Working Paper Series, N. 6, May 2020

Department of Statistical Sciences
University of Padua
Italy

Construction of a frailty indicator with partially ordered sets: a multiple outcome proposal based on administrative healthcare data

Margherita Silan
Department of Statistical Sciences
University of Padua
Italy

Giada Signorin
Department of Statistical Sciences
University of Padua
Italy

Elisa Ferracin
Epidemiology Unit
Regional Health Service ASL TO3
Italy

Elisabetta Listorti
Centre for research on health and social care management (CERGAS)
Bocconi University
Italy

Teresa Spadea
Epidemiology Unit
Regional Health Service ASL TO3
Italy

Giuseppe Costa
Epidemiology Unit
Regional Health Service ASL TO3
Italy

Giovanna Boccuzzo
Department of Statistical Sciences
University of Padua
Italy

Keywords: Frailty Indicator · administrative healthcare data · poset theory · measurement · aging · multiple outcomes

Final version (2019-10-31)

Contents

Department of Statistical Sciences
Via Cesare Battisti, 241
35121 Padova
Italy

tel: +39 049 8274168
fax: +39 049 8274170
http://www.stat.unipd.it

Corresponding author:
Margherita Silan
silan@stat.unipd.it

Construction of a frailty indicator with partially ordered sets: a multiple outcome proposal based on administrative healthcare data

Margherita Silan
Department of Statistical Sciences
University of Padua
Italy

Giada Signorin
Department of Statistical Sciences
University of Padua
Italy

Elisa Ferracin
Epidemiology Unit
Regional Health Service ASL TO3
Italy

Elisabetta Listorti
Centre for research on health and social care management (CERGAS)
Bocconi University
Italy

Teresa Spadea
Epidemiology Unit
Regional Health Service ASL TO3
Italy

Giuseppe Costa
Epidemiology Unit
Regional Health Service ASL TO3
Italy

Giovanna Boccuzzo
Department of Statistical Sciences
University of Padua
Italy

Abstract: Given the progressive ageing of Italian and European populations, chronic diseases attributable to ageing are rising steeply, calling for new strategies for health resources management and implementation of prevention policies. Among chronic patients, frail subjects have special and wider care needs, together with an increased risk of adverse health outcomes. Thus, their identification is a fundamental goal, claimed as the first step of the Italian National Program for Chronic Diseases.

The aim of this work is to construct an indicator that measures the frailty level of each individual in the population using administrative health data-flows of the Piedmont region. Following the multidimensional nature of frailty, in our proposal we adopted a multiple outcome approach, by considering six outcomes: death, emergency hospitalization, access to the emergency room with red code, avoidable hospitalization, hip fracture and disability. We identified a parsimonious set of 7 explanatory variables, able to predict simultaneously the six outcomes we considered, and we assembled them in a unique frailty indicator through the use of partially ordered set (poset) theory.

Our indicator has a good performance with respect to all the outcomes and is able to describe several individual characteristics that are not directly considered in the computation of the indicator. Thanks to its parsimony and to the use of administrative health data, our indicator allows all the stakeholders involved in the healthcare process, such as Italian Local Health Units, general practitioners and regional managers, to use it in order to target frailer individuals with better comprehensive healthcare actions.

Keywords: Frailty Indicator · administrative healthcare data · poset theory · measurement · aging · multiple outcomes

1 Introduction

Nowadays, the identification of frail individuals is quite a popular research theme. Despite the growing interest about this topic, a unique definition for frail individuals still does not exist and frailty is defined as a syndrome in desperate need of description and analysis (Gillick, 2001). Indeed, the identification of common criteria and guidelines for frailty has been described as a highly complex and demanding task (Bortz, 2002); this is because frailty is a complex and multidimensional concept that involves several functional domains of elderly (Gobbens et al., 2010).

The attention towards the definition, measuring and identification of frailty originates from the ageing phenomenon. In fact, the average age of the population has increased worldwide, together with the diffusion of age-related chronic conditions. The amount of people older than 65 years in Europe amounts to 101 millions of people (out of total 512 millions). Between 2018 and 2050, this number is projected to expand, by 17.6 % for those aged 65-74 years, and by 60.5% for those aged 75-84 years. As for the Italian context, it is the European country with the highest percentage of people aged 55 years or more, accounting for more than one third of the population. Among them, the portion referring a longstanding illness or health problem ranged from to 72.5% (people aged 85 years or more) to 56.9% (people aged 65-74 years) (Eurostat, 2019).

In this setting, in 2016 the Italian Ministry of Health issued a National Plan for Chronic Disease (NPCD) (Ministero della Salute, 2016), a strategic framework aiming at improving the organization of healthcare services that are related to old and chronic subjects, centering clinical decisions on patients, and calling into action all the involved stakeholders.

In order to reach these goals, the first step that has been established is the stratification of the population. The task consists in the definition of algorithms and tools that exploit administrative health data-flows to identify subgroups of the population that are homogeneous in terms of health needs and characteristics. The value added of the initial stratification is to undertake policy measures that can be tailored, based on the subgroup of the population to which individuals belong. Moreover, being based on administrative health data, the stratification represents a tool that can be exploited by all the actors of the healthcare sector.

