Journal of Clinical Epidemiology 187 (2025) 111921

GRADE CONCEPT SERIES

GRADE concept paper 9: rationale and process for creating a GRADE Ontology

Paul Whaley[a,b,*], Brian Alper[c,d,**], Joanne Dehnbostel[c,d,e], Carlos Alva-Diaz[f,g], Stavros Antoniou[h,i], Antonio Bognanni[j], Javier Bracchiglione[k,l,m], Therese Kristine Dalsbø[n], Sean Grant[o], Jennifer Hunter[p,q], Alfonso Iorio[j], Malgorzata Lagisz[r,s], Harold Lehmann[t], Sheyu Li[u], Joerg Meerpohl[v,w], Saphia Mokrane[x], Cauê Monaco[y], Ignacio Neumann[z], Kevin Pottie[aa,ab], Shahab Sayfi[j], Nigar Sekercioglu[ac,ad], Jasvinder Singh[ae,af], Bernardo Sousa-Pinto[ag], Janice Tufte[ah], Lenny Thinagaran Vasanthan[ai], Li Wang[aj], Jun Xia[ak], Xiaomei Yao[j], Holger Schünemann[al,am,***]

[a]*Lancaster Environment Centre, Lancaster University, Lancaster, UK*
[b]*The Evidence-Based Toxicology Collaboration at Johns Hopkins Bloomberg School of Public Health, Baltimore, MD, USA*
[c]*Computable Publishing LLC, Franklin, NC, USA*
[d]*Scientific Knowledge Accelerator Foundation, Franklin, NC, USA*
[e]*Mel and Enid Zuckerman College of Public Health, University of Arizona, Tucson, AZ, USA*
[f]*Grupo de Investigación NEMECS, Neurociencias, Metabolismo, Efectividad Clínica y Sanitaria, Universidad Científica del Sur, Lima, Peru*
[g]*Servicio de Neurología, Departamento de Medicina y Oficina de Apoyo a la Docencia e Investigación (OADI) Hospital Daniel Alcides Carrión, Callao, Peru*
[h]*Department of General Surgery, Papageorgiou General Hospital, Thessaloniki, Greece*
[i]*Papageorgiou GRADE Center, Thessaloniki, Greece*
[j]*Department of Health Research Methods, Evidence, and Impact, McMaster University, Hamilton, Ontario, Canada*
[k]*Iberoamerican Cochrane Centre, Institut de Recerca Sant Pau (IR Sant Pau), CIBERESP, Barcelona, Spain*
[l]*Universitat Autònoma de Barcelona, Barcelona, Spain*
[m]*Interdisciplinary Centre for Health Studies (CIESAL), Universidad de Valparaíso, Viña del Mar, Chile*
[n]*National Institute of Occupational Health, Oslo, Norway*
[o]*College of Education, University of Oregon, Eugene, OR, USA*
[p]*Health Research Group Pty. Limited, Sydney, Australia*
[q]*Faculty of Medicine and Health, The University of Sydney, Sydney, Australia*
[r]*Evolution & Ecology Research Centre, School of Biological, Earth & Environmental Sciences, Kensington Campus, Sydney, 2052, NSW, Australia*
[s]*Department of Biological Sciences, University of Alberta, Edmonton, Alberta, T6G 2E9, Canada*
[t]*Department of Medicine, Johns Hopkins University, Baltimore, MD, USA*
[u]*Department of Endocrinology and Metabolism, MAGIC China Centre, West China Hospital of Sichuan University, Chengdu, Sichuan Province, 610041, China*
[v]*Institute for Evidence in Medicine, Medical Center & Faculty of Medicine, University of Freiburg, Freiburg, Germany*
[w]*Cochrane Germany, Cochrane Germany Foundation, Freiburg, Germany*
[x]*Département de Médecine Générale, Faculté de Médecine, Université Libre de Bruxelles, Brussels, Belgium*
[y]*Centro Universitário São Camilo, School of Medicine, São Paulo, Brazil*
[z]*School of Medicine, Universidad San Sebastian, Santiago, Chile*
[aa]*CT Lamont Centre for Primary Care, Bruyère Health Research Institute, Ottawa, Ontario, Canada*
[ab]*Department of Family Medicine, Western University, London, Ontario, Canada*
[ac]*Department of Health Research Methods, Evidence and Impact, McMaster University, Hamilton, Canada*
[ad]*Division of Nephrology, Department of Internal Medicine, Health Sciences University, Istanbul, Turkey*
[ae]*Baylor College of Medicine, One Baylor Plaza, Houston, TX, 77030, USA*
[af]*Michael E. DeBakey Veterans Affairs Medical Center, 2002 Holcombe Blvd, Houston, TX, 77030, USA*
[ag]*Faculty of Medicine, University of Porto, Porto, Portugal*

Funding: The authors did not receive for this work any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.
* Corresponding author. Lancaster Environment Centre, Lancaster University, Lancaster LA1 4YQ, United Kingdom.
** Computable Publishing LLC, 41 Labor in Vain Rd, Ipswich, MA 01938-2623, USA.
*** Clinical and Epidemiology and Research Center (CERC), Department of Biomedical Sciences, Humanitas University & IRCCS Humanitas Research Hospital, Via Rita Levi Montalcini 4, Pieve Emanuele (MI) 20072, Italy.
E-mail address: paul@whaleyresearch.uk (P. Whaley).

https://doi.org/10.1016/j.jclinepi.2025.111921

[ah]Hassanah Consulting, Seattle, WA, 98122, USA
[ai]Christian Medical College Vellore, Ida Scudder Road, Vellore, 632004, Tamil Nadu, India
[aj]Department of Anesthesia, McMaster University, Hamilton, Ontario, Canada
[ak]School of Economics, University of Nottingham Ningbo China, 199 Taikang East Road, Ningbo 315100, China
[al]Clinical Epidemiology and Research Center (CERC), Humanitas University and Humanitas Research Hospital, Milan, Italy
[am]Fraunhofer Institute for Translational Medicine and Pharmacology ITMP, Allergology and Immunology, Berlin, Germany

Accepted 31 July 2025; Published online 5 August 2025

Abstract

Context: As the rate of research production accelerates, the ability to efficiently and unambiguously communicate judgments relating to the synthesis, evaluation, and use of scientific information becomes paramount.

