A rule-based framework for risk assessment in the health domain

Luca Cattelani, Federico Chesani, Luca Palmerini, Pierpaolo Palumbo, Lorenzo Chiari, Stefania Bandinelli

A rule-based framework for risk assessment in the health domain

In International Journal of Approximate Reasoning,Volume 119,2020,p. 242-259

A Rule-based Framework for Risk Assessment in the Health Domain

Luca Cattelani[a,*], Federico Chesani[a], Luca Palmerini[b], Pierpaolo Palumbo[b], Lorenzo Chiari[b], Stefania Bandinelli[c]

[a] *Department of Computer Science and Engineering, University of Bologna, Viale Risorgimento 2, 40136 Bologna, Italy.*
[b] *Department of Electrical, Electronic, and Information Engineering "Guglielmo Marconi", University of Bologna, Viale Risorgimento 2, 40136 Bologna, Italy.*
[c] *Local Health Unit Tuscany Center, Firenze, Piero Palagi Hospital, Viale Michelangelo 41, 50125 Firenze, Italy.*

Abstract

Risk assessment is an important decision support task in many domains, including health, engineering, process management, and economy. There is a growing interest in automated methods for risk assessment. These methods should be able to process information efficiently and with little user involvement. Currently, from the scientific literature in the health domain, there is availability of evidence-based knowledge about specific risk factors. On the other hand, there is no automatic procedure to exploit this available knowledge in order to create a general risk assessment tool which can combine the available quantitative data about risk factors and their impact on the corresponding risk. We present a Framework for the Assessment of Risk of adverse Events (FARE) and its first concrete applications FRAT-up and DRAT-up, which were used for fall and depression risk assessment in older persons and validated on four and three European epidemiological datasets, respectively. FARE consists of i) a novel formal ontology called On2Risk; and ii) a logical and probabilistic rule-based model. The ontology was designed to represent

*Corresponding author
Email addresses: luca.cattelani@unibo.it (Luca Cattelani), federico.chesani@unibo.it (Federico Chesani), luca.palmerini@unibo.it (Luca Palmerini), pierpaolo.palumbo@unibo.it (Pierpaolo Palumbo), lorenzo.chiari@unibo.it (Lorenzo Chiari), stefania.bandinelli@uslcentro.toscana.it (Stefania Bandinelli)

Preprint submitted to International Journal of Approximate Reasoning January 10, 2020

qualitative and quantitative data about risks in a general, structured and machine-readable manner so that this data may be concretely exploited by risk assessment algorithms. We describe the structure of the FARE model in the form of logic and probabilistic rules. We show how when starting from machine-readable data about risk factors, like the data contained in On2Risk, an instance of the algorithm can be automatically constructed and used to estimate the risk of an adverse event.

Keywords: depression, falls, formal ontology, logical rule-based system, missing data, risk assessment

1. Introduction

Risk assessment methodologies aim to estimate the risk that a specific situation will occur. Generally speaking, a risk is characterized in terms of the magnitude of the potential loss related to the occurrence of the situation [1], and in terms of the probability that the situation/loss will occur. In the health domain, a commonly considered risk type directly relates the concept of potential loss to the happening of a specific event, named the *adverse event*. Examples of adverse events (for simplicity, events, from now on) are the development of cardiovascular diseases [2], of Type II diabetes [3], falls in the elderly [4], and development of depression [5].

In epidemiological research, a common activity consists of identifying which are the *risk factors* associated with the happening of an event. A vast scientific literature documents which risk factors are (statistically) related with the appearance of a pathology within a certain time span; e.g., smoking is often referred to be a risk factor for cardiovascular diseases [6]. Although a number of risk assessment tools are available [7, 8], to the best of our knowledge, few or none of them exploit the huge amount of scientific literature that relates risks and risk factors.

In this paper we introduce the Framework for the Assessment of Risk of adverse Events (FARE), a framework supporting the evaluation of risks in the health domain. FARE exploits the existing literature in a field and provides as output the probability (the *risk score*) that a subject will experience an event within a given time span. To this end, FARE provides two major contributions: a formal ontology (On2Risk) for representing the available knowledge about a specific risk; and a methodology that computes a subject-specific risk by taking as input the mentioned ontology and subject-specific

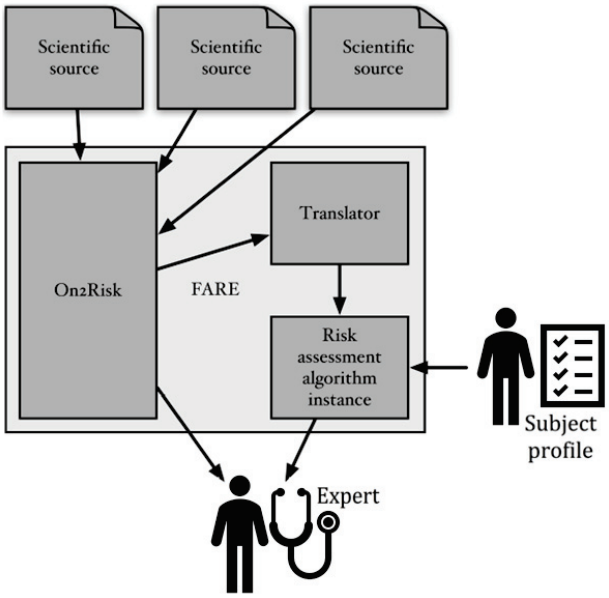


Figure 1: FARE main components and usages.

information (the subject's *profile*).

On2Risk organizes the existing knowledge in a structured manner: in this way knowledge can be easily processed by automatic software (e.g. [9, 10]). It supports important notions such as *Risk Factors*, *Odds Ratios* (statistical knowledge about the relation between the risk and the specific factor), *Estimators* (how risk factors are evaluated), *synergies* between Risk Factors (the concurrent exposure to multiple risk factors further increases the overall risk), and management of missing information (some data might be missing from the subject's profile). While we are aware of previous ontologies for medical data (see Section 2.1), to the best of our knowledge none of them includes such a rich and quantitative representation of risk factors to be directly used by a risk assessment tool.

The usage of FARE follows a two-step process. The first step takes as input an On2Risk instance with knowledge about a specific risk, and produces a probabilistic logic program that is able to compute the risk. The second step, subject-dependent, takes as input a subject's profile, feeds the logic program, and returns the subject's probability of experiencing the event within a defined time span (Figure 1). While the second step is executed one or more times for each subject, the first step is executed only when the

On2Risk instance changes, i.e. whenever novel insights are gained about a specific risk.

There are other possible approaches to risk assessment in the health domain, which can use rule-based procedures such as belief rule-based risk assessment tools [11], statistical techniques such as regression and survival analysis [12, 13], and machine learning tools, as in [14].

FARE has a number of advantages with respect to these other risk assessment implementations. First, it is generalizable: it can be applied to many different health risks, provided the knowledge about those risks is available. Methods like the ones based on machine learning instead, because of the need of a learning phase, are not directly generalizable. Second, it is not based on a specific dataset, but rather exploits the existing literature, thus not requiring parameter estimation (for statistical techniques and belief rule-based procedures) or learning phases (for machine learning). Therefore it is independent of the specific cohort. Third, it manages incomplete information about the subject (in clinical practice it can happen that not all the required subject's data are available). Finally, the first concrete applications of FARE, FRAT-up and DRAT-up, were successfully applied to assess fall and depression risks in the elderly, and they compared favorably with respect to state-of-the-art risk assessment tools [15, 16, 17, 18]. The FARE probabilistic logic rules formulation, the On2Risk ontology, and the architecture of the framework are novel contributions of this paper.

This paper is organized as follows. A brief background about the adopted technologies (ontologies and probabilistic logic programs) is given in Section 2. In Section 3 we describe On2Risk and its main concepts. Section 4 reports the mathematical foundation of FARE and then describes how a Probabilistic Logic Program (PLP) can effectively encode the risk evaluation algorithm. Overall implementation of the framework, together with glimpses of how a PLP is automatically generated starting from an On2Risk instance, are described in Section 5. Previous work on health risk evaluation, a brief glance on the falls and depression domains, and a summary of the validation results of FRAT-up and DRAT-up are in Section 6 and discussed in Section 7.

2. Background

In the following, we provide a brief introduction on the technologies that we exploited in the FARE framework. We do not provide an exhaustive

description, but rather we sketch the advantages of the chosen tools, and point the interested reader towards more detailed documentation.

2.1. Formal ontologies

A formal ontology is a conceptualization of the knowledge about a domain by means of a formal language [19]. A number of ontology languages exist and have been classified depending on their type of logic [20]. Description Logics (DLs) [21] in particular are a family of logic languages used to represent the knowledge of an application domain in a structured way, where the important notions of the domain are described by *concepts* and *binary relations* among them. Usually, the information contained in an ontology can be distinguished into the *terminological box* TBox, containing the definition of the concepts important for the domain, and the *assertion box* ABox containing facts and properties related to the *individuals* being part of the domain. For example, the notions of "Person" and "Student" would be represented as concepts. The relation "every student is a person" would be part of the TBox. Finally, "John is a student" would be an assertion in the ABox. The Semantic Web Initiative[1] has defined languages, standards, tutorials and best practices for representing logic ontologies, and it is a first starting point for the interested reader. A number of tools exist to help users to properly express the domain. We used Protégé[2], a free ontology editor supporting W3C standards.

In FARE the starting point is the knowledge contained in the scientific literature related to a specific risk. Unfortunately, almost the totality of scientific papers have their content organized for human readers: i.e., they convey their content by means of images, tables and natural language. Conversely, in the epidemiological field the use of ontologies has been recognized as an important step [22]. In the broader health sector, we mention the SNOMED-CT[3] ontology for precise definition of clinical health terms, and the Gene Ontology initiative[4].

2.2. Probabilistic Logic Programs

Rule-Based Systems (RBS) allow to define the available knowledge in terms of rules of the form *antecedent* $\rightarrow$ *consequent*, whose meaning can

[1] https://www.w3.org/standards/semanticweb/
[2] http://protege.stanford.edu/
[3] http://www.snomed.org/snomed-ct
[4] http://www.geneontology.org/

be intended as "if the antecedent is satisfied, then the consequent is satisfied too". An *inference engine* exploits such knowledge to perform some computation.