Among the older and chronic population, frail subjects are individuals with special care needs that are particularly important to identify in order to implement prevention policies and improve their quality of life (Bergman et al., 2002). This is why the construction of a tool that is able to measure frailty among individuals using administrative health data-flows assumes a fundamental role both in the national and international scenarios.

This topic was already been addressed in the article by Silan et al. (2019), where a composite indicator for frailty is proposed using administrative health data-flows coming from an Italian Local Health Units (LHUs) in the northern part of the Italian city of Padova (LHU 15 "Alta Padovana", now part of LHU 6 "Euganea"). The composite indicator proposed by Silan et al. (2019) is based on the fact that frail subjects are characterized by an increased susceptibility to experiment negative outcomes related with frailty condition (Fried et al., 2001). Indeed, given the absence of a unique and shared definition for frailty, the ability to predict adverse outcomes is considered the highest standard for a successful definition of frail individual (Rockwood, 2005). The two most cited negative outcomes are death and emergency hospitalization (Falasca et al., 2011). These two outcomes are also used for the definition of composite indicator by Silan et al. (2019), that is defined using 9 variables collected from administrative health data-flows and aggregated using partially ordered set theory.

However, from a policy maker point of view, the identification of frail subjects as individuals with higher risk to die or be emergency hospitalized offers little room for improvement in patients' health conditions, while the NPCD advocates concrete actions directed to a global improvement of the quality of life. Indeed, taking into account also less severe outcomes in the definition of frail elderly may result in a more useful tool in order to stratify the population and intervene with prevention policies to keep elderly health conditions from worsening. This is the reason why we decided to enlarge the list of outcomes considered in the construction of the frailty indicator, by including four more outcomes: access to the emergency room with red code, avoidable hospitalization, hip fracture and disability. Through the inclusion of multiple health outcomes, the indicator achieves the additional goal of representing the complexity of frailty condition. However, the inclusion of more outcomes increases also the complexity of the statistical methodology and some computational aspects involved in the variables selection.

In this paper it is described the methodological proposal for a frailty indicator that faces complexity in the theoretical definition of frailty, its representation using administrative data-flows of the Italian Piedmont region, the management of several outcomes and methodological and computational issues. The main goal of this work consists in disentangling all these issues, in order to propose a frailty indicator that involves a small set of variables that are easy to collect from administrative health data-flows, and therefore easy to implement and to replicate in any Italian Local

Health Unit. In particular, through the poset methodology, we answer the need for a composite indicator that is able to detect the frailest individuals compared with the population to which they belong.

2 Steps for the Construction of the Frailty Indicator

This section summarizes the main steps that have been undertaken for the construction of the frailty indicator 1.