Perspective: Scientific information can be viewed as a ''layered infrastructure'' of data, evidence, knowledge, and use. The GRADE approach serves as a de facto data standard for this infrastructure, supporting movement between layers by reducing ambiguity in claims to knowledge (in the form of judgements of certainty in the evidence when answering research questions) and level of commitment to possible solutions to problems (in the form of strength of recommendations for interventions).

Purpose: This GRADE concept paper outlines the structure, purpose, and potential benefits of the GRADE Ontology for (a) the creators of, educators in, and users of systematic reviews, health guidelines, and health technology assessments, and (b) the development of tools that help with conducting, finding, and summarising the same. This paper also presents the processes for the development and maintenance of the GRADE Ontology, a formalised terminology standard within GRADE that will support the efficiency, rigour, consistency, and interoperability of GRADE's use.

Keywords: GRADE approach; Certainty of evidence; Evidence-to-decisions; Ontology; Data standards

Plain Language Summary

The rate of research production is increasing exponentially. It is therefore becoming increasingly important to quickly, efficiently, and unambiguously communicate the judgments made and processes used when doing research and using evidence to inform policy decisions. GRADE is a widely used approach to assessing certainty of evidence when answering research questions and making recommendations for health interventions, designed to help with the efficient and transparent evaluation and use of evidence. However, the absence of a formalized terminology standard within GRADE limits the efficiency with which the results of its use can be communicated. In response, the GRADE Ontology is being created. This concept paper outlines what an ontology is, how it helps with communicating scientific information, the specific benefits of the GRADE Ontology, and the processes for developing and maintaining a useful, valid ontology that supports the use of the GRADE approach.

1. Introduction

The purpose of this GRADE concept paper is to provide an overview of the potential benefits and development process of the GRADE Ontology. It introduces the concept of ontology and its value in the context of a ''layered infrastructure'' of scientific information, presents the rationale for the GRADE Ontology specifically, and summarizes the protocol for the development, governance, and maintenance of the GRADE Ontology. It does not present the final GRADE Ontology itself, as this will be the purpose of a future publication accompanying the release of Version 1 of the GRADE Ontology.

To develop this concept article, we followed the general methods described by GRADE [1]. An interdisciplinary project group of GRADE Working Group members developed this concept paper through discussion in web conferences, iterative revision of a draft of the concept paper, and from feedback in large (approximately 100 people in May 2023) and small group sessions (approximately 5-10 participants) at GRADE Working Group meetings. After presentation at a GRADE Working Group meeting in May 2024 (approximately 55 attendees) in which GRADE Working Group members voted to approve the paper, the GRADE Guidance Group (GGG) reviewed and then approved the concept paper.

What is new?

Key findings

- A formalized terminology standard within GRADE should support the efficiency, rigor, consistency, and interoperability of GRADE's use.

What this adds to what is known

- This paper adds to the community's understanding of the importance of ontologies for the accurate communication of scientific information, not only for entities such as diseases and genes, but also for methods for evaluating scientific information in systematic reviews and the development of health guidelines.

What is the implication and what should change now?

- This paper outlines the conceptual approach to developing the GRADE Ontology, in anticipation of the release of version 1 of the Ontology. Software and information systems developers should anticipate supporting the use of method ontologies for improving the efficiency of reuse of scientific information.

1.1. An urgent need for a computable information infrastructure

In the 21st century, society faces complex, interacting, and increasingly urgent healthcare, public, and environmental health challenges that need evidence-informed policy responses. As researchers, we are producing ever-more scientific information to address these challenges. The documentation of this research is reaching readers faster than ever, via shorter journal publication timelines and directly through preprints and press releases. This creates an increasingly chaotic information environment that is becoming exponentially harder to navigate, making it ever-more difficult to ensure policy is based on contemporary analysis of the current best evidence [2].

Ensuring that our ability to synthesize, evaluate, and use scientific information scales with the rate of its production is becoming an increasingly high priority for research organizations around the world. Accordingly, the creation of computer-based infrastructures that support the generation, storage, conversion, and communication of scientific information has become a core feature of modern scientific endeavor.

Information infrastructure can be viewed as being layered (Fig 1), moving from (1) individual units of *primary data* that describe what was done and found in a scientific study or other data collection exercise, to (2) *evidence* made up from analyzed data within studies (study evidence) and across studies (evidence synthesis), to (3) *knowledge* that is the interpretation of evidence to answer a question or achieve some other goal of understanding, to (4) *utilization* that is the use of knowledge to solve practical problems and inform decisions. Movement between layers requires *evaluation*, where the trustworthiness of data, validity of claims to knowledge, and degree of commitment to one or another solution to a problem, need articulating in a way that maintains a clear link to prior layers of the infrastructure of research (Fig 1).

1.2. The importance of data standards

Data standards are an essential contributor to communicating information across the layers of information infrastructure, required because producers and consumers of information in one layer need to be able to understand and evaluate the claims and statements being made in any other layer. We follow the Open Data Institute in categorizing data standards for research under three broad types [3]:

- ***Standards for guidance***, that provide frameworks and recommendations for creating, collecting, analyzing, or reporting data. Standards for guidance aim at improving one or more aspects of the consistency, validity, transparency, and utility of scientific studies. Units and measures, codes of practice, standard operating procedures and protocols, and reporting guidelines are types of standards for guidance.
- ***Standards for terminology***, that specify preferred terms and definitions for the concepts used in a specific area or domain. Standards for terminology aim to provide a common language for expressing and communicating the information presented in scientific studies. Controlled vocabularies and ontologies are types of standards for terminology.
- ***Standards for exchanging data***, that provide common formats and rules to support the interoperability and reusability of data. Standards for exchanging data aim to minimize loss or change in the meaning of research data as it moves between locations, users, and information systems. File formats, data transfer protocols, and data types are types of standards for exchanging data.