Logic Programming (LP) [23] is based on First Order Logic (FOL), and exploits the rules (namely, *clauses*) in a backward manner. Rules are often represented in the form $head \leftarrow body$, whose intended meaning is that *head* is true if *body* is true. Rules may have the form of $head \leftarrow true$: in such cases the antecedent is always true, and the consequent is interpreted as a *fact*. Given a program P (a set of clauses) and a goal $\mathcal{G}$, an LP inference engine looks to prove that $\mathcal{G}$ is a logical consequence of P, i.e. that $P \models \mathcal{G}$. One of the most well known LP languages is Prolog [23]: a number of dialects, interpreters and compilers are available, supporting different program semantics, and providing a rich set of software libraries for many common tasks (from exposing web-based services compatible with modern web standards, to the definition of grammars and the automatic construction of language parsers). Indeed, Prolog is considered a fully fledged programming language.

Probabilistic Logic Programs (PLP) are an extension of LP, where also probabilities can be taken into account. A widely adopted semantics for PLP is the distribution semantics [24], where a probabilistic logic program defines a distribution on non-probabilistic logic programs called *worlds*. The probability distribution of a query in a probabilistic program is obtained by marginalization, from the query in the various worlds together with the probabilities of the worlds. In our framework we opted for *Logic Programming with Annotated Disjunctions* (*LPAD*) [25], a type of PLP that extends Prolog and allows the presence of disjunctions in the *head*. We follow the syntax of the *cplint* implementation [5] . Syntactically, the rules have the following

[5] *cplint* is currently available in three different flavors:

- a YAP-based implementation `http://ds.ing.unife.it/~friguzzi/software/cplint/manual.html` ;
- a SWI-based implementation `http://ds.ing.unife.it/~friguzzi/software/cplint-swi/manual.html` ;
- and as a web-app available online `http://cplint.ml.unife.it/p/sublist_sldnf.pl`

where YAP (`https://www.dcc.fc.up.pt/~vsc/yap/`) and SWI (`https://www.swi-prolog.org/`) are two Prolog interpreters and compilers.

form:

$$h_1 : p_1;\ h_2 : p_2;\ \ldots;\ h_n : p_n : -\ b_1,\ b_2,\ \ldots,\ b_m.$$

Where $h_1 : p_1;\ h_2 : p_2;\ \ldots;\ h_n : p_n$ are the head disjuncts, while $b_1,\ b_2,\ \ldots,\ b_m$ are the body conjuncts. The intended meaning is that each h_i has the probability p_i of being true if the body of the rule is true. We will skip all the caveats of the LPAD syntax. We shall recall only that terms starting with capital letters are variables.

We want to express our risk assessment algorithm in a probabilistic logic language, not only because these languages are particularly well suited to represent logical and probabilistic relationships, but also because every step of reasoning of their engines is clearly explainable, and they interact easily with formal ontologies, that are written in logical languages too. We opted for LPAD because it has a more general syntax than other similar languages [26], and is particularly well suited to represent probabilistic relations between causes and effects, where the causes for the same effect (e.g. risk factors for a risk) are causally independent. In a previous work [15], we presented the mathematical ground for our model, including how it exploits the assumption of causal independence.

3. On2Risk

On2Risk is an ontology about risks. It is defined in the OWL 2 Web Ontology Language [27] and is open access and freely downloadable[6]. Its aim is to organize scientific knowledge about risk factors, including quantitative information that can be automatically extracted and used by risk assessment algorithms. The concept of risk is mapped to the `Risk` class. An instance of this class is e.g. `StrokeIn1Year`, the probability to suffer a stroke at least once during the next year. In On2Risk we represent a scientific source with an individual (instance) of the class `Reference`. Each reference has annotations attached such as title or DOI identifier, following the Dublin Core standard [28].

A key concept in On2Risk is the `RiskFactor`. It is introduced in Section 3.1, together with a quantitative measure of its impact, diagnostic methods, a classification by value type, and a description of how missing information on its exposure is handled. On2Risk also rules the way risk factors are

[6] http://ffrat.farseeingresearch.eu/on2Risk

computed (Section 3.2). A comprehensive description of On2Risk is included as supplementary material.

3.1. Risk factors

A risk factor is a characteristic of subjects that is reported in scientific literature to be significantly associated with an augmented risk (e.g. `Diabetes` and `Age` are associated with several health risks). To support risk assessment algorithms, On2Risk includes a quantitative measure of the impact of a risk factor on a risk. Various measures can be used. We focus on the odds ratio, since in epidemiology it is the most commonly used measure to quantify the effect of being exposed to a risk factor. Other impact measures may be supported in a similar manner.

To represent the fact that there may be more than one procedure to estimate a risk factor, and that the exposure to a risk factor is a general property of a subject independently of how it is measured, we introduced the class `Estimator`. Estimators are used to evaluate risk factors. They may be as simple as a direct observation or a question to the subject (as in the case of age and sex) or more convoluted, like a CES-D questionnaire for depression (Center for Epidemiologic Studies Depression Scale [29]), that is an estimator for the risk factor `Depression`. Estimators and risk factors are the only concepts used to describe a subject in our formalism. In the following, we refer to them as "features".

On2Risk does not include specific subjects with specific feature values, but defines names, characteristics and relationships of features. Some features answer to typical yes/no questions, such as "is the subject affected by Parkinson disease?", while others expect a numerical value, like "number of drugs". To represent missing information, which can be frequent in the health domain, each feature instance may be assigned the value `unknown`. Each feature can be "ternary" or "scalar with unknown". A ternary feature may take a three-valued logic assignment: "true", "false", or "unknown". A "scalar with unknown" feature may take a numerical value belonging to a finite interval, or be assigned "unknown". Following this distinction, `RiskFactor` is subclassed by `TernaryRiskFactor` and `ScalarRiskFactor`, while `Estimator` is subclassed by `TernaryEstimator` and `ScalarEstimator`. In the following for brevity with "scalar" we will mean "scalar with unknown".

We want risk assessment algorithms based on On2Risk to be able to cope with missing data. To this end we take into account the possible values of an unknown factor in the population of interest, according to its prevalence (i.e.

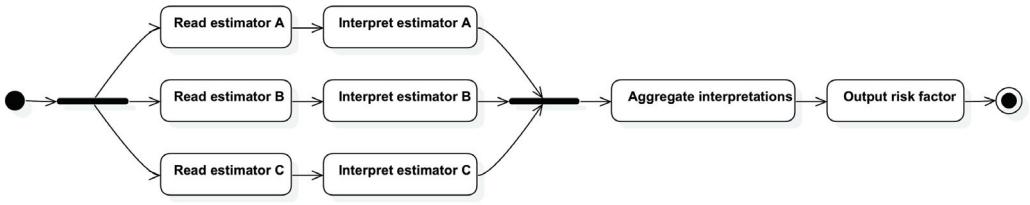


Figure 2: UML activity diagram showing operations of interpretation and aggregation.

the distribution of the values of the risk factor on a reference population). On2Risk includes the proportion of people affected by a dichotomic risk factor and the proportion of people affected by each level of a scalar risk factor.

3.2. Computing risk factors

The concepts of estimators and risk factors are tightly coupled with the methods that lead from the values of one or more estimators to the value of a risk factor. We want these methods to be quantitative functions so that this information may be interpreted and used automatically by risk assessment algorithms. In scientific literature it is also possible to find statistical data about risk factors that are derived from other risk factors. To use that information, we also map the procedures to obtain the value of a risk factor starting from the values of one or more other risk factors.

Instances of these procedures belong to the class `ToRiskFactor`. We call *direct risk factor*, part of the class `DirectRiskFactor`, a risk factor that is obtained by one or more procedures `ToRiskFactor` that start from estimators. We call *indirect risk factor*, part of the class `IndirectRiskFactor`, a factor obtained starting from other risk factors.[7]

Procedures for direct risk factors (with estimators as input) are composed of two phases: a first phase where interpretation functions are applied to single estimators, and a second phase where the interpretations of all the involved estimators are aggregated to produce a single output value for the risk factor. The output of an interpretation is used as input to an aggregation function (Figure 2).

[7] An indirect risk factor may be produced only with direct risk factors as input, and the two sets `DirectRiskFactor` and `IndirectRiskFactor` are a partition of `RiskFactor`. This structuring forbids cyclic dependencies of derivation between risk factors and provides a valid representation for the risk factors we considered in our practical applications.

Before continuing with the presentation of the domain concepts, let us discuss a simple example to better clarify why we need to distinguish between estimators, their interpretation, the aggregation of such interpretations and the consequent mapping into a risk factor. Let us consider the domain of *falls*: from Table 6 in [30], we get to know that *vision impairment* has been associated with an augmented risk of falling. In our terminology, we would say that *vision impairment* is an instance of the `TernaryRiskFactor` class (the subject might suffer of vision impairment, might not, or it could be unknown). However, how could we decide that a subject suffers/does not suffer of vision impairment? Again from the medical literature, we get to know that at least three different methods are used to assess vision impairment: the *visual acuity* test (aka the Monoyer's scale), the *visual stereognosis* test and the *contrast sensitivity* test. Each of these tests is an instance of the `Estimator` class. Of course, each test has its own meaning and should be properly interpreted: for the visual stereognosis, an outcome value lower than 4 indicates visual impairment; for the visual acuity, an outcome value lower than 6 indicates visual impairment. The `EstimatorInterpretation` instances take care of such difference, providing the right understanding for each different estimator. Finally, what if the sight capabilities of a subject have been tested with all the three mentioned tests? Obviously, if the subject failed all the three tests, we would say that she/he suffers of vision impairment only once, and not that she/he is thrice-visual-impaired! An instance of `AggregationOfInterpretations` class would take care exactly of such aspect, by aggregating different estimator interpretations towards the same risk factor.

Interpretations belong to the class `EstimatorInterpretation` and are partitioned in the `TernaryInterpretation` and `ScalarInterpretation` subclasses, depending on the interpretation output. The interpretation input is a ternary or scalar estimator. When the estimator is ternary, the interpretation function is often an identity function. In these cases the ternary estimator is used directly as the interpretation, so `TernaryEstimator` is a subclass of `TernaryInterpretation`.

Another typical case involves a threshold applied to a scalar estimator. In this case the input is scalar while the output is ternary. `EstimatorInequality-Interpretation` is a subclass of `TernaryInterpretation` capturing this concept. It has the properties

- `isAboutScalarEstimator exactly 1 ScalarEstimator`,

- `Inequality exactly 1 inequalityOperator`,
- `ScalarEstimatorValue exactly 1 int`.