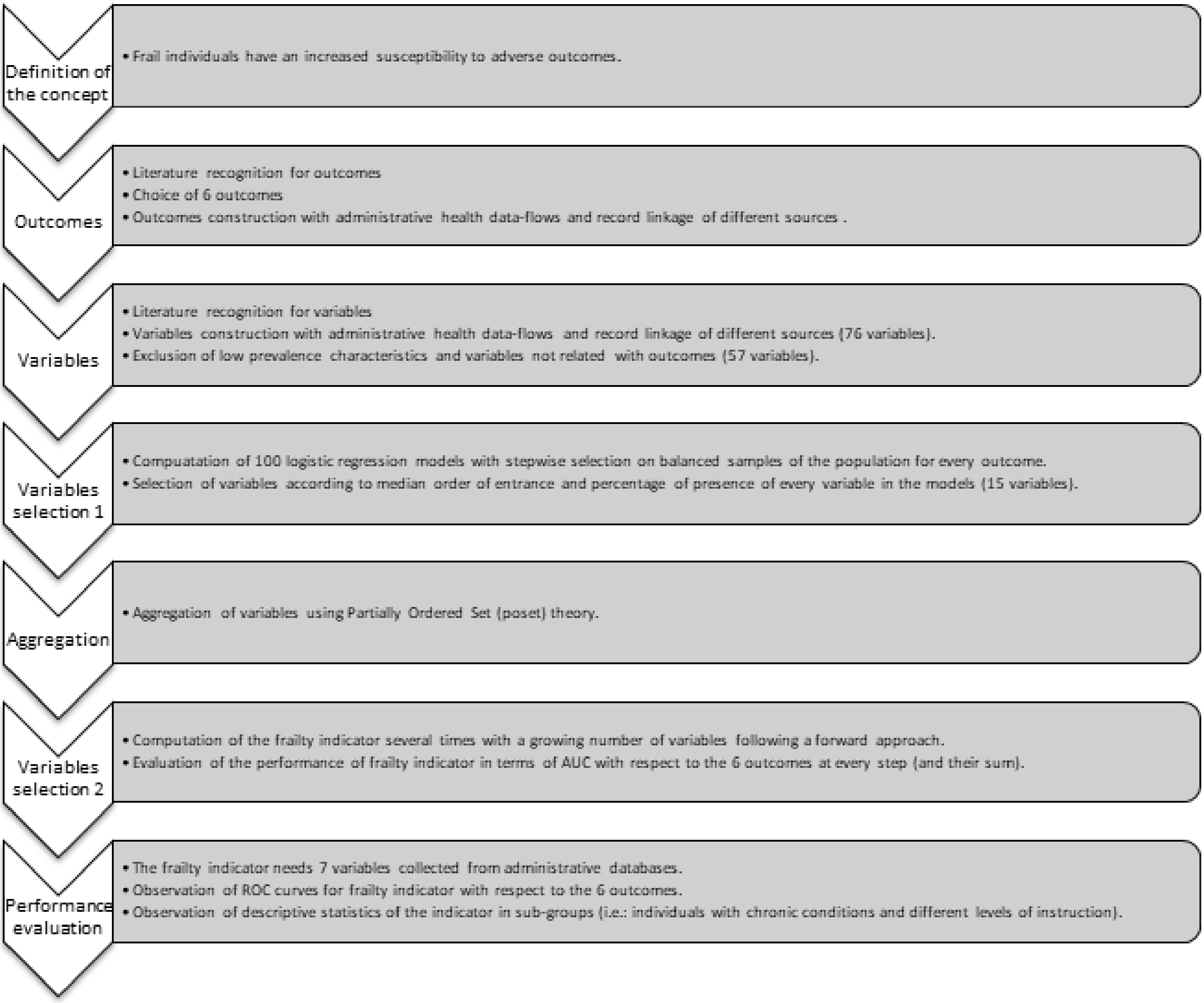


Figure 1: Steps for the construction of the frailty indicator.

The first step consists in the literature recognition looking for frailty definitions. As explained, there is no consensus on a unique definition for frailty, thus the composite indicator is based on the fact that frail individuals have an increased susceptibility to adverse negative outcomes related with frailty condition (more details are provided in section 3.1).

As a consequence, the second step for the construction of the frailty indicator consists in the choice of outcomes usually associated with frailty condition according to literature. After careful consideration, six outcomes were selected as related with frailty condition: death, emergency hospitalization, access to the emergency room (ER) with red code, avoidable hospitalization, hip fracture and disability (as explained in depth in section 3.2). These six outcomes are collected from administrative health data-flows, thus their practical definition goes together with the structure of administrative databases and in some cases it is quite delicate. For instance, the disability outcome is coded as the incidence (not prevalence) of disability in 2016. A brief description of data sources used in this work can be found in section 4.

Outcomes are not directly included in the computation of the frailty indicator, but they are fundamental in order to select the explanatory variables that constitute the composite indicator. Thus, the third step is strictly connected with the selected outcomes. It consists in a bibliographic research of predictive variables for the six outcomes (section 3.3). The crucial point of this step consists in the codification of the variables and their construction with administrative health databases. Also in this case, the practical definition of variables is quite delicate and strictly connected with some practical aspects related with the structure of administrative databases. For instance, chronic conditions are dichotomous variables equal to one if the subject suffers from the disease. Variables referred to chronic conditions are the result of the union of information coming from different data sources such as the participation in the prescription charges, the territorial drug prescriptions and diagnosis assigned in hospital discharge records and in accident & emergency databases.

The final dataset contains 76 explanatory variables collected and computed thanks to a deterministic record linkage (performed with a personal identification code and described in section 4.3) of several different data-sources (with respect to 2014 and 2015) and the six negative outcomes observed on administrative databases of 2016.

The step number four consists in a variables selection based on their ability to predict all the six considered outcomes. This task is carried on by the estimation of 100 logistic regression models on balanced samples of the whole population for every outcome with a stepwise selection criterion (described in detail in section 5.1). The stepwise criterion allows the observation of both the presence of every variable in estimated models, but also order of entrance of each variable in the models. Thus, the selection of variables is guided by both median order of entrance and percentage of presence of every variables in the models (the reduced set of variables contains 15 variables).

Then, in the fifth step, the variables should be aggregated using Partially Ordered Set theory (poset), fully described in paragraph 5.2. However, the number of variables selected is still too high both because of computational reasons and because of the inclusion of all of them would worsen the performance of the indicator. Thus, the following step is again about variables selection, but this time it follows a forward approach in order to maximize the sum of the Area Under the ROC Curves (AUC) of the six outcomes, the steps of this procedure are described in section 5.3.

The final composite indicator for frailty is composed by 7 variables aggregated through the use of poset theory. The final step for its construction consists on performance evaluation and validation, described in section 6. This step is mainly based on checking for some characteristics that according to literature (Rockwood, 2005) make a definition of frailty successful.

In the following, the steps for the construction of the frailty indicator are described in detail.

3 Frailty: a literature recognition

3.1 Definition of frailty

Within the epidemiological literature, the concept of frailty emerges with a multidimensional nature (Gobbens et al., 2010). As such, it does not refer to a unique context, rather it includes a combination of physical, sensory, psychological and social domains. The presence of multiple facets that need to be considered together generates the challenge of determining a unique conceptual definition of frailty, which at the moment is still missing.