Data standards used in research often draw on features from more than one of the above types of standards to achieve their aims.

Without data standards, ambiguity in processes, meaning, and structure of data risks loss or change of information at every move between layers. If the interrelationships between studies are difficult to discern because their subjects, methods, and findings are described incompletely or inconsistently, then data are less readily transformed into evidence. If data relevant for understanding study findings,

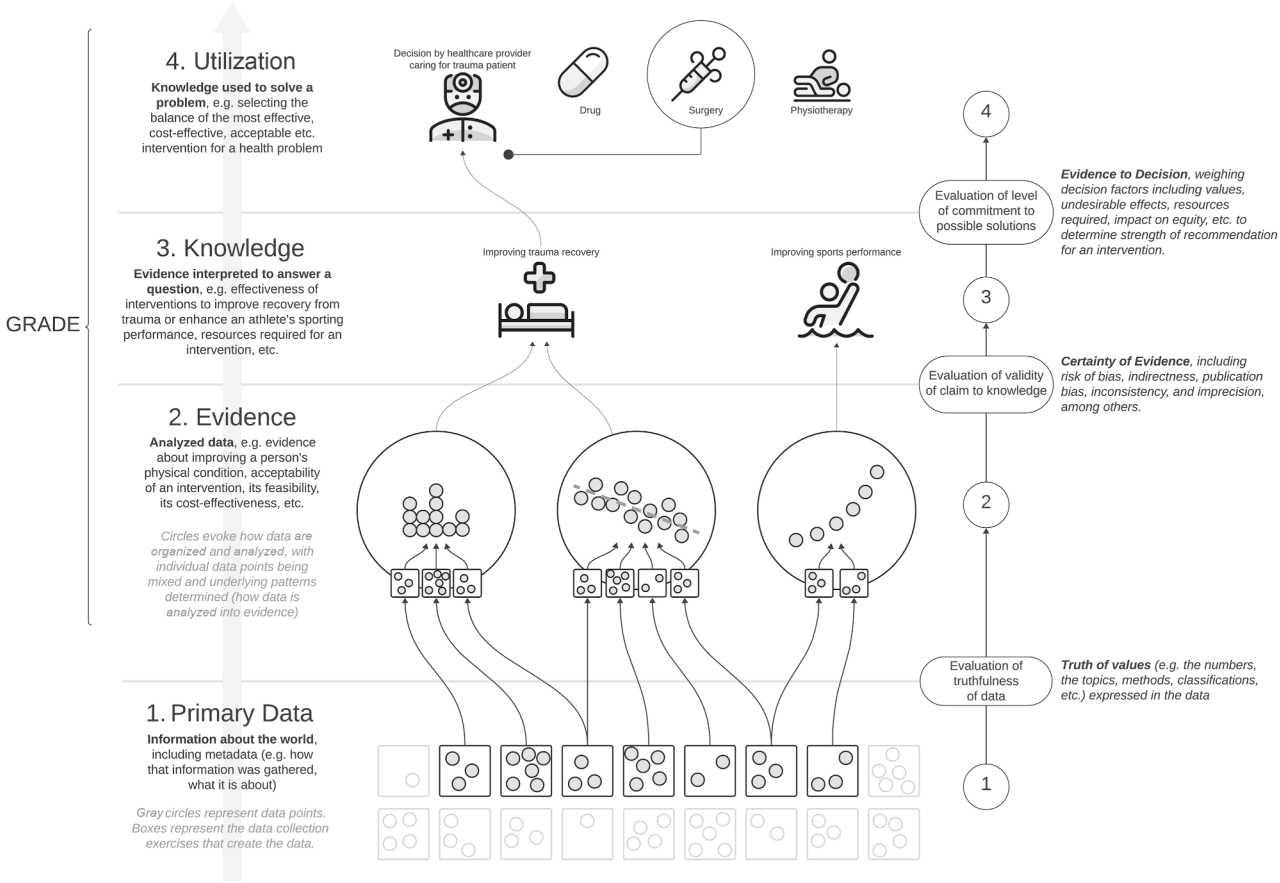

Figure 1. The layered infrastructure of research information. The infrastructure of creation, storage, conversion, and communication of information in research, represented as a flow through four layers from primary data (the raw product of research, including how the research was done) to evidence (analyzed data) to knowledge (evidence interpreted to answer a question) to utilization (knowledge used to solve a problem). What needs to be evaluated and communicated in the move between each layer (truth of values, certainty of evidence, weighing of decision factors) is shown on the right of the figure. Figure generated in Lucidchart (lucidchart.com). Icons by Streamline (streamlinehq.com).

including the interpretation of their validity, are inconsistently described or unavailable, then evidence is less readily transformed into knowledge. Finally, if information relevant to determining the appropriate level of commitment to one solution among the various that are available is inconsistently described or unavailable, knowledge is less readily applied to practical problems.

1.3. GRADE as a de facto data standard

GRADE (the Grading of Recommendations, Assessment, Development, and Evaluations) is an approach to the systematic development and presentation of summaries of evidence and health-related recommendations including health policy recommendations (book.gradepro.org). The GRADE approach consists of two elements: a Certainty of Evidence (CoE) assessment for evaluating and describing certainty in the evidence for effects on health of an intervention or exposure [4,5]; and an Evidence-to-Decision (EtD) framework for selecting a key set of prespecified but flexible criteria when developing recommendations about health interventions, health policy or coverage decisions, or occupational and environmental health policy, and for describing how strongly the recommendation is being made [6,7].

Because the GRADE approach presents a structured approach to creating data and a vocabulary for concepts, GRADE can be interpreted as being a *de facto* data standard that combines elements of standards for guidance and standards for terminology. Expressed in the terms of the layered data infrastructure described above, GRADE's CoE assessment reduces ambiguity in claims to knowledge (moving from evidence to knowledge), while GRADE's EtD framework reduces ambiguity in the degree of commitment to solutions (moving from knowledge to application).