All code excerpts follow the OWL 2 Manchester syntax [31]. The inequality may be "=", ">", "≥", "<", or "≤". The interpretation of an instance of an estimator is true if its value respects the given inequality with the given "ScalarEstimatorValue".

To cover the cases where there is an interpretation of a scalar estimator that maintains the scalar nature but renders it more coarse grained, we introduce the specialization of `ScalarInterpretation` called `DiscreteLevelsEstimatorInterpretation`. It has the following properties:

- `isAboutScalarEstimator exactly 1 ScalarEstimator`,
- `Step1Start exactly 1 int`,
- `StepSize exactly 1 int`, and
- `LastStep exactly 1 int`.

The output is 0 for estimator values less than `Step1Start`, is 1 when corresponding to `Step1Start` and increases by 1 for each increase of `StepSize` of the estimator value, up to a maximum output value of `LastStep`. E.g. `AgeDiscreteLevels` is defined with `Step1Start` 70, `StepSize` 5 and `LastStep` 4. This means that the output value is 1 for each input value between 70 and 74, 2 between 75 and 79, and so on. `DiscreteLevelsEstimatorInterpretation` covers also the cases when the estimator value starts from 0 or from 1 and must not change after the interpretation. It is sufficient to set `Step1Start` as 1, `StepSize` as 1 and `LastStep` as the maximum estimator value.

The second step to compute a risk factor starting from one or more estimators is to aggregate the interpretations. An aggregation function is expressed with an instance of the class `AggregationOfInterpretations`, specializing `ToRiskFactor`. It has the property `hasOutputRiskFactor exactly 1 DirectRiskFactor`. It may be specialized to represent a wide range of functions, like AND, min, sum, mean, and so on. We present two specializations, denoting an OR between ternary estimators (in Kleene's three-valued logic [32]) and a max between scalar estimators. The max function when all inputs are known numbers produces the maximum between them, while if there are

unknown inputs produces an unknown output. These aggregation functions when having a single input are identity functions.

E.g. Let us consider again the example related to the visual impairment and the risk of falling. As mentioned earlier, various tests target different visual functionalities, but all diagnose a visual impairment when their score is low. We model their aggregation with `AcuityContrastStereognosisToVisualImpairment` where the risk factor `VisualImpairment` is given by applying the OR operator to three inequalities, on `VisualAcuity`, `ContrastSensitivity` and `VisualStereognosis`. `AcuityContrastStereognosisToVisualImpairment` is an individual of class `LogicalOREstimatorsToFactor`, subclass of `AggregationOfInterpretations`. It has the properties:

- `hasEstimatorInequality VisualAcuityInequality`,
- `hasEstimatorInequality ContrastSensitivityInequality`,
- `hasEstimatorInequality VisualStereognosisInequality`, and
- `hasOutputTernaryRiskFactor VisualImpairment`.

Indirect factors are computed with functions represented by members of the class `FactorsToFactors` that take direct factors as input, and have the property `hasOutputRiskFactor exactly 1 IndirectRiskFactor`.

3.3. Synergic factors

In some cases, the presence of more risk factors determines a risk that is higher than what would be expected by taking into account the contributions of the individual risk factors [33]. In other terms, the presence of more risk factors acts as an additional risk factor. On2Risk models this effect with the class `SynergicFactorsToFactor`, a specialization of `FactorsToFactors`. It represents a function that receives two or more direct ternary risk factors as input and outputs an indirect scalar risk factor, as expressed by the properties:

- `hasOutputRiskFactor exactly 1 IndirectScalarRiskFactor` and
- `hasSynergicRiskFactor min 2 DirectTernaryRiskFactor`.

The intended semantic is that the output scalar risk factor takes, for a specific subject, a value that is the number of the input factors to which he/she

is affected.
E.g. Suppose there is the need to represent a synergy between dizziness and vertigo, Parkinson, history of stroke and depression. It is possible to define an instance of `SynergicFactorsToFactor` called `ToComorbidity`, with properties:

- `hasOutputRiskFactor ComorbidityRiskFactor`,
- `hasSynergicRiskFactor DizzinessAndVertigo`,
- `hasSynergicRiskFactor Parkinson`,
- `hasSynergicRiskFactor HistoryOfStroke`, and
- `hasSynergicRiskFactor Depression`.

If a subject is affected by Parkinson and Depression but not by the other two factors, he/she has a value of 2 for `ComorbidityRiskFactor`.

4. A Rule-based model and computation of risk

On2Risk has been designed to support algorithms for quantitative risk assessment. Quantitative impact of risk factors on a risk is expressed by means of odds ratios. The most commonly expected output of a risk assessing algorithm is a probability of the event, in a given time period. On2Risk represents known data about risk and is not linked to a specific algorithm. The algorithm we present follows a rule based approach. We describe the function used to assess the risk and how we compute its parameters starting from risk factor odds ratios (Section 4.1). The whole process of risk assessment, starting from estimator values, can be implemented with logical and probabilistic rules. Having a whole mapping of the algorithm in PLP allows for easier interpretability of the model and faster prototyping of new features. The principal data flow is depicted in Figure 3. We present such an implementation based on a three-layered architecture (Section 4.2).

4.1. From ORs to Probabilities

The mathematical foundation of our approach is expressed in an equation that computes the probability of happening of an event starting from the odds ratios of the risk factors. Its full derivation is presented in [15].

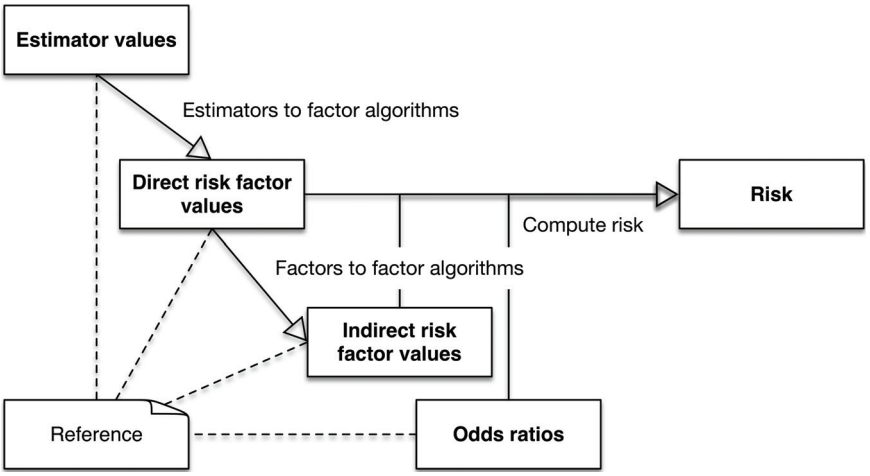


Figure 3: Simplified diagram depicting how risk is computed starting from estimators. Each quantitative value has one or more scientific references.

For simplicity we will restrict here to the case where all risk factors are dichotomous without unknown values. Let $E_0, E_1, \ldots, E_n$ be $n+1$ dichotomous random variables, representing risk factors, with values in $\{0; 1\}$ and $E = (E_0, E_1, \ldots, E_n)$. We say that the i^{th} risk factor is present if $E_i = 1$. d is the event. The odds ratio relative to the risk factor E_i, with $i = 1, \ldots, n$, is defined as

$$OR_i := \frac{P\left(d|E_i = 1\right)}{1 - P\left(d|E_i = 1\right)} \frac{1 - P\left(d|E_i = 0\right)}{P\left(d|E_i = 0\right)}.$$

We define C_0 as the event probability relative to a subject with none of the accounted risk factors. The event probability given the information on the subject's risk factors is computed as

$$P\left(d|E\right) = 1 - (1 - C_0) \prod_{i=1}^{n} \left(1 - E_i C_0 \frac{OR_i - 1}{1 - C_0 + C_0 OR_i}\right). \tag{1}$$

We call

$$C_i = C_0 \frac{OR_i - 1}{1 - C_0 + C_0 OR_i} \tag{2}$$

the *contribution* of the i^{th} risk factor to the risk.

This methodology may be applied also to scalar and synergy risk factors

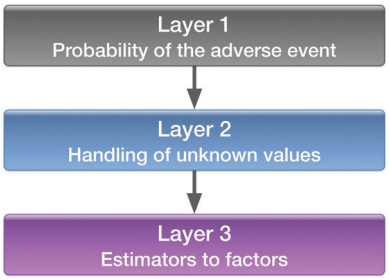


Figure 4: Layers of the risk assessment algorithm.

and to inputs with unknown values, as detailed in [15].

4.2. Three-layered architecture

In this and the following subsections we describe a scheme of logical and probabilistic rules by which it is possible to code the FARE risk assessing algorithm. With our approach being a rule-based one, the mathematical model expressed in Section 4.1 was designed to have a natural translation in form of rules.

The whole algorithm computes the probability of happening of the event, in the given time interval, starting from a list of estimator values. We structure it in three layers (Figure 4).

1. The first layer computes the probability of happening of the event starting from risk factor values without "unknown". That is, from risk factors expressed in Boolean values or with scalars without "unknown". It is described in Section 4.3.
2. The second layer removes the `unknown` values, using risk factor prevalences. The procedure is described in Section 4.4.
3. The third layer computes risk factor values starting from the values of the estimators. Unknown values in estimators may cause unknown values in risk factors. The procedure is described in Section 4.5.

4.3. First-layer: dealing with the probability of the adverse event

Rules of the first layer compute the probability of the event starting from risk factor values without "unknown". The semantics are identical to those of equation (1) and fit naturally in LPAD rules.

One rule is used to express the fixed base probability of the event, called C_0 in (1). The parameter C_0 is found by imposing that the probability of

the event for a subject with all estimators `unknown`, is equal to the prevalence of the event in the population (i.e., the adverse event probability). We express the probability to suffer an "event E" of a subject `S` with the term `adverseEventE(S)`. If, for example, the base probability C_0 of the event E is 0.075, it is defined by the following rule.

```
adverseEventE(S) : 0.075.
```

For each Boolean-valued risk factor, its contribution is defined in equation (2) and is taken into account with a rule. Equation (2) includes two parameters: OR_i and C_0. In FARE, the odds ratio are extracted from On2Risk, that is filled with data from scientific sources. For instance, if the odds ratio of "risk factor R" to the "adverse event E" is 2.32, from Equation (2) its contribution is 0.09 and is factored with the following rule.

```
adverseEventE(S) : 0.09 :- factorB(S, 'risk factor R', t).
```

Term `factorB(S, 'risk factor R', t)` is true when the presence of the risk factor R is true for subject `S`. In our coding `t` stands for `true`, `f` for `false` and `u` for `unknown`.