In terms of usage and citations, the most known definition of frailty refers to Fried et al. (2001), which has considered it in its clinical aspect: frailty is defined as a biologic syndrome of decreased reserve and resistance to stressors, resulting from cumulative declines across multiple physiologic systems, causing vulnerability to adverse outcomes. In fact, the measurement system proposed by Fried follows this clinical approach: the diagnosis is medical and based on presence of symptoms of physical weakness (including weak muscles strength, slow gait speed, unintentional weight loss, low physical activity). Another well-known definition is the one made by Rockwood (2005), who has highlighted a more comprehensive view: frailty is considered as a combination of aging, disease and other factors that make some people vulnerable, whose consequences affect the functional status as well as physical and cognitive deficits and psychosocial factors. Other definitions have been added (Buchner and Wagner, 1992; Campbell and Buchner, 1997; Strawbridge et al., 1998). Based on the experiment performed by Gobbens et al. (2010), the conceptual definition most accepted by experts was the one produced by Schuurmans et al. (2004): frailty as a loss of resources in several domains of functioning, which leads to a declining reserve capacity for dealing with stressors. Eventually, Gobbens himself has produced a more complete definition of frailty, which entails, beyond the physical and psychological aspect, also the social one (Gobbens et al., 2010).

Depending on the definition adopted, frailty prevalence can range among 7% (Fried et al., 2001) and 45% (Roppolo et al., 2015) of the adults aged 65 years and older, with the occurrence of frailty increasing with age. Independently from the definition used, frailty is associated with increased need for assistance with mobility, self-care, and activities of daily living, with an associated progressive loss of self-confidence, leading to social isolation, reduced physical activity, progressive isolation, and decreased social interaction, further exacerbating the level of frailty. Moreover, frailty is recognized as a continuum process, a relative state that can change over time (Gobbens et al., 2010). As such, it is possible to intervene in that pathway so that frailty is delayed, reduced, or prevented from becoming worse (Morley et al., 2002).

All in all, two main points recur in researchers' works, which are crucial for the decision on our operational definition of frailty, as it will be explained in the next paragraph. First of all, all definitions of frailty refer to a state of susceptibility to adverse health outcomes. Secondly, experts' opinions converge on treating frailty as an accelerator in this process of worsening of health conditions. In fact, frailty is said to confer significantly increased risk for poor health outcomes, such as hospitalization and mortality.

The two elements together compose the general meaning of frailty as a personal condition of having unfavorable conditions in the future.

Despite these two common characteristics of paramount importance, frailty indeed encloses all the previously mentioned facets, and thus requires to be tackled as an umbrella term. Next section will explain how the health outcomes considered in this analysis aim to include the multidimensional nature of frailty.

3.2 Adverse health outcomes related with frailty condition

As mentioned in the previous section, frailty does not own a unique conceptual definition. Likewise, frailty has not yet run into a unique measurement system, and our contribution is in the research of a better method to identify it and to reach consensus in the clinical community. Our approach to determine our operational definition is to focus on the envisaged goal: we reckon that the objective of acting upon frailty in a specific context conducts to the proper operational definition of frailty for that context. In our case, the need for the definition of frailty is driven by the implementation at the regional level of the NPCD, which calls for action in anticipating and slowing down the clinical worsening of elderly people. Hence, the main point we are interested in is frailty seen as a global deterioration of health conditions, which in turn causes adverse clinical events. The higher the level of frailty, the higher the probability that people may incur in one or more of these events. Through the mapping of these clinical events, frailer subjects can be found, and to them a preventive and supportive policy can be addressed, in order to reduce the probability of incurring in additional negative health events.

Table 1 summarizes the main outcomes that have been used in recent works on frailty. In particular, since our aim is to use only administrative health data, we reported only those papers that use prediction models based on this type of data.

The most widely used outcome is death, which ultimately resumes all the processes that can worsen the health status. However, the inspiring principle of the NPCD is to slow down the process leading to such a hard outcome, before it becomes unavoidable. Hence, we decided to add to our analysis other types of outcomes, in order to highlight aspects of frailty that are preventable and that the NPCD can deal with. In particular, three types of events are of interest for this approach: 1) emergency admissions and urgent hospitalizations, since they often refer to the exacerbation of a chronic disease, hence to failures in continuity of care; 2) avoidable hospitalizations, which represent the volume of hospital care potentially preventable by accessing to timely and effective primary and outpatient care; 3) events related to

Table 1: Literature sources for negative outcomes related with frailty condition.