We acknowledge that some GRADE guidance, such as about how to express judgments of certainty of evidence, has elements of a more formal standardization [8]. However, GRADE has not yet disseminated a fully developed standardized terminology. This may be a cause of inconsistency in application and interpretation of GRADE guidance, impeding successful uptake and implementation of GRADE.

The GRADE Working Group has therefore begun the process of creating the "GRADE Ontology," a human-

and machine-readable set of preferred terms, definitions for those terms, and guidance on application of those terms, that covers the core set of unique concepts under the GRADE approach.

2. The GRADE ontology

The term ‘‘ontology’’ originates in the branch of philosophy that investigates what kinds of things exist, that is, the types and structures of objects that are in the universe [9]. In 1995, the term was repurposed by computer scientists to describe any formal representation of a set of concepts within a domain and the relationships between those concepts [10,11]. Ontologies are fundamental to information infrastructure because they present a shared understanding of a domain space that can be communicated not only between people but also between people and the computer-based services, applications, and systems that can make the use of information more efficient [12].

The concepts captured in an ontology can be objects, categories, properties, or relationships. Each concept is typically denoted by a human-readable preferred term and a unique identifier. This unique identifier is essential: as a formally structured code that exclusively denotes the concept in the system, it makes the concept machine-readable (ie, processable by a computer without human involvement), and therefore allows for local variations in display of a concept while enabling a machine to recognize that the same concept is being used in each locale. Health researchers may already be familiar with ontologies such as SNOMED, which provides a consistent approach to recording patient clinical information [13], or with the Medical Subject Headings (MeSH) terms used by the US National Library of Medicine for indexing health research [14].

The GRADE Ontology will specify preferred and alternative terms for core concepts in the determination and expression of certainty of evidence and strength of recommendations for the GRADE approach, provide definitions for those concepts, and provide comments for application that will help users apply correct terminology when using GRADE concepts. Each concept will be associated with a unique machine-readable code. Terms will be related in a hierarchical logical structure with a minimum number of types of relationship between terms.

2.1. Potential benefits of the GRADE Ontology

2.1.1. Benefits to the creators and users of systematic reviews and guidelines, including educators and learners

Much work across many publications has been completed by a large community to ensure the GRADE approach is *sensible*, that is, is practical, functional, and likely to be of benefit. However, GRADE is in some aspects quite technical, complex, and has changed and evolved over time. Variation in how people interpret GRADE terms when developing systematic reviews and health guidelines may result in misapplication of GRADE guidance when making judgments, and errors in communication of the content and reasons for those judgments. This threatens to undermine the effectiveness and transparency of use of the GRADE approach.

The concept of dissemination bias illustrates how shifts in meaning and terminology can challenge users of GRADE. Historically, the term ‘‘publication bias’’ has been used for systematic differences in the evidence that is available to researchers at the time of conducting evidence reviews compared to evidence that is not available. The differences in availability arise when publishing incentives favor evidence that shows particular results compared to evidence that does not show those results (eg, positive vs negative findings). Recently, the term "publication bias" has been replaced in the GRADE approach by the term ‘‘dissemination bias’’. "Dissemination bias" covers the concept of publication bias but also includes the concept of reporting bias and other biases related to dissemination, which were previously handled under the limitations in study design and execution domain of GRADE—a domain which itself was recently renamed from ‘‘risk of bias’’.

An authoritative, actively updated ontology helps users navigate these shifts in meaning by explaining the preferred term for a concept, associating a concept with terms for other similar concepts, and explaining in a comment for application how the concept should be used.

While such a resource is not sufficient for ensuring the correct use of GRADE concepts and terms, it at least makes it more straightforward for those teaching, using, and learning about GRADE to stay up to date if they have a single reference point for doing so. For example, implementing GRADE Ontology outputs will assist in standardizing the content presented in GRADE workshops and other training initiatives, such as the Guideline International Network’s INGUIDE program [15]. Furthermore, the outcomes of the Ontology project will inform the update of the GRADE Book (https://book.gradepro.org/) and glossary, promoting a more consistent and clear educational experience.

There are also potential efficiencies in access to evidence and guidelines that could be gained in the deployment of a GRADE Ontology. Because there is a high rate of publication of guideline documents and systematic reviews that apply GRADE, it can be a significant manual burden to keep track of new publications that, due to their conclusions about certainty of evidence or strength of recommendations, have implications for standard of care or setting of threshold limits. This challenge could be addressed by absorbing the GRADE Ontology into bibliographic indexing systems such as MeSH. This would make possible the creation of GRADE-related literature filters that enable researchers to automatically identify systematic reviews with high certainty conclusions or guidelines that make strong recommendations for an

intervention. Automating the identification of such publications could reduce the lag time for translating such findings into practice.

2.1.2. Benefits to tool developers

The software application GRADEpro GDT (GRADEpro Guideline Development Tool, www.gradepro.org) has been developed to support GRADE users in consistently applying the structure and vocabulary of the GRADE approach [16]. GRADEpro GDT is an official GRADE tool and provides a controlled vocabulary insofar as users select from a preset list of terms when recording judgments made using the GRADE approach. This is a first step in using computers to reduce the cost of reusing data from applying GRADE, as the increased consistency of process outputs implies less human work needs to be done later to interpret the meaning of, correct, search for, screen, or otherwise make use of that data.

However, systematic reviewers and guideline developers also use other software applications, such as iSoF (Interactive Summary of Findings Tables, https://m-isof.epistemonikos.org/#/) or MAGICapp (the MAking GRADE the Irresistible Choice Authoring and Publishing Platform, https://magicevidence.org/magicapp), or many others. The range, output, and contextual diversity of software tools to support evidence synthesis is increasing, leading to a knowledge and application ecosystem that is rapidly growing in complexity. As for GRADEpro GDT, many of these tools provide users with sets of controlled terms from which to select when recording judgments. These software tools, more or less developed around the GRADE approach, incorporate the concepts and terminology of GRADE to varying degrees, and in many cases provide similar terms without coordination for consistent application of the concepts across tools. This lack of coordination increases the potential for the kinds of errors in judgment and communication already discussed above.