A scalar risk factor with K+1 levels (without loss of generality, from 0 to K) is coded as a set of K dichotomous risk factors. Namely, the N^{th} of these coding dichotomous risk factors is true if the scalar risk factor is at least at level N. The following example shows how a scalar risk factor `factor T` contributes to the risk of event E.

```
adverseEventE(S) : 0.01 :- factorS(S, 'factor T', N), number(N), N >= 1.
adverseEventE(S) : 0.01 :- factorS(S, 'factor T', N), number(N), N >= 2.
adverseEventE(S) : 0.01 :- factorS(S, 'factor T', N), number(N), N >= 3.
```

Where `factorS(S, 'factor T', N)` unifies when subject `S` has risk factor T at level `N`. The contribution 0.01 of each dichotomous risk factor is calculated from the odds ratio of one unit increase of factor T as from Equation (2).

The synergy between two or more dichotomous risk factors is implemented with a set of rules, one to list the risk factors involved and one for each possible level of the synergy. The following example contains rules to define

the participating factors and the quantitative contribution of a synergy factor called "synergy F".

```
synergy('synergy F', L) :- L =
   ['factor A', 'factor B',
    'factor C', 'factor D'].
adverseEventE(S) : 0.015 :-
   synergy('synergy F', L), at_least_true_factors(S, L, 2).
adverseEventE(S) : 0.015 :-
   synergy('synergy F', L), at_least_true_factors(S, L, 3).
adverseEventE(S) : 0.015 :-
   synergy('synergy F', L), at_least_true_factors(S, L, 4).
```

`at_least_true_factors(S, L, N)` is true if subject `S` is affected by at least `N` risk factors from the list `L`. The contributions of the levels of a given synergy have all the same magnitude. Thus, the level of a synergy affects the computation of the risk in a way that is analogous to the level of a scalar risk factor.

4.4. Second-layer: dealing with unknown values

The second layer, when encountering a known value, simply passes it up to the first layer. When encountering an `unknown` value, it splits the remaining of the computation, leveraging LPAD expressiveness, into two or more classes of worlds. Each class of worlds has one of the possible values of the risk factor in place of the unknown one. The probability of each class of worlds is multiplied by the probability of the corresponding risk factor value. While the first layer depends on the risk to be assessed, the second and third layers have the objective of assigning "known" values to risk factors and are independent of the specific risk of interest.

When a ternary factor is `t` or `f`, it is simply passed up as a factor in Boolean values with the corresponding assignment (`t` or `f`). The following 3 rules show how the Parkinson disease risk factor is converted from a three-valued form to a Boolean-valued form, with two rules that directly pass up the `true` and `false` values, and a third rule that, when the original value is `unknown`, assigns `true` with 0.01 probability and `false` with `0.99` probability. This prevalence may be read from On2Risk and is ultimately extracted from

a scientific source [34].

```
factorB(S, 'parkinson', t) :- factor3(S, 'parkinson', t).
factorB(S, 'parkinson', f) :- factor3(S, 'parkinson', f).
factorB(S, 'parkinson', t) : 0.01 ;
factorB(S, 'parkinson', f) : 0.99 :- factor3(S, 'parkinson', u).
```

The term `factorB` is used when factor values are Boolean. `factor3` has similar meaning, but factor values are ternary.

Factors with values that are scalar with `unknown`, are simply passed up when the value is a known number. In case of `unknown` the probability is distributed between the possible levels, including level 0, according to the prevalence. The following example is about the risk factor `age`.

```
factorS(S, 'age', N) :- factorSU(S, 'age', N), number(N).
factorS(S, 'age', 0) : 0.25 ; factorS(S, 'age', 1) : 0.25 ;
   factorS(S, 'age', 2) : 0.20 ; factorS(S, 'age', 3) : 0.16 ;
   factorS(S, 'age', 4) : 0.14 :- factorSU(S, 'age', u).
```

`factorS` is used when the factor value is scalar without possibility of `unknown`, whereas `factorSU` is used when `unknown` is a possible value.

4.5. Third-layer: dealing with estimators

The third layer contains algorithms that, starting from the values of one or more input estimators, compute the exposure to a risk factor, with the possibility of `unknown` values.

For some factors an identity function is enough, like in the following example.

```
factor3(S, 'parkinson', X) :- member(estimator3('parkinson', X), S).
```

Where `estimator3('parkinson', X)` means that the estimator named `parkinson` is associated, in three-valued logic, with the value `X`, that can be `t`, `f`, or `u`. `S` is the subject the check is made on.

A scalar risk factor may result from an interpretation that discretizes a scalar estimator, as seen in the description of On2Risk (Section 3.2). Such an interpretation is obtained with 4 rules: a pair of rules handle cases resulting in a risk factor with level 0 or with the maximum level, a rule for cases resulting in intermediate levels, and a fourth that applies when the estimator is unknown, resulting in an unknown risk factor. The following example

shows how the risk factor `age` is computed starting from the estimator with the same name.

```
factorSU(S, 'age', 0) :-
   member(estimatorSU('age', N), S), number(N), N < 70.
factorSU(S, 'age', 4) :-
   member(estimatorSU('age', N), S), number(N), N >= 85.
factorSU(S, 'age', L) :-
   member(estimatorSU('age', N), S), number(N),
   N >= (70 + (5 * (L - 1))), N < (70 + (5 * L)).
factorSU(S, 'age', u) :- member(estimatorSU('age', u), S).
```

`factorSU(S, 'age', L)` represents that subject `S` is affected by the risk factor named `age` with level `L`, where `L` is a scalar with the possibility of being unknown. We might notice that the rules above have to be mutually exclusive: if not, a wrong probability would be calculated. Such property is guaranteed by the fact that these rules are not directly defined by the developer, but rather they are generated automatically by an algorithm, starting from the ontology.

A scalar estimator may be used to compute a three-valued risk factor applying a threshold. Three rules are used to state that: the factor is `true` when the threshold is satisfied, the factor is `false` when the threshold is not satisfied, and the factor is `unknown` when the estimator is equally undetermined. The following example shows the rules inferring the value of the `depression` risk factor from the `CESD` estimator.

```
factor3(S, 'depression', t) :-
   member(estimatorSU('CESD', N), S), number(N), N > 20.
factor3(S, 'depression', f) :-
   member(estimatorSU('CESD', N), S), number(N), N =< 20.
factor3(S, 'depression', u) :-
   member(estimatorSU('CESD', u), S).
```

The call `member(estimatorSU('CESD', X), S)` unifies X to the value of estimator `CESD`. `SU` stands for "scalar valued with the possibility of being unknown".

As seen in Section 3.2, an aggregation may compute a risk factor value using one of a variety of rules such as logical OR and arithmetic max, that present one or more interpretations. In cases where there is just one interpretation it is not necessary to represent the OR or max explicitly, so the code may be simplified using the interpretation directly, as in the previous

examples. When there are two or more interpretations instead, the aggregation must be coded explicitly. We explain how the OR kind of aggregation is implemented. The max may be implemented in a similar way. The implementation has 3 parts.

1. The first part contains a rule for each estimator that checks if the interpretation is `true`. If at least one interpretation is `true`, the factor is also `true`.
2. The second part similarly has one rule for each estimator, this time checking two conditions: if the factor is not already `true` (following a unification in a rule of the first part) and if the interpretation is `unknown`. If a rule of the second part unifies, the factor is `unknown`.
3. The third part has just one rule that checks if the factor has not an assigned value of `true` or of `unknown` and by exclusion assigns `false`.

This implementation is compliant with Kleene's three-valued logic [32], the same logic chosen for On2Risk (Section 3.1).

In the following example the factor `vision impairment` is computed starting from 3 estimators: `visual acuity 3 m`, `contrast sensitivity`, and `visual stereognosis`. To each estimator is applied a threshold interpretation and the interpretations are aggregated by logical OR.

```
factor3(S, 'vision impairment', t) :-
   member(estimatorSU('visual acuity 3 m', N), S), number(N), N =< 5.
factor3(S, 'vision impairment', t) :-
   member(estimatorSU('contrast sensitivity', N), S), number(N), N =< 16.
factor3(S, 'vision impairment', t) :-
   member(estimatorSU('visual stereognosis', N), S), number(N), N =< 3.

factor3(S, 'vision impairment', u) :-
   member(estimatorSU('visual acuity 3 m', u), S),
   !, \+ (factor3(S, 'vision impairment', t)).
factor3(S, 'vision impairment', u) :-
   member(estimatorSU('contrast sensitivity', u), S),
   !, \+ (factor3(S, 'vision impairment', t)).
factor3(S, 'vision impairment', u) :-
   member(estimatorSU('visual stereognosis', u), S),
   !, \+ (factor3(S, 'vision impairment', t)).

factor3(S, 'vision impairment', f) :-
   \+ (factor3(S, 'vision impairment', t)),
   \+ (factor3(S, 'vision impairment', u)).
```

5. Implementation of the system

FARE allows the user to assess the risk of a subject whose data is manually inserted, run analyses on datasets in batch mode, compile and use an LPAD risk assessing program or run the assessment directly in the Java environment. The system may be seen as composed of 2 main parts.

1. A core Java application that creates the risk assessing algorithm, accesses data sets of subjects, runs the assessment directly or produces an LPAD program able to compute the risk, and translates the provided information about subjects in Prolog. The subjects are then ready to be assessed by the LPAD program.
2. An LPAD environment where the LPAD version of the risk assessment is executed, with Prolog scripts to run the assessment on whole sets of subjects and aggregate the results.

For manual insertion of subject data, it is possible to use a Web interface, like the one available for FRAT-up [8]. For a study on its usability see [35].

The core application of FARE is a Java program that includes a number of features.

1. Composing a Java object that represents a program computing the probability of an event, "assessment program object" from now on. The input is a source with quantitative data about estimators and risk factors, and with functions from estimators to risk factors, like the ontology On2Risk described in Section 3, that can be accessed with a query language like SPARQL.
2. Constructing/Generating an LPAD program that computes the probability of an event, starting from an assessment program object of point (1).
3. Reading estimators or, directly, risk factors about sets of subjects from a database or from a comma-separated values (csv) file.
4. Writing Prolog rules defining the estimator values of a set of subjects from point (3), to be feed to a risk assessment LPAD program of point (2).
5. Deriving the values of a subject risk factors starting from his/her estimators, supplying the estimators to factor functions contained in an assessment program object of point (1).