Outcome	**Literature Sources**
Death	Paw et al. (1999), Fried et al. (2001), Saliba et al. (2001), Mitnitski et al. (2002), Klein et al. (2005), Mazzaglia et al. (2007), Ravaglia et al. (2008), Rothman et al. (2008), Ávila-Funes et al. (2009), Kamaruzzaman et al. (2010), Falasca et al. (2011), Cavazza and Malvi (2014)
Disability	Paw et al. (1999), Woods et al. (2005), Carrière et al. (2005), Ravaglia et al. (2008), Rothman et al. (2008), Ávila-Funes et al. (2009)
Institutionalization	Brody et al. (1997), Brody et al. (2002), Jones et al. (2004), Rockwood et al. (2006), Rothman et al. (2008), Luppa et al. (2009), Kamaruzzaman et al. (2010)
Hip Fracture	Woods et al. (2005), Ravaglia et al. (2008)
Hospitalization	Shelton et al. (2000), Fried et al. (2001), Landi et al. (2004), Damush et al. (2004), Walker et al. (2005), Mazzaglia et al. (2007), Ravaglia et al. (2008), Ávila-Funes et al. (2009), Crane et al. (2010), Kamaruzzaman et al. (2010)
Emergency Hospitalization	Shelton et al. (2000), Damush et al. (2004), Walker et al. (2005), Crane et al. (2010), Inouye et al. (2008), Falasca et al. (2011), López-Aguilà et al. (2011), Billings et al. (2013), Hippisley-Cox and Coupland (2013), Cavazza and Malvi (2014), Ahn et al. (2018)
Dementia	Buchman et al. (2007), Ávila-Funes et al. (2009), Song et al. (2014), Avila-Funes et al. (2012), Solfrizzi et al. (2013), Gray et al. (2013)
Comorbidity	Tammemagi et al. (2004)
Total of hospitalized days	Rockwood (2005), Makary et al. (2010)
Avoidable Hospitalization	Louis et al. (2014), Gao et al. (2014), Hibbard et al. (2017)

disability, unfavourable housing conditions or poor family support, usually triggered by failure in health-social integration. All these events often hide failures either in organization or in coverage of essential care, which can be subjected to improvement policies, such as those that the new NPCD wants to takes charge of. We decided to exclude outcomes related to the mental status of individuals, since they denote a more complex status that includes a combination of more biological and social aspects, all of which are different to predict or control.

Thus, we decided to include these six adverse events: death, urgent hospitalizations (Em. Hosp.), admissions in emergency department with red code (ER Red code), hip fracture interventions (Hip Fract.), avoidable hospitalizations (Avoid. Hosp.) and disability. In this way, i) the multidimensional nature of frailty is kept into account, ii) we reinforce the idea of frailty as a continuum process, whose progress can be positively influenced by actions directed to population and/or individuals.

3.3 Explanatory variables

Within each of the dimensions that have been described as involved in the frailty process, there are a number of risk factors. Among them, several and complex interactions exist, such that the total level of frailty is therefore not equivalent to the sum of its components (De Vries et al., 2013).

Four main categories can be outlined as containing all the risk factors that have been considered in the literature: personal, social, clinical and psychological or behavioral. The majority of the risk factors that have been related to frailty refer to evaluations that can be performed only through face-to-face visits. Since in our algorithm we use administrative health data, we face the limits imposed by them, such as the absence of information about the physical status of people, their lifestyles, etc. Table 2 reports the risk factors that have been explored in the literature and that can be measured with administrative health data.

Table 2: Variables considered for the frailty indicator, divided by category.

Category	**Variables considered**
Personal	Age, sex, nationality, education level
Social	Civil status, income, deprivation level, condition of property of house, utilization of mental health services, utilization of social services
Clinical	Charlson index, previous hospital/emergency department admissions, multiple prescriptions of drugs, past acute events (e.g., heart attack, hip fracture, etc.), presence of a chronic disease (coronary art disease, cancer, peripheral vascular disease, diabetes, etc.), disability
Psychological or behavioural	Depression, dementia

4 Data

4.1 Definition of the population

Our analysis concerns people aged 65 years and older, considering age computed at December 31, 2016. This specific age group was chosen mainly because people's health conditions and morbidities differ according to their age. Our observation period covered 3 years, where 2014 and 2015 were used to observe explanatory variables to be included in the indicator, while in 2016 negative outcomes related with frailty condition were observed. We used Piedmont population assisted register to select people aged 65 years and older at the end of December 2016. The total population counts 1,095,613 individuals.

4.2 Administrative Healthcare Databases

This research has been performed using the longitudinal study of the Piedmont (SLP) resident population (more than 4 million). The SLP is a record-linkage among population census (2011 data) and the main administrative healthcare databases. The variables had have been used to create the frailty index were observed in 2015 and in same cases in 2014 too. The outcomes were observed in 2016. More specifically, in order to describe individual health conditions, we used several databases:

1. archives of hospital discharges: these contain all the information collected on hospitalizations, including the principal diagnosis and up to five secondary diagnoses (2014-2015);

2. archives of drug prescriptions: these list all prescribed drugs and their quantities and classifies each drug according to its therapeutic, pharmacological and chemical properties, based on the Anatomical Therapeutic Chemical Classification (ATC) system (2014-2015);

3. first Aid: these records contain all the information collected on emergency room, including the principal diagnosis, triage color code access, color code exit (2014-2015);

4. income exemptions (2015): this dataset collects information on service activation, beneficiaries, starting and ending dates, exemptions code;

5. home care service (2015): it collects information on service activation, beneficiaries, starting and ending dates, treatments and diagnoses;

6. medical exemptions (2015): this dataset collects information on service activation, beneficiaries, starting and ending dates, medical exemptions code.