A GRADE Ontology that can be used across software tools could contribute to overcoming these challenges by allowing the following:

- For CoE and EtD support tools to send data using the same controlled vocabulary, and by association for all receiving tools to expect use of the same controlled vocabulary. This reduces the cost of and potential error from term translation and concept mapping between systems.
- Supportive content in the GRADE Ontology (eg, term definitions and comments for application) can automatically be provided in CoE and EtD support tools, making the GRADE approach easier to understand at the point of using and applying the controlled vocabulary.
- The GRADE Ontology can be stored in the form of the data exchange standard most suitable for the exchange of evidence and guidance for healthcare, namely the HL7(R) FHIR(R) standard, significantly easing the burden of writing code to transfer common concepts from one system or application to another.
- Terms from non-GRADE systems can be mapped to GRADE CoE and EtD judgments, making it easier to determine when judgments are comparable, increasing the interoperability of the overall evidence and guideline ecosystem even in relation to non-GRADE approaches.

2.2. Development and maintenance of the GRADE Ontology

The methods for the development of the GRADE Ontology were created in collaboration between the Health Evidence Knowledge Accelerator (HEvKA) GRADE Ontology Working Group and the GRADE Working Group. Development of the ontology follows a publicly available protocol [17] that is an adaptation of the protocol used to create the Scientific Evidence Code System (SEVCO) taxonomy [18]. The ontology is being developed under the GRADE Ontology Project Group.

In summary, a group of self-designated GRADE experts meet virtually each week in open public meetings to curate terms from published GRADE manuscripts, establish preferred and alternative terms for included concepts, define the concepts, and write comments that guide users in correctly applying the terms. To secure the collective feedback of an international and varied stakeholder group around the development and use of GRADE, the approach to creating the ontology involves live discussion on weekly calls and asynchronous voting and commenting between weekly calls. Any dissenting comments or votes are discussed until unanimous agreement is reached (ie, there are zero "no" votes in a round of voting). We anticipate that after publication of Version 1 of the Ontology, the same group will maintain and support the use and further development of the ontology.

The process for developing the GRADE ontology covers the following steps:

1. Defining the project team and governance;
2. Defining the scope of the ontology (ie, the concepts to be included within it);
3. Establishing preferred and alternative terms for the included concepts, defining each concept, and writing guidance for how each concept should be applied by the user;
4. Approving terms, definitions, and guidance for application;
5. Publication of Version 1 of the ontology;
6. Supporting use of the ontology; and
7. Maintenance and development of the ontology (Version 1+).

An example of a draft term is shown in Figure 2.