[8] http://ffrat.farseeingresearch.eu

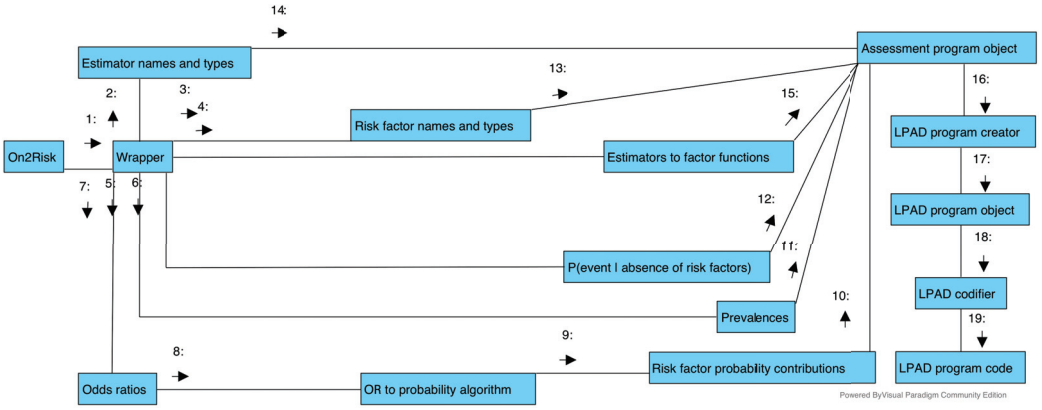


Figure 5: UML Collaboration Diagram of LPAD production steps. Messages are numbered in a consistent chronological order.

6. Computing the probability that an event will happen to a subject starting from the subject's risk factor values, applying an assessment program object from point (1).
7. Applying the assessment to sets of subjects data (subjects' profiles) and computing various quality measures, like AUC and Brier score [36].

In feature (2), the LPAD rules for risk assessment are automatically generated from the Java application. The process of producing and writing the LPAD program goes through a sequence of steps as depicted in the UML Collaboration Diagram of Figure 5 and encompasses the composition of an assessment program object of feature (1). First a data source, like On2Risk, is read through a wrapper and data are extracted; they include (a) the names and types of the risk factors and the estimators; (b) the prevalences of the risk factors; (c) the probability of the event in absence of risk factors C_0; (d) the odds ratios associated to the risk factors OR_i; and (e) for each risk factor a function that specifies how it is computed starting from one or more estimators.

Applying the mathematical model described in Section 4.1, the probability contribution C_i associated to each risk factor is computed starting from the probability of the event in absence of risk factors C_0 and the odds ratio of the risk factor OR_i, using equation (2). This is required because our approach is based on probabilities (it is indirectly based on odds ratios). The probability contributions together with data (a), (b), (c), and (e) compose the assessment program object of feature (1).

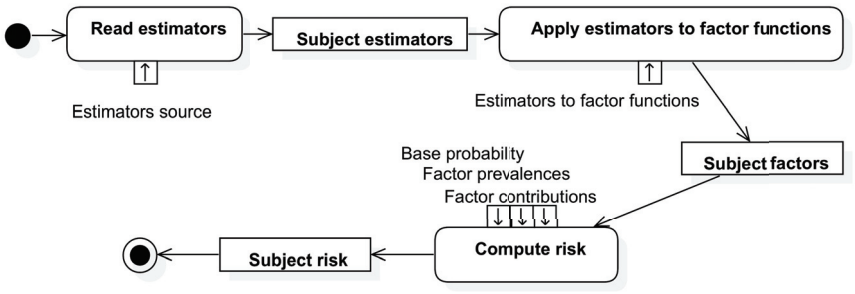


Figure 6: UML Activity Diagram of FARE risk assessment.

In the next step an LPAD program creator algorithm uses the assessment program object to build a representation in a composite Java object of an LPAD program, a tree that mirrors the syntactical tree of a well-formed LPAD program. Finally in the last step an LPAD codifier writes the actual LPAD program according to customizable formatting rules.

The core application permits to assess the risk of a subject directly in Java, starting from his/her values of the relevant estimators. Data contained in an assessment program object from feature (1) is used first to compute risk factor values starting from estimator values (feature (5)), and then to compute the event probability, starting from risk factor values (feature (6)). The whole procedure is depicted in the UML Activity Diagram of Figure 6. The resulting event probability is equal to applying the LPAD rules described in Section 4, and the diagram applies to them too, looking at the rule layers backwards from layer 3 to layer 1.

Subject estimators are received from a dataset or from user input as in the FRAT-up Web application. Using the estimator to factor functions, subject's risk factors are computed, possibly with unknown values. Using the probability of the event in absence of risk factors, the prevalence of the risk factors, and their contributions to the risk, the unknown values are handled and the probability of the event is computed, according to Equation (1).

Both the LPAD and the Java implementations have good computational performances in typical cases where almost all the subject factors are known, with a computation lasting less than 15 milliseconds on a low cost machine (e.g. on a MacBook Pro 2010, 2.4 GHz Intel Core i5). When the subject factors contain many unknown values the computation is slower, requiring in the LPAD implementation more than 10 seconds in the worst case where all estimators are unknown using FRAT-up. This is due to the fact that

the presence of unknown values makes the problem harder (a brute force implementation would compute a probability of the event for each of the possible assignments of the unknown factors). The LPAD interpreter features general purpose optimizations and is a lot faster than a brute force algorithm, but still lacks an important domain specific optimization, that instead we could deploy in the Java version [9].

6. Validation of the approach on two health-related domains

Tools for risk evaluation are key components of preventive medicine, as they inform prevention policies about people at higher risk. They have been developed for several specific medical conditions, e.g. cardiovascular diseases [2] or type 2 diabetes [3]. Two risk prediction models (RPMs) have been developed according to the framework presented in this study for two health-related domains: FRAT-up, a fall risk assessment tool for community-dwelling older adults [37, 15, 16], and DRAT-up, a depression risk assessment tool for community-dwelling subjects aged 60-75 [17]. They have been validated using four and three large European epidemiological datasets, respectively. The datasets had different sets of risk estimators. In order to enable FARE to operate on them, for each dataset we only had to write a different set of third-layer rules (without changing the other two layers). These rules were written before testing FARE and without being influenced by its performance.

6.1. Validation on fall risk

Falls are the most burdensome cause of injury among older adults[10]. Not only are they responsible for physical injuries and consequent physical disability, but also for an increase in fear of falling, and a decrease in self-confidence, physical activity and social participation.

[9] This optimization exploits properties of statistical independence: it can be proved that it is not necessary to split the problem in two or more possible worlds for each unknown variable, but it is sufficient to split on the factors that are part of a synergy (11 out of 26 in FRAT-up, 0 out of 5 in DRAT-up). This reduces the computation by orders of magnitude and allows to compute the risk in less than 15 milliseconds even in the worst case in the FRAT-up Java version.

[10] The Global Burden of Disease Study 2010. `http://www.healthdata.org/gbd`

Table 1: AUC (95% CI) for FRAT-up on different datasets.

ActiFE	ELSA	InCHIANTI	TILDA
0.562	0.699	0.636	0.685
(0.530, 0.595)	(0.680, 0.718)	(0.594, 0.681)	(0.660, 0.709)

A number of risk factors have been identified to contribute to proneness to falling (e.g. cognition impairment, visual loss, etc.) [30, 38] and various preventive programs have been shown to be effective [39, 40, 41].

The odds ratios of the fall risk factors were taken from a review and meta-analysis [30] and the parameter C_0 was calculated assuming an average probability of falling of 0.31 [42].

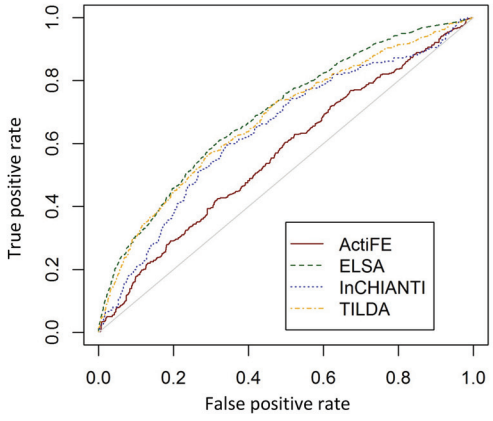
We have compared its performance against common indicators of fall risk and data-driven models. Common fall risk indicators were history of falls, gait speed, and the Short Physical Performance Battery [18], which can be used by FRAT-up as risk factor estimators. Data-driven models were: a Poisson Lasso regression trained on an extensive number of variables of the InCHIANTI dataset (NCT01331512) [18] and stepwise logistic regressions trained on FRAT-up risk factors and their two-way interactions of four datasets of epidemiological studies of older adults [16]: Activity and Function in the Elderly in Ulm (ActiFE-Ulm) [43, 44], English Longitudinal Study of Ageing (ELSA) [45], InCHIANTI [46], and The Irish Longitudinal Study on Ageing (TILDA) [47, 48].

Regarding validity, results have shown that FRAT-up significantly outperforms simple fall risk indicators and is equivalent to the data-driven Poisson Lasso regression. Namely, the AUC (95% confidence interval (CI)) for FRAT-up was 0.638 (0.610–0.666), for the Poisson Lasso was 0.639 (0.611–0.667), and for gait speed was 0.594 (0.566–0.622) [16].

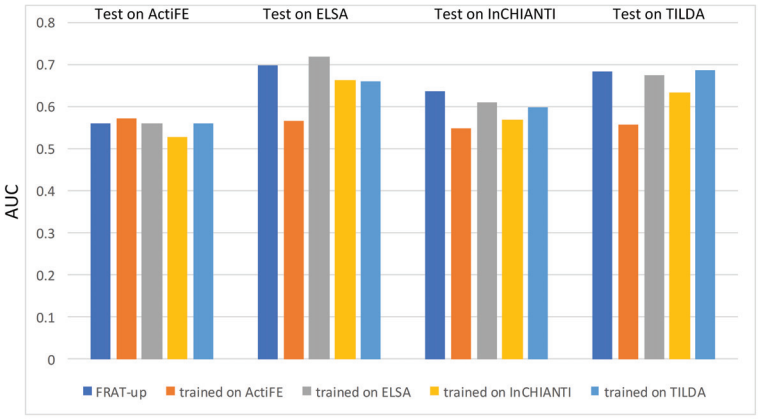
On the four different datasets, FRAT-up attained different values of AUC, ranging approximately from 0.56 to 0.70 as reported in Table 1. On each dataset, FRAT-up AUC was always greater than the AUC of data-driven models trained on external datasets (Figure 7).

