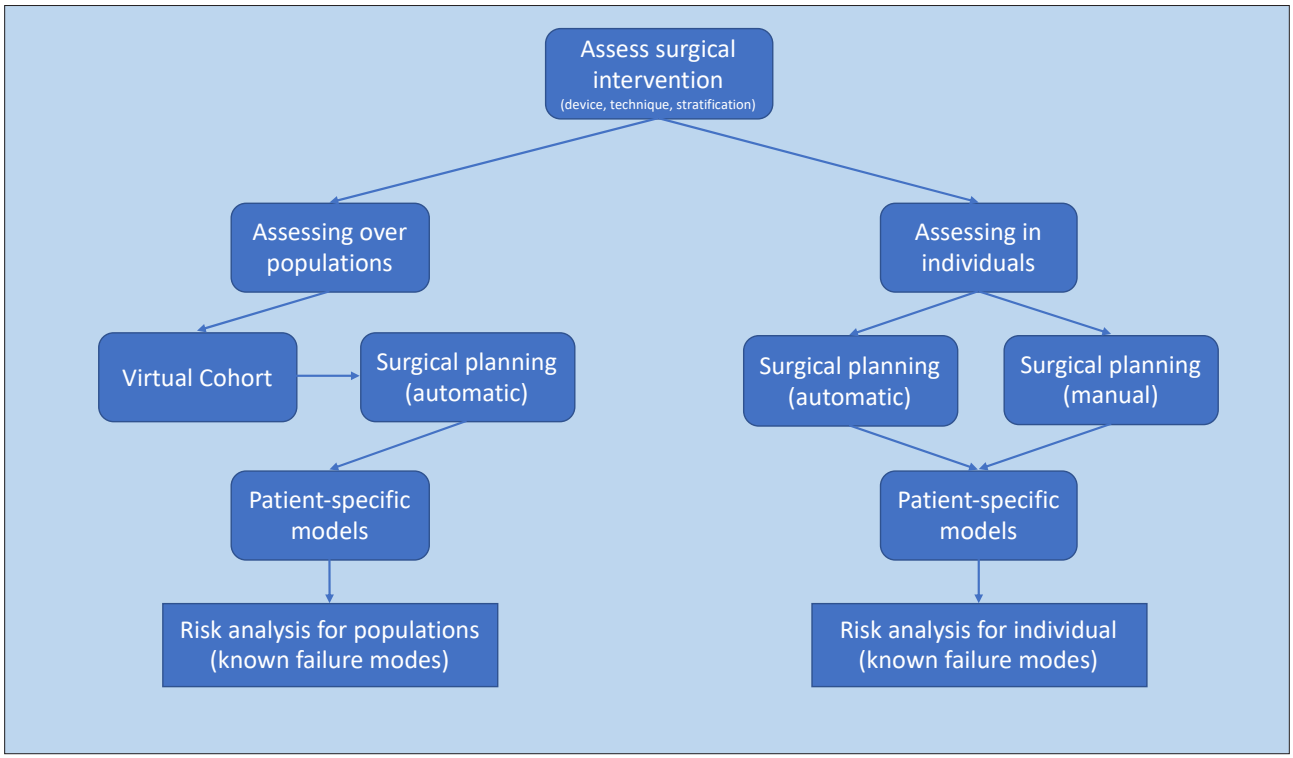
On the contrary phenomenological models are entirely based on data. Because it is unlikely that we have large volumes of observational data for the new device being tested, such models must be built and validated using data obtained for "similar" devices. However, here we need to define similarity by homology, e.g. devices that function similarly to the one of interest; too frequently defining such functional similarity is extremely difficult, and we end up resorting to similarity by analogy, devices that look similar to the one of interest. This in some cases may produce inaccurate conclusions.

Less sensitive to this problem are grey-box models such a Bayesian statistics model. In these models the probability of an adverse event is calculated as the sum of the likelihood (the probability as observed in an experiment) and the prior (probability as observed in prior studies or derived from mechanistic knowledge); while the problem of choosing the prior by similarity remains, if the probability of the prior is considerably different from that of the likelihood, this suggest the prior might have been erroneously chosen.

With respect to subquestion b), the experts agree that In Silico Trials can and should be used to increase the exploratory power of pre-clinical studies. Predictive models can be used to explore the effect of operational conditions to find the worst-case conditions under which to run the experimental study (e.g. (2)). Predictive models can also be used to expand the results of an experimental study by showing that model can predict the experimental results, and then use the model to explore a wider range of variability factors (patient-related, surgery-related, device related) (e.g. (3)).

5. Part 2

a. Graphical Abstract



b. Search Strategy

Due to the limited literature available no systematic literature review was possible.

c. Rationale

An orthopaedic surgery planning system is composed of two parts: one that defines the size and the pose (position + orientation) of the implant components with respect to the patient's anatomy; the other that predicts the probability that if that implant with that size is placed in that pose in that patient anatomy, a given complication may occur.