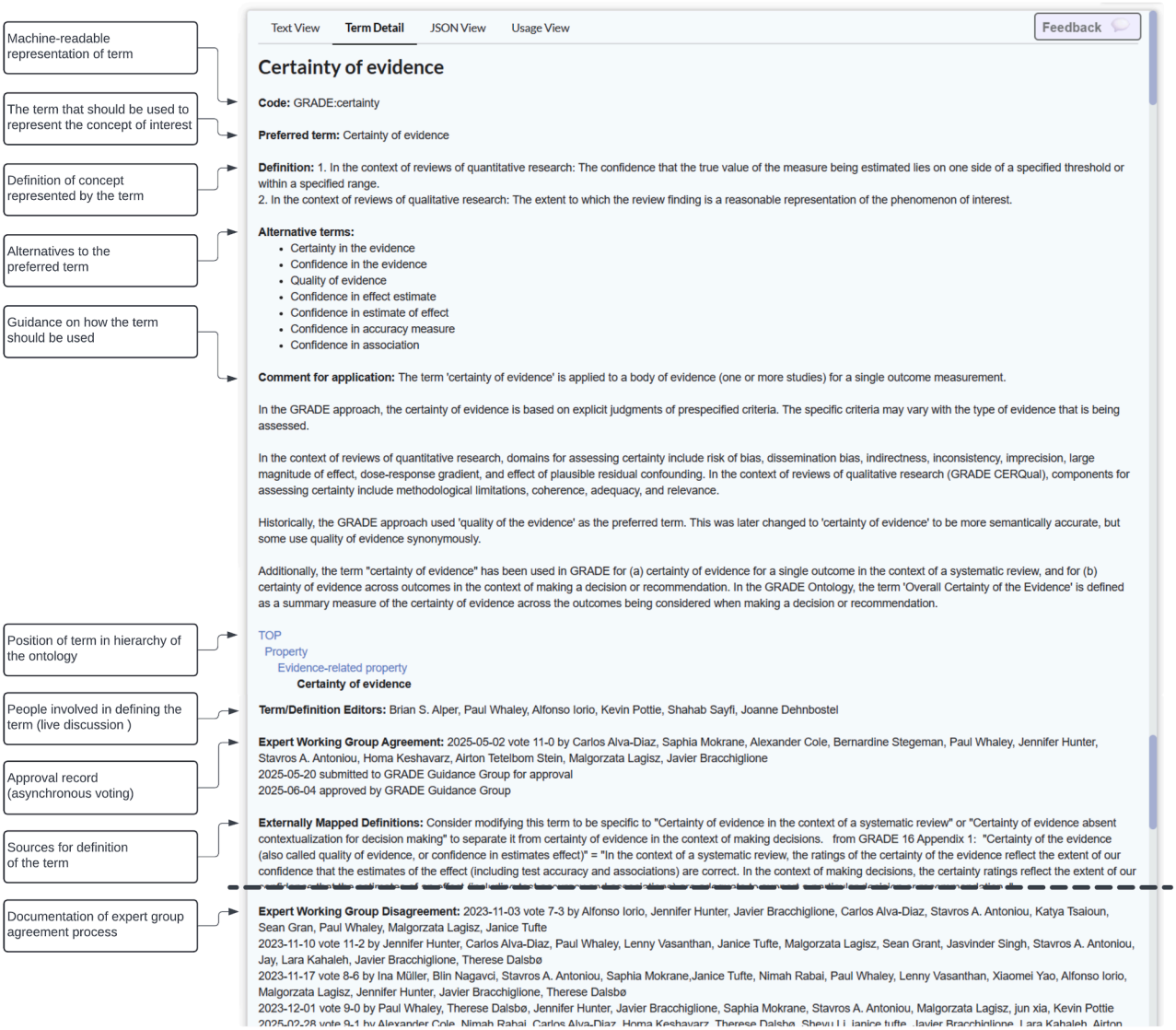

Figure 2. Example structure of a term in the GRADE Ontology as shown in the FEvIR Platform Term Viewer [19]. This consists of: the term itself; an indication of whether the term is draft (not applicable to this example as it is an approved term); the unique ID for the term; the preferred term (also shown in the display heading at top); alternative terms; the definition of the term; the comment for application that gives guidance on how to correctly use the term; the position of the term in the ontology in relation to parent and child terms (current model is hierarchical); names of term editors for giving credit; agreement record for documenting approval and giving credit to contributors; source definitions for the term; disagreement record for documenting process and giving credit to contributors.

2.2.1. Team and governance

The development of the ontology (protocol, project management, and infrastructure) is managed by the GRADE Ontology Project Group (henceforth ''Project Group''). Project Group members must be members of the GRADE Working Group and are entitled to coauthorship of any GRADE-approved manuscripts produced by the Project Group, so long as coauthorship requirements are fulfilled (see ''Publication'' section below).

The ontology itself (terms, definitions, and comments for application) is developed by the GRADE Ontology Working Group (henceforth ''Ontology Working Group''), a subworking group of the HEvKA. Membership of the Ontology Working Group is open to any interested party via an open web link (https://fevir.net/resources/Project/111563). Membership only requires registration in the Ontology Working Group system on the FEvIR Platform, which allows access to the voting software and enables the leaders of the Project Group to communicate with the members of the wider Ontology Working Group. Members of the Ontology Working Group can comment and vote on terms (asynchronous discussion) and/or participate in weekly Working Group calls where comments and votes are discussed as part of the consensus process (live discussion). Each is considered an equally valuable contribution. Ontology Working Group members need not be members

Table 1. Examples of preferred terms and term definitions for the GRADE Ontology

Preferred term	ID/Code	Definition
Certainty of evidence	GRADE:certainty	1. In the context of reviews of quantitative research: The confidence that the true value of the measure being estimated lies on one side of a specified threshold or within a specified range 2. In the context of reviews of qualitative research: The extent to which the review finding is a reasonable representation of the phenomenon of interest
Limitations in study design and execution (draft)	GRADE:02001	Issues in the design, conduct, or analysis of one or more studies that may introduce systematic error into a body of evidence
Inconsistency (draft)	GRADE:02002	Variation in the findings of individual studies from which the synthesis result was derived
Indirectness (draft)	GRADE:02003	Differences in the characteristics of the observed evidence and the characteristics of the target situation to which the observed evidence is being applied
Imprecision	GRADE:02004	The degree of variation or spread among the probable values for the estimate of effect

Terms in the table that have been approved by the Ontology Working Group but are pending final approval from the GGG are indicated as draft. Terms that have been approved by the GGG have no draft label. The codes are hyperlinked to the relevant entry in the GRADE Ontology where comments for application, voting record, source terms, and discussion can be reviewed.

Note that while a term may have been approved by the GGG, the GRADE Ontology has not yet been endorsed by the GRADE Working Group as a final product, and should therefore not be cited or otherwise treated as a final work. The draft terms are up-to-date at the time of manuscript proof corrections (2 September 2025).

of the GRADE Working Group to participate in calls or vote but must join the GRADE Working Group if they wish to be credited as coauthors of related GRADE publications.

Final adoption of the GRADE Ontology will be determined by the standard GRADE Guidance approval process (80% of GRADE Working Group members vote in favor, with final approval by the GGG).

2.2.2. Scope

Version 1 of the GRADE Ontology will consist of approximately 86 terms relating to CoE and EtD judgments. The current list of proposed terms can be viewed on the FEvIR platform at https://fevir.net/resources/CodeSystem/27833#top. Any individual can propose additional terms. Terms determined to be redundant or unnecessary will be excluded. At the time of writing, the list includes: 17 certainty of evidence domain terms (eg, ''certainty of evidence,'' ''risk of bias,'' ''inconsistency''); 6 certainty of evidence rating terms (eg, ''high certainty,'' ''moderate certainty''); 13 certainty of evidence domain rating terms (eg, ''serious concern,'' ''absent,'' ''rate down once''); 2 recommendation domain terms; 7 recommendation rating terms; 12 decision factor terms; and 28 decision factor and decision factor rating terms. The initial GRADE Ontology is being created in English. A protocol is being developed for the translation of the terms and alternative terms into other languages. Examples of draft term definitions are shown for illustration in Table 1.

2.2.3. Specifying preferred terms, definitions, and comments for application

Preferred and alternative terms, term definitions, and guidance for the application of terms will be specified by first identifying draft definitions in the GRADE literature. Each of these will be discussed and refined in weekly Ontology Working Group meetings. When the group of experts present in the meeting agree that a preferred term and its components are ready, it is put out to the whole Ontology Working Group to vote on. If a unanimous vote occurs in favor of the term with a minimum of 5 votes, the term is provisionally accepted into the Ontology. The term is then referred to the GGG for review. Terms that are unanimously agreed by all GGG members who participate in voting in this round of review will be considered accepted. Once all terms are accepted, the Ontology will be presented for approval by the GRADE Working Group.

To help progress consensus, if a person votes against a term, they are required to explain in an accompanying comment the reason for the negative vote. If negative votes are cast, the accompanying comments are discussed in the next open meeting, and the term is revised and reopened for vote. This process continues iteratively until unanimous approval is obtained. A contingency approval process can be used in the event of stalled consensus at the discretion of the experts present on a call, as described in clause 6.e.ii.3 of the protocol [17]. To reduce disincentive to vote negatively, voting is anonymous to all participants. The only people who are aware of the identity of voters are the project administrators (B.A., J.D.), who need to be aware of voter identity in order to follow-up with individual voters for clarification of any issues that may be preventing consensus.