6.2. Validation on depression risk

In recent years, the importance of prevention and early identification of depression has been increasingly recognized [49]. Late life depression is particularly disabling: it causes personal and familial suffering, heightens



(a)



(b)

Figure 7: Results from FRAT-up validation on four epidemiological datasets of older adults. (a) Receiver Operating Characteristic curves for FRAT-up. (b) AUC of FRAT-up and four cohort-specific risk models trained with stepwise logistic regression. More detailed numerical results can be found in [16].

Table 2: DRAT-up performance. 95% confidence intervals in brackets.

	ELSA	InCHIANTI	TILDA
AUC	0.761	0.736	0.768
	(0.746, 0.775)	(0.703, 0.769)	(0.717, 0.815)
Brier score	0.054	0.133	0.041
	(0.052, 0.056)	(0.118, 0.148)	(0.036, 0.045)

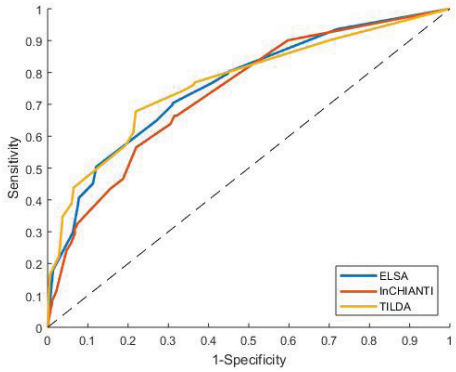


Figure 8: ROC curve of DRAT-up on the three validation datasets.

suicide risk, worsens the outcomes of associated physical conditions and increases healthcare costs [50]. Treatments include antidepressant drugs [51] or psychosocial interventions [52]. Clinical outcomes remain suboptimal due to lack of resources and frequent under-diagnosis [53].

DRAT-up was validated on 24689 samples, relative to 11704 individuals, from the three datasets ELSA, InCHIANTI, and TILDA. The probability contributions C_i were derived from the odds ratios reported in [54] and $C_0 = 0.061$ was derived from [55].

The AUC and the Brier score attained by DRAT-up on the 3 datasets, including 95% confidence intervals (CI), are reported in Table 2. The AUC ranged from 0.736 in InCHIANTI to 0.768 in TILDA. DRAT-up was characterized by a fair performance in all three datasets, with similar Receiver Operating Characteristic (ROC) curves, displayed in Fig. 8.

A previous RPM, by Okamoto and Harasawa, was based on a stepwise linear discrimination analysis of a sample of 754 subjects aged 65 and older[56]. We compared DRAT-up with the model developed by Okamoto and Harasawa on the ELSA and TILDA datasets. We judged that the InCHIANTI

dataset could not be harmonized to calculate the Okamoto score. The AUCs for the Okamoto score on the ELSA and TILDA datasets are respectively 0.672 (95% CI 0.657-0.690) and 0.683 (95% CI 0.628-0.735). Thus, upon direct comparison with a previously-proposed RPM, DRAT-up showed a higher discriminative ability.

6.3. Validation conclusions

The rules derived from the literature produced valid fall and depression risk assessment tools. Their predictive accuracy was greater than commonly-used risk indicators (e.g. gait speed for fall risk), previously-proposed models (the Okamoto score for risk of depression), or data-driven models trained on external datasets. In this regard, we highlight that since our risk models were not trained on the datasets we used for validation but were derived from literature knowledge, our results have external validity. Furthermore, our approach proved to be adaptable and to work on different domains and on different datasets, which where not specifically designed to study fall nor depression. Finally, it was able to seamlessly cope with missing data, an ubiquitous issue that affects real-world and research-based datasets. All these represent clear advantages over other similar tools.

7. Conclusion and Discussion

Risk assessing protocols have been experimented with differing results in numerous application fields including ecology [57], suicide [58], cardiovascular diseases [59], health as a whole [60], and many others.

In this article we presented a general approach for risk assessment. This approach is characterized by an ontology (On2Risk) that structures the available scientific literature about risk factors. The ontology also stores information that allows to handle missing data and to quantify synergies between risk factors. The starting point of our approach is always some medical knowledge already published and accepted by the medical community. Such knowledge is usually made available through scientific papers and books, i.e., it is represented in natural language texts, images, tables summarizing odds ratio, bibliographic references, etc. Although easily accessible by human health experts, such knowledge is not automatically exploitable through an algorithm. The On2Risk ontology is our proposal for a suitable formal representation of such a knowledge.

The adoption of On2Risk brings another advantage. An ontology-based approach allows any user (or any algorithm) to inspect the knowledge risk tools are based on, by simply navigating the ontology itself: in this sense, On2Risk might be seen as an answer to the urgent need of having explainable, transparent AI tools, especially in the health field.

Formal ontologies are not the only way for representing knowledge. E.g., we could have represented the knowledge directly by means of a PLP. While from a computer science perspective such alternative sounds plausible and reasonable, we might notice that formal ontologies (with various degrees of expressiveness) are becoming a standard in the health domain.

In any case, our approach requires the extraction of machine-readable data from sources that are only human-readable. Creating and updating a formal ontology requires human work and sometimes a huge effort. This is an unavoidable limitation of our approach: when building a risk tool for a new domain, the developer should identify proper knowledge, possibly merging different knowledge sources, and then proceed to create a formal representation. Such process is known to be time-consuming. Conversely, the lack of a training phase is a clear practical advantage.

The ontology is the input (other inputs could be possible in the general approach) to a probabilistic rule-based system. The mathematical model underneath the rule-based system was designed to have a natural translation in form of rules. The scheme of the rules was written in the LPAD language. The presented system has a three-layered architecture that from a list of risk estimators computes the probability of happening of the event. This can be done also in the presence of unknown values of risk factors.

We detailed and explained FARE characteristics in the light of its first practical implementations, FRAT-up and DRAT-up. The chosen fields of application are fall and depression risk assessments, both crucial because of the severe related problems, like death, disability, hospitalization and quality of life reduction. FRAT-up and DRAT-up were validated with good results in a total of seven different assessments, with data from four different datasets, showing the effective FARE modularity. In fact, in order to apply the tools to different datasets, only a single layer had to be re-written (leaving the higher-level layers unchanged). The performance of FRAT-up and DRAT-up was compared to the performance of other tools showing the effectiveness of the proposed risk assessment method. Despite its excellent promises, the effectiveness of the method in other health-related areas remains to be explored.

It should be noted that the FRAT-up and DRAT-up implementations are based only on information present in scientific literature and no learning is made on the available data. We believe that adding data-driven capabilities (automatically learn from data) to these implementations could allow to exploit further potentially useful data sources (like wearable devices) which could result in future performance improvements. Solutions have been developed in the past for learning the probabilities of the single risk factors. For instance, approaches based on Inductive Logic Programming ([61] is among the earliest, to the best of our knowledge) have been proposed, while recently also solutions based on Expectation Maximization algorithm [62] have appeared. These approaches focus on learning the *parameters*, hence On2Risk might be exploited as a background knowledge (once it has been stripped of odds ratio). Interestingly, in [63] a solution for learning the *structure* of the program (i.e., the rules) is presented: this could open up the possibility of mixing previous knowledge coming from On2Risk with new knowledge (still in the form of logic rules) learned from other data sets, by introducing (learning) novel risk factors.

8. Conflict of Interest

L. Palmerini and L. Chiari have a significant financial interest in mHealth Technologies, a company that may have a commercial interest in the results of this research. All other authors declare no competing interests. All authors are aware of the submission of this manuscript.

9. Acknowledgments

The research leading to these results has received funding from the European Union Seventh Framework Programme (FP7/2007-2013) under grant agreement FARSEEING No. 288940.

The authors would like to thank the teams of the InCHIANTI, ActiFE, ELSA, and TILDA studies for their datasets. Special thanks to Flavio Zuffa for his work on On2Risk and its SPARQL connection and to Julia Marshall Leach for her tips on the linguistic style.

The analyses, views and interpretations expressed in this paper are of the authors, and responsibility for any errors or omissions rests with the authors alone.

The InCHIANTI Study is registered under ClinicalTrials.gov NCT01331512 (The InChianti Follow-Up Study); http://clinicaltrials.gov/show/NCT01331512 . It is sponsored by the National Institute on Aging (NIA).

The ActiFE-study was supported by grants from the Ministry of Science, Research and Arts, state of Baden-Wuerttemberg, Germany, as part of the Geriatric Competence Center, Ulm University.

The data relative to the English Longitudinal Study of Ageing (ELSA) were made available through the UK Data Archive (UKDA). ELSA was developed by a team of researchers based at the NatCen Social Research, University College London and the Institute for Fiscal Studies. The data were collected by NatCen Social Research. The funding is provided by the National Institute of Aging in the United States, and a consortium of UK government departments co-ordinated by the Office for National Statistics.

TILDA is an inter-institutional initiative led by Trinity College Dublin. TILDA data have been co-funded by the Government of Ireland through the Office of the Minister for Health and Children, by Atlantic Philanthropies, and by Irish Life; have been collected under the Statistics Act, 1993, of the Central Statistics Office. The project has been designed and implemented by the TILDA Study Team. ©Department of Health and Children. Copyright and all other intellectual property rights relating to the data are vested in TILDA. Ethical approval for each wave of data collection is granted by the Trinity College Research Ethics Committee. TILDA data is accessible for free from the following sites: Irish Social Science Data Archive (ISSDA) at University College Dublin http://www.ucd.ie/issda/data/tilda/ ; Interuniversity Consortium for Political and Social Research (ICPSR) at the University of Michigan http://www.icpsr.umich.edu/icpsrweb/ICPSR/studies/34315 .

This study has been conducted despite the regulatory (Italian law 240/2010) and financial [64] problems of the Italian research system.