Implant sizing and positioning are usually done manually by the surgeon, using some sort of surgical navigation interface (4–6). When this operation is automated, it usually relies on statistical (7) or machine learning (8) algorithms. As the aim is to reproduce the typical average choice of a trained surgeon, these systems are frequently supervised, e.g. the surgeon can make small corrections or approve the automated plan.

Models that predict the risk of a given complication for a given surgical plan are becoming common (9–13).

Many are based on mechanistic (typically biophysics-based) models, and are designed to predict the risk for a given complication (e.g.

dislocation, intra-operative fracture, etc.). These models can be easily validated with ex vivo (10) or intra-operative (14, 15) experiments, by showing that the model accurately predicted the range of motion, the primary stability, etc. The main shortcoming of these mechanistic models is that they require detailed information on the patient, to be build. Typically, 3D anatomy is required (by mean of CT or MRI imaging), and frequently also preoperative functional information (e.g. gait analysis, range of motion, dynamometry, etc.). While such detailed examination of each patient is possible in clinical studies, it is hard to imagine this translated into a clinical pathway, for cost and time reasons. While a widespread clinical adoption of these detailed mechanistic models as patient-specific planning tools is unlikely, these methods can be used to predict the effect of implant design choices or of surgical techniques by including statistical representations of the patient-related and surgery-related variability (16–19). It is easy to imagine the development of these methods into full-blown In Silico Trials, which have been recently reported for other medical devices (e.g. (20)).

The use of data-based models for such purpose is more recent. Artificial neural networks (ANN) and support vector machine (SVM) methods were used to optimise implant design with respect to stress shielding (21); various machine learning (ML) algorithms were used to optimise soft tissue balancing (22); Deep learning was used to identify the primary implant model in revision surgeries (23); multiple ML algorithms were tested to recognise size mismatch (24).

In principle one could train an Artificial Neural Network (ANN) to recognise from a post-operative radiograph, which implants will fail over a given time frame; the surgical planning software could then generate a synthetic x-ray that AI system use to predict the risk of failure. However, there are two issues, with this approach. The first is that to train well an ANN you need to have at least a few hundred positive cases, cases that failed within the time frame. Considering that the failure rate for joint replacements is usually less than 5% nowadays, you would need more than 2000 cases to collect 100 failures. The second issue is that such system would warn that that surgical plan is at risk, but would not specify what is wrong, or which failure mode is more likely. To add such explanatory power, we would need to train the ANN on sub-groups (aseptic loosening, massive wear, recurrent dislocation, etc.), and this would exacerbate the first problem. The development of ANN predictors that can discriminate between failure modes would require the creation of a very large (≈100,000 cases) collection of anonymised radiographs, annotated with some patient information and if failed the time of failure and the failure mode. The EFORT would be in an ideal position to create and manage such previous resource.

6. References

1. Viceconti M, Emili L, Afshari P, Courcelles E, Curreli C, Famaey N et al. **Possible Contexts of Use for In Silico trials methodologies: a consensus- based review; 2021. (PP).**

2. Trieb K, Huber D, Sonntag R, Kretzer JP. **Finite-Elemente-Analyse und biomechanische Testung des neuen MiniMIS-Kurzschafts. Z Orthop Unfall 2019; 157(2):188–93.**

3. Martelli S, Taddei F, Schileo E, Cristofolini L, Rushton N, Viceconti M. **Biomechanical robustness of a new proximal epiphyseal hip replacement to patient variability and surgical uncertainties: a FE study. Medical Engineering & Physics 2012; 34(2):161–71. Available from: URL: https://www.sciencedirect.com/science/article/pii/s1350453311001639.**

4. Bono JV. **Digital Templating in Total Hip Arthroplasty. JBJS 2004; 86(suppl_2):118. Available from: URL: https://journals.lww.com/jbjsjournal/fulltext/2004/12002/digital_templating_in_total_hip_arthroplasty.16.aspx.**

5. Lattanzi R, Viceconti M, Zannoni C, Quadrani P, Toni A. **Hip-Op: an innovative software to plan total hip replacement surgery. Medical Informatics and the Internet in Medicine 2002; 27(2):71–83.**

6. Colombi A, Schena D, Castelli CC. **Total hip arthroplasty planning. EFORT Open Reviews 2019; 4(11):626–32.**

7. Otomaru I, Nakamoto M, Kagiyama Y, Takao M, Sugano N, Tomiyama N et al. **Automated preoperative planning of femoral stem in total hip arthroplasty from 3D CT data: atlas-based approach and comparative study. Medical Image Analysis 2012; 16(2):415–26. Available from: URL: https://www.sciencedirect.com/science/article/pii/s1361841511001435.**

8. Jodeiri A, Zoroofi RA, Hiasa Y, Takao M, Sugano N, Sato Y et al. **Fully automatic estimation of pelvic sagittal inclination from anterior-posterior radiography image using deep learning framework. Computer Methods and Programs in Biomedicine 2020; 184:105282. Available from: URL: https://www.sciencedirect.com/science/article/pii/s0169260719305024.**

9. Testi D, Simeoni M, Zannoni C, Viceconti M. **Validation of two algorithms to evaluate the interface between bone and orthopaedic implants. Computer Methods and Programs in Biomedicine 2004; 74(2):143–50. Available from: URL: https://www.sciencedirect.com/science/article/pii/s0169260703000786.**