We aim for 100% consensus during the voting process because the concepts under discussion are technical and multifaceted. Quick majorities would therefore risk missing important issues. In the experience of the Project Group, a single negative vote with accompanying explanation from

the voter often results in major changes that materially improve term definitions. We believe an exacting, iterated consensus process is essential for ensuring the utility and acceptability of the GRADE Ontology.

2.2.4. Publication and dissemination

When all terms for version 1 of the GRADE Ontology have been approved, the GRADE Ontology Project Group will develop a GRADE Guidance paper about the use and maintenance of the Ontology. This will be submitted for review and approval by the GRADE Working Group as per the general methods described by GRADE [1]. After approval, we will publish the ready-for-public-use version of the code system and seek publication of introductory articles about the GRADE Ontology in the biomedical literature. At the time of publication, attribution of contributorship will be shared by:

1. Listing as authors anyone who contributed as a ''term editor'' on any of the terms.
2. Listing as endorsers anyone who contributed as a ''voter'' on any of the final-approval votes for any terms.
3. Listing as reviewers anyone who contributed as a ''commenter'' or a ''voter'' on any of the nonunanimous votes on any of the terms.

To ensure completeness of reporting of the ontology, documentation of the release version of the ontology (v1.0) will be compliant with the MIRO Guidelines [20].

2.2.5. User testing and support for implementation

After Version 1 of the GRADE Ontology has been published, the GRADE Ontology Working Group will identify tools and systems that could use the GRADE Ontology and offer support for implementation. We also plan to measure the proportion of systems that implement the GRADE Ontology, seek to evaluate ease of use of the Ontology, and generate GRADE Ontology change requests as needed.

As a voluntary initiative without dedicated funding, at the time of writing there are no plans for extensive user testing of the GRADE Ontology. However, we acknowledge such work would be important for evaluating the extent to which the GRADE Ontology is achieving its objectives. A framework for evaluation of the GRADE Ontology could include user surveys before and after implementation of ease of use of GRADE terminology, linguistic analysis of publications by the GRADE Working Group and users of the GRADE approach for improvements in consistency and accuracy of use of GRADE terms, in-task usability testing, and adoption metrics such as the number of software tools and guidelines using the ontology or integration of the ontology into research indexes such as MeSH.

2.2.6. Maintenance and development

For ongoing maintenance and further development of the code system, we will maintain an open invitation for GRADE Ontology users to join the GRADE Ontology Working Group for recognized contribution. GRADE Ontology changes may be initiated by change requests from the community. We will maintain an open invitation for anyone to share comments regarding specific GRADE Ontology terms for continued feedback. This will include prompting authors of new GRADE papers to check the ontology for new terms, potential changes in definitions, and deprecation of redundant terms. The Project Group will validate that change requests are appropriate for group deliberation (such as fitting the purpose of the code system, having sufficient rationale, and avoiding duplication). Valid change requests will lead to drafting a preferred display term, a list of alternative terms, and a definition, with the approval process running as described above. We will also maintain an open invitation for anyone to suggest additional tools or systems that use concepts with a usefully close match to the GRADE Ontology.

3. Conclusion

The GRADE approach to making certainty-of-evidence and evidence-to-decision judgments is already a *de facto* data standard. The creation and management of a GRADE Ontology is intended to support the efficiency, rigor, consistency, and interoperability of use of the GRADE approach. The development and maintenance processes for the GRADE Ontology will facilitate discovery of terms where inconsistency, ambiguity, or misunderstanding occurs and help resolve such discrepancies via consensus methods for selecting preferred terms, defining them, and providing comments about their appropriate use. Support for the implementation of the GRADE approach will help end-users (the developers and users of systematic reviews and guidelines), software developers (to incorporate into software tools), and computers (for automated use by software tools) successfully record and communicate judgments made using the GRADE approach, further advancing the GRADE Working Group's mission of developing a common, sensible, and transparent approach to grading certainty of evidence and strength of recommendations.

Declaration of generative AI and AI-assisted technologies in the writing process

During the preparation of this work the authors used Windows 11 Copilot to create the first draft of the abstract for this manuscript. After using this service, the authors reviewed and edited the content as needed and take full responsibility for the content of the published article.

CRediT authorship contribution statement

Paul Whaley: Writing − original draft, Visualization, Project administration, Methodology, Investigation, Conceptualization. **Brian Alper:** Writing − review & editing, Software, Project administration, Methodology, Investigation, Data curation, Conceptualization. **Joanne Dehnbostel:** Writing − review & editing, Project administration, Methodology, Investigation, Data curation. **Carlos Alva-Diaz:** Writing − review & editing, Investigation. **Stavros Antoniou:** Writing − review & editing, Investigation. **Antonio Bognanni:** Writing − review & editing, Investigation. **Javier Bracchiglione:** Writing − review & editing, Investigation. **Therese Kristine Dalsbø:** Writing − review & editing, Investigation. **Sean Grant:** Writing − review & editing, Investigation. **Jennifer Hunter:** Writing − review & editing, Investigation. **Alfonso Iorio:** Writing − review & editing, Methodology, Investigation, Conceptualization. **Malgorzata Lagisz:** Writing − review & editing, Investigation. **Harold Lehmann:** Writing − review & editing, Visualization, Investigation. **Sheyu Li:** Writing − review & editing, Investigation. **Joerg Meerpohl:** Writing − review & editing, Investigation. **Saphia Mokrane:** Writing − review & editing, Investigation. **Cauê Monaco:** Writing − review & editing, Investigation. **Ignacio Neumann:** Writing − review & editing, Investigation, Conceptualization. **Kevin Pottie:** Writing − review & editing, Methodology, Investigation, Conceptualization. **Shahab Sayfi:** Writing − review & editing, Investigation. **Nigar Sekercioglu:** Writing − review & editing, Investigation. **Jasvinder Singh:** Writing − review & editing, Investigation. **Bernardo Sousa-Pinto:** Writing − review & editing, Investigation. **Janice Tufte:** Writing − review & editing, Investigation. **Lenny Thinagaran Vasanthan:** Writing − review & editing, Investigation. **Li Wang:** Writing − review & editing, Investigation. **Jun Xia:** Writing − review & editing, Investigation. **Xiaomei Yao:** Writing − review & editing, Investigation. **Holger Schünemann:** Writing − review & editing, Supervision, Investigation, Conceptualization.