10. Bibliography

[1] T. Aven, Risk assessment and risk management: Review of recent advances on their foundation, European Journal of Operational Research 253 (1) (2016) 1–13. doi:10.1016/j.ejor.2015.12.023.
URL http://dx.doi.org/10.1016/j.ejor.2015.12.023

[2] J. S. Berger, C. O. Jordan, D. Lloyd-Jones, R. S. Blumenthal, Screening for cardiovascular risk in asymptomatic patients., Journal of the American College of Cardiology 55 (12) (2010) 1169–77. doi:10.1016/j.

jacc.2009.09.066.
URL http://www.ncbi.nlm.nih.gov/pubmed/20298922

[3] D. Noble, R. Mathur, T. Dent, C. Meads, T. Greenhalgh, Risk models and scores for type 2 diabetes : systematic review, Br 7163 (November) (2011) 1–31. doi:10.1136/bmj.d7163.

[4] T. Masud, R. O. Morris, Epidemiology of falls, Age and Ageing 30 (2001) 3–7. doi:10.1093/ageing/30.1.3.

[5] E. Tedeschini, Y. Levkovitz, N. Iovieno, V. E. Ameral, J. C. Nelson, G. I. Papakostas, Efficacy of antidepressants for late-life depression: a meta-analysis and meta-regression of placebo-controlled randomized trials., The Journal of clinical psychiatry 72 (12) (2011) 1660–8. doi:10.4088/JCP.10r06531.
URL http://article.psychiatrist.com/?ContentType=START{&}ID=10007683

[6] R. R. Huxley, M. Woodward, Cigarette smoking as a risk factor for coronary heart disease in women compared with men: a systematic review and meta-analysis of prospective cohort studies, The Lancet 378 (9799) (2011) 1297–1305.

[7] S. R. Lord, H. B. Menz, A. Tiedemann, A physiological profile approach to falls risk assessment and prevention., Physical therapy 83 (3) (2003) 237–52.
URL http://ptjournal.apta.org/content/83/3/237.full

[8] S. E. Chiuve, N. R. Cook, C. M. Shay, K. M. Rexrode, C. M. Albert, J. E. Manson, W. C. Willett, E. B. Rimm, Lifestyle-Based Prediction Model for the Prevention of CVD: The Healthy Heart Score, Journal of the American Heart Association 3 (6) (2014) e000954–e000954. doi:10.1161/JAHA.114.000954.
URL http://www.ncbi.nlm.nih.gov/pubmed/25398889

[9] S. Scheuer, D. Haase, V. Meyer, Towards a flood risk assessment ontology – Knowledge integration into a multi-criteria risk assessment approach, Computers, Environment and Urban Systems 37 (2013) 82–94. doi:10.1016/j.compenvurbsys.2012.07.007.

URL http://linkinghub.elsevier.com/retrieve/pii/S0198971512000762

[10] J. D. Ferreira, C. Pesquita, F. M. Couto, M. J. Silva, Bringing epidemiology into the Semantic Web, in: Proceedings of the International Conference on Biomedical Ontologies, PMC, Bethesda MD, USA, 2010, pp. 1–5.

[11] G. Kong, D.-L. Xu, R. Body, J.-B. Yang, K. Mackway-Jones, S. Carley, A belief rule-based decision support system for clinical risk assessment of cardiac chest pain, European Journal of Operational Research 219 (3) (2012) 564–573. doi:10.1016/j.ejor.2011.10.044.
URL https://linkinghub.elsevier.com/retrieve/pii/S0377221711009842

[12] K. E. Freedland, M. J. Hesseler, R. M. Carney, B. C. Steinmeyer, J. A. Skala, V. G. Dávila-Román, M. W. Rich, Major Depression and Long-Term Survival of Patients With Heart Failure, Psychosomatic Medicine 78 (8) (2016) 896–903. doi:10.1097/PSY.0000000000000346.
URL https://insights.ovid.com/crossref?an=00006842-201610000-00003

[13] S. Johnson, S. Kelly, D. Rasali, Differences in fall injury hospitalization and related survival rates among older adults across age, sex, and areas of residence in Canada, Injury Epidemiology 2 (1) (2015) 24. doi:10.1186/s40621-015-0056-1.
URL https://injepijournal.biomedcentral.com/articles/10.1186/s40621-015-0056-1

[14] B. A. Goldstein, A. M. Navar, R. E. Carter, Moving beyond regression techniques in cardiovascular risk prediction: applying machine learning to address analytic challenges, European Heart Journal 38 (23) (2016) ehw302. doi:10.1093/eurheartj/ehw302.
URL http://eurheartj.oxfordjournals.org/lookup/doi/10.1093/eurheartj/ehw302

[15] L. Cattelani, P. Palumbo, L. Palmerini, S. Bandinelli, C. Becker, F. Chesani, L. Chiari, FRAT-up, a Fall-Risk Assessment Tool for Elderly People Living in the Community, Journal of Medical Internet Re-

search 17 (2) (2015) e41. doi:10.2196/jmir.4064.
URL http://www.jmir.org/2015/2/e41/

[16] P. Palumbo, J. Klenk, L. Cattelani, S. Bandinelli, L. Ferrucci, K. Rapp, L. Chiari, D. Rothenbacher, Predictive Performance of a Fall Risk Assessment Tool (FRAT-up) for Community-Dwelling Older People in 4 European Cohorts, Journal of the American Medical Directors Association 17 (12) (2016) 1106–1113. doi:10.1016/j.jamda.2016.07.015.
URL http://www.sciencedirect.com/science/article/pii/S1525861016302936

[17] L. Cattelani, M. Belvederi Murri, F. Chesani, L. Chiari, S. Bandinelli, P. Palumbo, Risk Prediction Model for Late Life Depression: Development and Validation on Three Large European Datasets., IEEE journal of biomedical and health informatics 23 (5) (2018) 2196–2204. doi:10.1109/JBHI.2018.2884079.
URL https://ieeexplore.ieee.org/document/8552433/

[18] P. Palumbo, L. Palmerini, S. Bandinelli, L. Chiari, Fall Risk Assessment Tools for Elderly Living in the Community: Can We Do Better?, PLOS ONE 10 (12) (2015) e0146247. doi:10.1371/journal.pone.0146247.
URL http://journals.plos.org/plosone/article?id=10.1371/journal.pone.0146247

[19] A. Gomez-Perez, M. Fernández-López, O. Corcho, Ontological Engineering: with examples from the areas of Knowledge Management, e-Commerce and the Semantic Web, Springer Science & Business Media, London, 2006.

[20] J. R. G. Pulido, M. A. G. Ruiz, R. Herrera, E. Cabello, S. Legrand, D. Elliman, Ontology languages for the semantic web: A never completely updated review, Knowledge-Based Systems 19 (7) (2006) 489–497. doi:http://dx.doi.org/10.1016/j.knosys.2006.04.013.
URL http://www.sciencedirect.com/science/article/pii/S0950705106000736

[21] F. Baader, D. Calvanese, D. L. McGuinness, D. Nardi, P. F. Patel-Schneider, The Description Logic Handbook: Theory, Implementation, and Applications, Cambridge university press, 2003.

[22] J. D. Ferreira, D. Paolotti, F. M. Couto, M. J. Silva, On the usefulness of ontologies in epidemiology research and practice, J Epidemiol Community Health 67 (5) (2013) 385–388.

[23] J. W. Lloyd, Foundations of Logic Programming, 2nd Edition, Springer, 1987.

[24] T. Sato, A statistical learning method for logic programs with distribution semantics, in: L. Sterling (Ed.), Proceedings of the Twelth International conference on logic programming, MIT Press, Tokyo, 1995, pp. 715–729.
URL http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.17.4408

[25] J. Vennekens, S. Verbaeten, M. Bruynooghe, Logic programs with annotated disjunctions, in: B. Demoen, V. Lifschitz (Eds.), Logic Programming, Springer Berlin Heidelberg, Berlin, Heidelberg, 2004, pp. 431–445. doi:10.1007/978-3-540-27775-0_30.

[26] F. Riguzzi, T. Swift, Well–definedness and efficient inference for probabilistic logic programming under the distribution semantics, Theory and Practice of Logic Programming 13 (02) (2012) 279–302. arXiv:1110.0631, doi:10.1017/S1471068411000664.

[27] W3C OWL Working Group, OWL 2 Web Ontology Language Document Overview (2013).
URL http://www.w3.org/TR/owl2-overview/

[28] Association for Information Science and Technology, Dublin Core Metadata Initiative (2015).
URL http://dublincore.org/

[29] L. S. Radloff, The CES-D Scale: A Self-Report Depression Scale for Research in the General Population, Applied psychological measurement 1 (3) (1977) 385–401. doi:10.1177/014662167700100306.
URL http://apm.sagepub.com/content/1/3/385.short

[30] S. Deandrea, E. Lucenteforte, F. Bravi, R. Foschi, C. La Vecchia, E. Negri, Risk Factors for Falls in Community-dwelling Older People: A Systematic Review and Meta-analysis, Epidemiology 21 (5) (2010) 658–668.

[31] M. Horridge, P. F. Patel-Schneider, OWL 2 Web Ontology Language Manchester Syntax (Second Edition) (2012).
URL http://www.w3.org/TR/owl2-manchester-syntax

[32] S. C. Kleene, N. G. de Bruijn, J. de Groot, A. C. Zaanen, Introduction to metamathematics, van Nostrand New York, New York, 1952.

[33] L. Z. Rubenstein, K. R. Josephson, Falls and their prevention in elderly people: what does the evidence show?, The Medical clinics of North America 90 (5) (2006) 807–24. doi:10.1016/j.mcna.2006.05.013.
URL http://www.ncbi.nlm.nih.gov/pubmed/16962843

[34] K. L. Stone, S. Ancoli-Israel, T. Blackwell, K. E. Ensrud, J. a. Cauley, S. Redline, T. a. Hillier, J. Schneider, D. Claman, S. R. Cummings, Actigraphy-measured sleep characteristics and risk of falls in older women., Archives of internal medicine 168 (16) (2008) 1768–75. doi:10.1001/archinte.168.16.1768.
URL http://www.ncbi.nlm.nih.gov/pubmed/18779464

[35] A. Nawaz, J. L. Helbostad, L. Chiari, F. Chesani, L. Cattelani, User Experience (UX) of the Fall Risk Assessment Tool (FRAT-up), in: Computer-Based Medical Systems (CBMS), 2015 IEEE 28th International Symposium on, IEEE, Sao Carlos, 2015, pp. 19–22. doi:10.1109/CBMS.2015.63.