10. Reggiani B, Cristofolini L, Varini E, Viceconti M. **Predicting the subject-specific primary stability of cementless implants during pre-operative planning: preliminary validation of subject-specific finite-element models. Journal of Biomechanics 2007; 40(11):2552–8. Available from: URL: https://www.sciencedirect.com/science/article/pii/s0021929006004404.**

11. Palit A, King R, Gu Y, Pierrepont J, Simpson D, Williams MA. **Subject-Specific Surgical Planning for Hip Replacement: A Novel 2D Graphical Representation of 3D Hip Motion and Prosthetic Impingement Information. Ann Biomed Eng 2019; 47(7):1642–56. Available from: URL: https://link.springer.com/article/10.1007/s10439-019-02260-x.**

12. Viceconti M, Ascani D, Mazzà C. **Pre-operative prediction of soft tissue balancing in knee arthoplasty part 1: Effect of surgical parameters during level walking. Journal of Orthopaedic Research® 2019; 37(7):1537–45.**

13. Asseln M, Eschweiler J, Trepczynski A, Damm P, Radermacher K. **Evaluation and validation of 2D biomechanical models of the knee for radiograph-based preoperative planning in total knee arthroplasty. PLOS ONE 2020; 15(1):e0227272. Available from: URL: https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0227272.**

14. Tanino H, Sato T, Nishida Y, Mitsutake R, Ito H. **Hip stability after total hip arthroplasty predicted by intraoperative stability test and range of motion: a cross-sectional study. BMC Musculoskelet Disord 2018; 19(1):373. Available from: URL: https://bmcmusculoskeletdisord.biomedcentral.com/articles/10.1186/s12891-018-2289-y.**

15. Matsumoto T, Takayama K, Ishida K, Kuroda Y, Tsubosaka M, Muratsu H et al. **Intraoperative Soft Tissue Balance/Kinematics and Clinical Evaluation of Modified Kinematically versus Mechanically Aligned Total Knee Arthroplasty. J Knee Surg 2020; 33(8):777–84.**

16. Viceconti M, Brusi G, Pancanti A, Cristofolini L. **Primary stability of an anatomical cementless hip stem: a statistical analysis. Journal of Biomechanics 2006; 39(7):1169–79. Available from: URL: https://www.sciencedirect.com/science/article/pii/s0021929005001648.**

17. Del-Valle-Mojica J-F, Alonso-Rasgado T, Jimenez-Cruz D, Bailey CG, Board TN. **Effect of Femoral Head Size, Subject Weight, and Activity Level on Acetabular Cement Mantle Stress Following Total Hip Arthroplasty. Journal of Orthopaedic Research® 2019; 37(8):1771–83.**

18. Koh Y-G, Park K-M, Lee H-Y, Park J-H, Kang K-T. **Prediction of wear performance in femoral and tibial conformity in patient-specific cruciate-retaining total knee arthroplasty. J Orthop Surg Res 2020; 15(1):24. Available from: URL: https://josr-online.biomedcentral.com/articles/10.1186/s13018-020-1548-4.**

19. Ali AA, Clary CW, Smoger LM, Dennis DA, Fitzpatrick CK, Rullkoetter PJ et al. **Computational framework for population-based evaluation of TKR-implanted patellofemoral joint mechanics. Biomech Model Mechanobiol 2020; 19(4):1309–17. Available from: URL: https://link.springer.com/article/10.1007/s10237-020-01295-7.**

20. Karanasiou GS, Tsobou PI, Tachos NS, Antonini L, Petrini L, Pennati G et al. **Design and implementation of in silico clinical trial for Bioresorbable Vascular Scaffolds. In: 2020 42nd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC). IEEE; 2020.**

21. Cilla M, Borgiani E, Martínez J, Duda GN, Checa S. **Machine learning techniques for the optimization of joint replacements: Application to a short-stem hip implant. PLOS ONE 2017; 12(9):e0183755. Available from: URL: https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0183755.**

22. Verstraete MA, Moore RE, Roche M, Conditt MA. **The application of machine learning to balance a total knee arthroplasty. Bone & Joint Open 2020; 1(6):236–44.**

23. Karnuta JM, Luu BC, Roth AL, Haeberle HS, Chen AF, Iorio R et al. **Artificial Intelligence to Identify Arthroplasty Implants From Radiographs of the Knee. The Journal of Arthroplasty 2021; 36(3):935–40. Available from: URL: https://www.sciencedirect.com/science/article/pii/s0883540320311190.**

24. Polce EM, Kunze KN, Paul KM, Levine BR. **Machine Learning Predicts Femoral and Tibial Implant Size Mismatch for Total Knee Arthroplasty. Arthroplasty Today 2021; 8:268-277.e2. Available from: URL: https://www.sciencedirect.com/science/article/pii/s2352344121000066.**

EFORT Head Office
ZA La Pièce 2
1180 Rolle, Switzerland

Phone +41 (0)21 343 4400
Fax +41 (0)21 343 4411

office@efort.org
www.efort.org