To describe social characteristics, we used population census in 2011, where are collected individual information about education, occupation, civil status, family type, house condition. To define death outcome we used the mortality register, who contains information on date of death and causes of death.

4.3 Record linkage

The abovementioned datasets are designed for administrative purposes; thus, records represent events, rather than patients. The first step needed to use these data for our purposes was to rearrange the rows of data so that they represent individuals. For example, the Territorial Drug Prescriptions dataset can contain several rows, one for each prescription, for the same patient, and, ideally, we needed a single row containing all the drug prescriptions for a given individual.

The following step consists on linking all the datasets to obtain all the available data for each individual. A deterministic record linkage of all the training datasets with the Piedmont population assisted register was run. The linkage was based on individual codes and it generated a single table that contains all the relevant events (from 2014 or 2015), characteristics of people living in the sample area, and their health outcomes in 2016.

5 Methods

5.1 Variables selection

In order to be as much confident as possible that all aspects related with frailty condition were taken into account in our analysis, all variables found in literature that we can collect from administrative health data-flows were considered in the first stage of the variables selection. Thus, we started with 76 variables. This set of variables includes age (divided in 6 five-years age-classes), some categorical and ordinal variables describing mostly socio-economic characteristics, some dummy variables that indicate the presence of a condition (for instance a chronic disease) and some counting variables that quantify the use of some services (hospitalizations, access to the emergency room according to the color code, prescriptions).

However, since we needed to select a small set of the most meaningful variables that better describe frailty condition in the whole population, we first discarded low prevalence characteristics (less than 1% cases in the whole population) and variables with no significant association with outcomes. Indeed, low prevalence characteristics, even if they might represent critical aspects for some individuals, do not help in the description of the health of the majority of the population and may be also somehow be otherwise represented by other more general variables. In this way the number of considered variables has been reduced to 57.

Some of the variables that we are taking into account are counting variables, that should be reduced in classes to be included in the construction of a composite indicator. In order to end up with ordinal variables with a restricted number of meaningful classes, we conducted a decision tree analysis that helped us to identify the groups of values that better classify individuals according to the risk of experimenting the considered outcomes. Having a large number of outcomes, the consensus about the best classification according to all of them is not easy to reach, thus in doubtful situation the most severe outcome, death, was used to make the ultimate choice.

The final goal of this step was to identify a parsimonious set of variables that is also able to predict simultaneously the six outcomes we considered. However, the six outcomes may have quite different risks. Thus, the first step was observing and selecting explanatory variables for each outcome separately. In order to do so, we used logistic regressions where the outcome was the dependent variable and all the 57 variables were considered for the regression and then selected by a stepwise selection criterion. Having a large population, a great number of variables had enough significance to be included in the final model. For this reason and in order to overcome over-fitting, we decided to implement a three-step procedure to select important explanatory variables for every outcome:

1. Selection of a random sample without replacement of 75% of the whole population, in order to avoid over-fitting.

2. Balance the sample in order to have the same amount of subjects that experimented the outcome, (cases), and of those subjects that do not experimented it (controls). The balanced sample contains all the cases and a random sample of controls of size equal to the number of the cases.

3. Logistic regression with stepwise selection criterion, this step allowed both to observe if the variable was included in the final model and its rank of inclusion, as a proxy of its importance in the prediction of the outcome.

These three steps were repeated 100 times for each outcome in order to get general results, not strictly linked with the analyzed population.

As a result of this three-steps procedure iterated 100 times for each outcome, we get two measures, repeated for each outcome, for every explanatory variable: percentage of presence and the median order of entrance (Table 1). The percentage of presence is the number of times that the variable was selected in the final version of the 100 models for a given outcome. The median order of entrances is the median rank of entrance in the models, considering all the 100 logistic regression models performed for every outcome; in order to get this measure, we assigned the maximum possible rank (that theoretically is 57, as the total number of considered variables) when the variable was not selected in the final model, however these extreme values cannot affect the median. When the median is equal to the maximum possible rank, an asterisk is reported in Table 3.

Having two measures that summarize the importance for every variable in the prediction of every outcome, the following step consists in the selection of a unique relevant set of variables with respect to all the six outcomes. Thus, we selected a small set of variables relevant in the prediction of all the six outcomes according to the two defined measures: variables with percentage of presence $\geq 75\%$ for at least 3 outcomes and median order of entrance ≥ 15 for at least 3 outcomes. In this way, 15 relevant variables were selected (enlisted in Table 3).

5.2 Partially ordered set (POSET) theory to aggregate variables

Once variables are selected, the need is to aggregate dichotomous and ordinal variables, in order to create a composite indicator for frailty. To this aim, partially ordered set (poset) theory is a powerful tool. Poset is, in mathematics, a set of

Table 3: Percentage of presence and median order of entrance of the selected variables for the six outcomes.