[36] T. A. Gerds, T. Cai, M. Schumacher, The performance of risk prediction models., Biometrical journal. Biometrische Zeitschrift 50 (4) (2008) 457–79. doi:10.1002/bimj.200810443.
URL http://www.ncbi.nlm.nih.gov/pubmed/18663757

[37] L. Cattelani, F. Chesani, P. Palumbo, L. Palmerini, S. Bandinelli, C. Becker, L. Chiari, FRAT-Up, a Rule-Based System Evaluating Fall Risk in the Elderly, 2014 IEEE 27th International Symposium on Computer-Based Medical Systems 1 (2014) 38–41. doi:10.1109/CBMS.2014.35.
URL http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=6881844

[38] S. R. Lord, C. Sherrington, H. B. Menz, Falls in older people. Risk factors and strategies for prevention., Vol. 35 Suppl 2, Cambridge Uni-

versity Press, New York, 2001. doi:10.1093/ageing/afl084.
URL http://www.ncbi.nlm.nih.gov/pubmed/16926202

[39] I. M. Miake-Lye, S. Hempel, D. A. Ganz, P. G. Shekelle, Inpatient fall prevention programs as a patient safety strategy: a systematic review., Annals of internal medicine 158 (5 Pt 2) (2013) 390–6. doi:10.7326/0003-4819-158-5-201303051-00005.
URL http://annals.org/article.aspx?articleid=1656443

[40] L. D. Gillespie, M. C. Robertson, W. J. Gillespie, C. Sherrington, S. Gates, L. M. Clemson, S. E. Lamb, Interventions for preventing falls in older people living in the community, Cochrane database of systematic reviews (9) (2012) CD007146. doi:10.1002/14651858.CD007146.pub3.
URL http://onlinelibrary.wiley.com/doi/10.1002/14651858.CD007146.pub2/pdf/standard

[41] I. D. Cameron, L. D. Gillespie, M. C. Robertson, G. R. Murray, K. D. Hill, R. G. Cumming, N. Kerse, Interventions for preventing falls in older people in care facilities and hospitals, Cochrane database of systematic reviews (Online) 12 (2012) CD005465. doi:10.1002/14651858.CD005465.pub3.
URL http://www.ncbi.nlm.nih.gov/pubmed/23235623

[42] World Health Organization Department of Ageing and Life Course, WHO global report on falls prevention in older age, World Health Organization, Geneva, 2008.

[43] M. D. Denkinger, S. Franke, K. Rapp, G. Weinmayr, E. Duran-Tauleria, T. Nikolaus, R. Peter, Accelerometer-based physical activity in a large observational cohort–study protocol and design of the activity and function of the elderly in Ulm (ActiFE Ulm) study., BMC geriatrics 10 (1) (2010) 50. doi:10.1186/1471-2318-10-50.
URL http://www.biomedcentral.com/1471-2318/10/50

[44] J. Klenk, N. Kerse, K. Rapp, T. Nikolaus, C. Becker, D. Rothenbacher, R. Peter, M. D. Denkinger, Physical Activity and Different Concepts of Fall Risk Estimation in Older People —Results of the ActiFE-Ulm Study., PloS one 10 (6) (2015) e0129098. doi:10.1371/journal.pone.0129098.
URL http://www.ncbi.nlm.nih.gov/pubmed/26058056

[45] A. Steptoe, E. Breeze, J. Banks, J. Nazroo, Cohort profile: the English Longitudinal Study of Ageing., International journal of epidemiology 42 (6) (2013) 1640–8. doi:10.1093/ije/dys168.
URL http://www.ncbi.nlm.nih.gov/pubmed/23143611

[46] L. Ferrucci, S. Bandinelli, E. Benvenuti, A. Iorio, C. Macchi, T. B. Harris, J. M. Guralnik, Subsystems contributing to the decline in ability to walk: bridging the gap between epidemiology and geriatric practice in the InCHIANTI study, Journal of the American Geriatrics Society 48 (12) (2000) 1618–1625.
URL http://www.ncbi.nlm.nih.gov/pubmed/11129752

[47] R. A. Kenny, B. J. Whelan, H. Cronin, Y. Kamiya, P. Kearney, C. O'Regan, M. Ziegel, The design of the Irish longitudinal study on ageing (2010).
URL http://hdl.handle.net/10147/301640

[48] A. Nolan, C. O'Regan, C. Dooley, D. Wallace, A. Hever, H. Cronin, E. Hudson, R. A. Kenny, The over 50s in a changing Ireland: Economic circumstances, health and well-being, The Irish Longitudinal Study on Ageing, Dublin, 2014.

[49] F. Bernardini, L. Attademo, S. D. Cleary, C. Luther, R. S. Shim, R. Quartesan, M. T. Compton, Risk Prediction Models in Psychiatry: Toward a new frontier for the prevention of mental illnesses, J. Clin. Psychiatry 78 (5) (2017) 572–583. doi:10.4088/JCP.15r10003.
URL https://doi.org/10.4088/JCP.15r10003

[50] M. Luppa, C. Sikorski, T. Motzek, A. Konnopka, H.-H. König, S. G. Riedel-Heller, Health Service Utilization and Costs of Depressive Symptoms in Late Life - A Systematic Review, Curr. Pharm. Des. 18 (36) (2012) 5936–5957. doi:10.2174/138161212803523572.
URL http://www.ncbi.nlm.nih.gov/pubmed/22681171

[51] R. M. Kok, C. F. Reynolds, Management of Depression in Older Adults: A Review., JAMA 317 (20) (2017) 2114–2122. doi:10.1001/jama.2017.5706.
URL http://jama.jamanetwork.com/article.aspx?doi=10.1001/jama.2017.5706

[52] F. Holvast, B. Massoudi, R. C. Oude Voshaar, P. F. M. Verhaak, Non-pharmacological treatment for depressed older patients in primary care: A systematic review and meta-analysis, Plos One 12 (9) (2017) e0184666. doi:10.1371/journal.pone.0184666.
URL http://dx.plos.org/10.1371/journal.pone.0184666

[53] M. Belvederi Murri, M. Amore, M. Respino, G. S. Alexopoulos, The symptom network structure of depressive symptoms in late-life: Results from a European population study, Molecular Psychiatrydoi:10.1038/s41380-018-0232-0.
URL http://www.nature.com/articles/s41380-018-0232-0

[54] M. G. Cole, N. Dendukuri, Risk Factors for Depression Among Elderly Community Subjects: A Systematic Review and Meta-Analysis, Am. J. Psychiatry 160 (6) (2003) 1147–1156. doi:10.1176/appi.ajp.160.6.1147.

[55] J. R. M. Copeland, A. T. F. Beekman, A. W. Braam, M. E. Dewey, P. Delespaul, R. Fuhrer, C. Hooijer, B. A. Lawlor, S.-L. Kivela, A. Lobo, H. Magnusson, A. H. Mann, I. Meller, M. J. Prince, F. Reischies, M. Roelands, I. Skoog, C. Turrina, M. W. DeVries, K. C. M. Wilson, Depression among older people in Europe: the EURODEP studies., World psychiatry : official journal of the World Psychiatric Association (WPA) 3 (1) (2004) 45–9.
URL https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1414664/

[56] K. Okamoto, Y. Harasawa, Prediction of symptomatic depression by discriminant analysis in Japanese community-dwelling elderly, Archives of Gerontology and Geriatrics 52 (2) (2011) 177–180. doi:10.1016/j.archger.2010.03.012.

[57] S. Chen, B. Chen, B. D. Fath, Ecological risk assessment on the system scale: A review of state-of-the-art models and future perspectives, Ecological Modelling 250 (19) (2013) 25–33. doi:10.1016/j.ecolmodel.2012.10.015.
URL http://dx.doi.org/10.1016/j.ecolmodel.2012.10.015

[58] E. M. Haney, M. E. O'Neil, S. Carson, A. Low, K. Peterson, L. M. Denneson, C. Oleksiewicz, D. Kansagara, Suicide Risk Factors and Risk Assessment Tools: A Systematic Review, VA Evidence-based Synthesis

Program Reports, Department of Veterans Affairs, Washington (DC), Washington (DC), 2012.
URL http://europepmc.org/books/NBK92671

[59] D. C. Goff, D. M. Lloyd-Jones, G. Bennett, S. Coady, R. B. D'Agostino, R. Gibbons, P. Greenland, D. T. Lackland, D. Levy, C. J. O'Donnell, et al., 2013 ACC/AHA Guideline on the Assessment of Cardiovascular Risk, Journal of the American College of Cardiology 63 (25) (2014) 2935–2959. doi:10.1016/j.jacc.2013.11.005.
URL http://linkinghub.elsevier.com/retrieve/pii/S0735109713060312

[60] S. S. Lim, T. Vos, A. D. Flaxman, G. Danaei, K. Shibuya, H. Adair-Rohani, M. A. AlMazroa, M. Amann, H. R. Anderson, K. G. Andrews, et al., A comparative risk assessment of burden of disease and injury attributable to 67 risk factors and risk factor clusters in 21 regions, 1990–2010: a systematic analysis for the Global Burden of Disease Study 2010, The Lancet 380 (9859) (2012) 2224–2260. doi:10.1016/S0140-6736(12)61766-8.
URL http://linkinghub.elsevier.com/retrieve/pii/S0140673612617668

[61] S. Muggleton, Learning stochastic logic programs, Electron. Trans. Artif. Intell. 4 (B) (2000) 141–153.
URL http://www.ep.liu.se/ej/etai/2000/015/

[62] E. Bellodi, F. Riguzzi, Expectation maximization over binary decision diagrams for probabilistic logic programs, Intell. Data Anal. 17 (2) (2013) 343–363. doi:10.3233/IDA-130582.
URL https://doi.org/10.3233/IDA-130582

[63] N. D. Mauro, E. Bellodi, F. Riguzzi, Bandit-based monte-carlo structure learning of probabilistic logic programs, Machine Learning 100 (1) (2015) 127–156. doi:10.1007/s10994-015-5510-3.
URL https://doi.org/10.1007/s10994-015-5510-3

[64] G. Parisi, Balance research funds across Europe, Nature 530 (2016) 33.
URL http://dx.doi.org/10.1038/530033d