Declaration of competing interest

The authors declare they have no competing interests in relation to the development of this work.

Acknowledgments

The authors would like to thank Khalid Shahin for software support.

Data availability

Data is available from the linked resources in the article.

References

[1] Schünemann HJ, Brennan S, Akl EA, Hultcrantz M, Alonso-Coello P, Xia J, et al. 'The development methods of official GRADE articles and requirements for claiming the use of GRADE - a statement by the GRADE guidance group'. J Clin Epidemiol 2023;159:79−84. https://doi.org/10.1016/j.jclinepi.2023.05.010.

[2] Greenhalgh T, Howick J, Maskrey N, Evidence Based Medicine Renaissance Group. 'Evidence based medicine: a movement in crisis?'. BMJ 2014;348:g3725. https://doi.org/10.1136/bmj.g3725.

[3] Open Data Institute. Types of open standards for data, Open Standards for Data Guidebook. 2023. Available at: https://standards.theodi.org/introduction/types-of-open-standards-for-data/. Accessed August 17, 2023.

[4] Guyatt G, Oxman AD, Akl EA, Kunz R, Vist G, Brozek J, et al. 'GRADE guidelines: 1. Introduction-GRADE evidence profiles and summary of findings tables'. J Clin Epidemiol 2011;64(4):383−94. https://doi.org/10.1016/j.jclinepi.2010.04.026.

[5] Guyatt GH, Oxman AD, Vist GE, Kunz R, Falck-Ytter Y, Alonso-Coello P, et al. 'GRADE: an emerging consensus on rating quality of evidence and strength of recommendations'. BMJ 2008;336(7650):924−6. https://doi.org/10.1136/bmj.39489.470347.AD.

[6] Alonso-Coello P, Schünemann HJ, Moberg J, Brignardello-Petersen R, Akl EA, Davoli M, et al. 'GRADE evidence to decision (EtD) frameworks: a systematic and transparent approach to making well informed healthcare choices. 1: introduction'. BMJ 2016;353:i2016. https://doi.org/10.1136/bmj.i2016.

[7] Senerth E, Whaley P, Akl E, Beverly B, Alonso-Coello P, Rooney A, et al. 'GRADE guidance 40: the GRADE evidence-to-decision framework for environmental and occupational health'. Environ Int 2025;197(109314):109314. https://doi.org/10.1016/j.envint.2025.109314.

[8] Santesso N, Glenton C, Dahm P, Garner P, Akl EA, Alper B, et al. 'GRADE guidelines 26: informative statements to communicate the findings of systematic reviews of interventions'. J Clin Epidemiol 2020;119:126−35. https://doi.org/10.1016/j.jclinepi.2019.10.014.

[9] Hofweber T. 'Logic and Ontology', The Stanford Encyclopedia of Philosophy. Summer 2023. In: Zalta EN, Nodelman U, editors. Metaphysics Research Lab, Stanford University; 2023. Available at: https://plato.stanford.edu/archives/sum2023/entries/logic-ontology/. Accessed February 1, 2025.

[10] Gruber TR. 'Toward principles for the design of ontologies used for knowledge sharing?'. Int J Human Computer Stud 1995;43(5):907−28. https://doi.org/10.1006/ijhc.1995.1081.

[11] Whetzel PL, Noy NF, Shah NH, Alexander PR, Nyulas C, Tudorache T, Musen MA. 'BioPortal: enhanced functionality via new Web services from the National Center for Biomedical Ontology to access and use ontologies in software applications'. Nucleic Acids Res 2011;39:W541−5. https://doi.org/10.1093/nar/gkr469.

[12] Whaley P, Edwards SW, Kraft A, Nyhan K, Shapiro A, Watford S, et al. 'Knowledge Organization Systems for Systematic Chemical Assessments'. Environ Health Perspect 2020;128(12):125001. https://doi.org/10.1289/EHP6994.

[13] Stearns MQ, Price C, Spackman KA, Wang AY. SNOMED clinical terms: overview of the development process and project status. Proc AMIA Symp 2001662−6.

[14] Lowe HJ, Barnett GO. 'Understanding and using the medical subject headings (MeSH) vocabulary to perform literature searches'. JAMA 1994;271(14):1103−8. https://doi.org/10.1001/jama.1994.03510380059038.

[15] Schünemann HJ, Nieuwlaat R. 'The INGUIDE International Guideline Training and Certification Programme'. Clin Public Health Guidel 2024;1(1):e12008. https://doi.org/10.1002/gin2.12008.

[16] McMaster University and Evidence Prime. GRADEpro GDT: GRADEpro Guideline Development Tool. 2022. Available at: gradepro.org. Accessed February 1, 2025.

[17] Alper B, Dehnbostel J, Whaley P, GRADE Ontology Project Group. Protocol for development of the GRADE Ontology. Zenodo. 2024. Available at: https://doi.org/10.5281/zenodo.11002448.

[18] Alper BS, Dehnbostel J, Afzal M, Subbian V, Soares A, Kunnamo I, et al. 'Making science computable: developing code systems for statistics, study design, and risk of bias'. J Biomed Inform 2021;115(103685):103685. https://doi.org/10.1016/j.jbi.2021.103685.

[19] Alper BS. *FEvIR®: codesystem builder/viewer* [fast evidence interoperability resources (FEvIR) platform, FOI 29874]. 2024. Available at: https://fevir.net/resources/Project/29874. Accessed July 1, 2025.

[20] Matentzoglu N, Malone J, Mungall C, Stevens R. MIRO: guidelines for minimum information for the reporting of an ontology. J Biomed Semant 2018;9:6. https://doi.org/10.1186/s13326-017-0172-7.