Variables	**Percentage of Presence (%)**						**Median Order of Entrance**					
	Death	Em. Hosp.	ER Red code	Avoid. Hosp.	Hip Fract.	Disability	Death	Em. Hosp.	ER Red code	Avoid. Hosp.	Hip Fract.	Disability
Blood system diseases	100	100	100	100	14	100	9	7	7	8	*	10
Circulatory diseases	100	20	83	100	3	100	25	*	10	12	*	13
Age	100	100	100	100	100	100	1	2	1	2	1	1
N. ER access with Green code	100	100	42	100	100	100	8	4	*	4	8	7
N. ER access with Yellow code	100	100	100	100	6	100	16.5	13	5	11	*	17
Diabetes	100	100	100	100	32	100	17	11	10	9	*	16
Disability	100	100	98	98	38	-[a]	3	6	4	6	*	*
Charlson Index[b]	100	100	100	100	44	100	2	5	2	3	*	5
Neoplasia	97	87	10	17	34	85	4	8	*	*	*	8
Nervous system diseases	100	13	9	2	98	100	15	*	*	*	3	3
Parkinson	100	95	98	100	99	100	7	24	12	15	9	6
Poliprescriptions	100	100	92	100	5	100	6	1	3	1	*	4
Respiratory diseases	100	100	100	100	0	97	14	11	7	7	*	25
Marital status	100	91	54	100	54	99	11	25	42	15	14	19
Accommodation occupancy title	100	100	100	100	10	100	28	17	11	15	*	13

[a]The Disability variable is not included in models with Disability as dependent variable, indeed subjects already disabled in 2016 are not even considered in the model estimation.

[b]The Charlson Index is the most widely used comorbidity index (Huang et al., 2014). It covers 19 issues (De Groot et al., 2003) such as diabetes, congestive heart failure, peripheral vascular disease, chronic pulmonary disease, and mild or severe liver disease, each weighted according to their potential influence on mortality. It is the sum of the weights assigned to each condition affecting a given patient (Charlson et al., 1987). For the purposes of the present study, the Index was calculated according to the Deyo et al. (1993) adaptation so that it could be used with the recorded ICD-9-CM diagnoses.

elements where a binary relation that indicates an order can be traced, the word "partial" refers to the fact that not every pair of elements could be comparable.

In order to easily explain poset theory and its use to compute the composite indicator for frailty, we propose a toy example, that involves six subjects and three dichotomous characteristics (Figure 2).

Let us suppose that we have a population comprising six individuals characterized by three dichotomous variables, as represented in Figure 2 (a): age (which takes a value of 0 for individuals who are between 60 and 70 years old, and 1 if they are older); drugs (which takes a value of 0 if individuals had no drug prescriptions in 2015, and 1 otherwise); and First Aid (which takes a value of 1 if individuals had access to ER at least once in 2015, and 0 otherwise). These variables are ordered, a value equal to 1 corresponds to worse health conditions, and we are assuming that they are able to summarize individuals' frailty condition. Thus, every subject is represented in the poset by its profile, given by the set of its characteristics.

The comparison of the individuals in the population gives rise to a list of comparabilities and incomparabilities (Davey and Priestley, 2002), which can be represented in a graphic form called Hasse diagram (Figure 2 (b)). This diagram represents the elements in a poset: each node is an element, two or more equal elements still form one node, and every line segment is an order relation between comparable objects. When two individuals are comparable, they are connected by line segments in the diagram, like A and B , B and E or C and E, whereas there is no ascending or descending path between incomparable elements, like B and C. Indeed, B and E are the same age, but E has more drugs prescriptions than B, thus we expect E to be more frail than B; C and E have the same value for drugs prescriptions, but C is younger than E, thus we expect E to be more frail than C. However, the comparison between C and B is not possible because C is younger than B, but also has a higher value for drugs prescriptions.

As a consequence of incomparabilities, several possible rankings of subjects according to their frailty level are possible. All the possible rankings of elements in the poset that respect its comparabilities (the connections in the Hasse diagram) and incomparabilities (Brüggemann and Patil, 2011) are called linear extensions (Figure 2, part (c)). In order to summarize all the information provided by linear extensions and to give an idea about the position that each node (profile) assumes in the whole poset of elements, it is possible to compute the average rank (AR). The AR of a node represents the mean of all the ranks that the same element occupies in all possible linear extensions, starting from the known order relations, as listed in Figure 2 part (d).

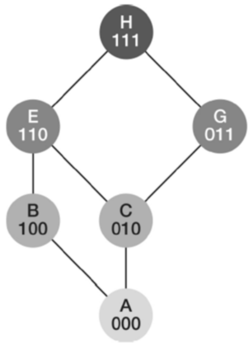
Then, the AR can be normalized in order to make it vary between 0 and 1, for sake of interpretability, using the min-max normalization method (Figure 2 part (d)). Thus, the normalized AR becomes an indicator that describes the relative position of an individual in the distribution of the latent concept, that in this work is frailty condition.

If the numbers of individuals and variables increase, the linear extensions become too many to be examined thoroughly, and it becomes computationally almost impossible to find the exact AR, as in the example in Figure 2. That said, researchers interested in the computation of AR have used two main approaches to obtain a computationally efficient calculation of the AR, by sampling linear extensions (Fattore, 2016; Lerche and Sorensen, 2003) or by the definition of an approximation formula (Brüggemann and Carlsen, 2011; De Loof et al., 2011).

Different approximation formulas have been proposed in the literature, such as the Local Partial Order Model (Brüggemann and Carlsen, 2011), or the one based on Mutual Probabilities (De Loof et al., 2011). The present work is based on De Loof's approach (2011) because it provides better results than other methods in terms of accuracy with a large sample size (De Loof et al., 2011). The approximated AR was computed using R software, using the package proposed by Caperna (2019), that can cope with large datasets (Boccuzzo and G., 2017; Caperna and Boccuzzo, 2018; Caperna, 2016).

	Age	Drugs	First Aid
A	0	0	0
B	1	0	0
C	0	1	0
E	1	1	0
G	0	1	1
H	1	1	1

(a)



(b)

H	H	H	H	H
E	E	E	G	G
G	G	B	E	E
C	B	G	B	C
B	C	C	C	B
A	A	A	A	A

(c)

H	→	6	→	1
E	→	4.6	→	0.72
G	→	4.2	→	0.64
B	→	2.8	→	0.36
C	→	2.4	→	0.28
A	→	1	→	0

(d)

Figure 2: A toy example to understand poset theory: part (a) contains a group of observations, part (b) represents the Hasse diagram of subjects in the table (a); part (c) lists all linear extensions for the subjects; and in part (d) the exact average rank for individuals of table (a) and normalized values.

Given that the main aim of our composite indicator is to detect the frailest individuals compared with the population to which they belong, the AR fits the purpose since it orders individuals according to their frailty level. However, it is important to remember that the AR summarizes the relative position of subjects in the specific population they belong, thus, the indicator is strictly population dependent and it is not possible to interpret it in absolute terms, nor it is possible to use it to measure a distance between individuals because it represents ranks.

5.3 Variables selection based on POSET

A poset created with all the 15 selected variables could become too complicated, leading to information potentially being lost because of an excess of incomparabilities between the profiles. Moreover, from a practical point of view, the collection of such a high number of variables to replicate the indicator may be too demanding. Thus, it is preferable to reduce the number of variables that compose the frailty indicator. This second selection process follows a forward logic, in order to select the most parsimonious set of variables that maximizes the indicator's performance. At this stage the indicator's performance was evaluated by observing the Area Under the ROC curve (AUC) with respect to the six outcomes and, in order to have a unique measure to compare performances, we considered the sum of the six AUCs. This selection procedure consisted in computing the frailty indicator several times with a growing number of variables, at every iteration the variables were selected maximizing the sum of the AUCs of the six outcomes.

Given the fact that the age variable is fundamental to identify frail individuals, we selected age as the first variable and we computed 14 composite indicators with only two variables (age and another variable chosen among the remaining 14). The indicator composed by age and poliprescriptions was the one with the highest AUC among the 14 considered indicators. Thus, in order to select the third variable, we computed 13 composite indicators with three variables (age, poliprescriptions and another variable chosen among the remaining 13). The selection process is summarized in Table 4, where are reported all the AUCs at every step and the added variables. We continued with the same procedure until we produced all the

Table 4: Steps of the variables selection process with forward approach: AUCs for all the six outcomes and added variables.

Added Variables	N. of var.	AUC Death	Em. Hosp.	ER Red code	Avoid. Hosp.	Hip Fract.	Disability
Age + Poliprescriptions	2	0.780	0.667	0.675	0.742	0.743	0.739
+ Charlson Index	3	0.798	0.680	0.690	0.759	0.743	0.754
+ Disability	4	0.811	0.684	0.699	0.764	0.744	0.764
+ N. ER access with Yellow code	5	0.815	0.687	0.705	0.769	0.749	0.769
+ N. ER access with Green code	6	0.815	0.695	0.705	0.773	0.750	0.776
+ Parkinson	7	0.816	0.695	0.706	0.773	0.753	0.781
+ Neoplasia	8	0.820	0.696	0.702	0.769	0.746	0.783

indicators with 8 variables and we noticed that even the indicator that maximizes the sum of the AUCs, worsens the performance of the indicator with respect to the one with 7 variables (as shown in Figure 3). This is because the addition of the eighth variable increases the incomparabilities and the entropy of the poset. Moreover, the computational time to produce the composite indicator with 8 variables is around 30 minutes (as shown in Figure 3), that is considerably higher that the time used by the same R function to compute the indicator with only one variable less (around 10 minutes).

For these reasons and for sake of parsimony, we preferred the indicator that involves only 7 variables: age, poliprescriptions, Charlson Index, disability, number of accesses to the emergency room with yellow code, number of accesses to the emergency room with green code and the presence of the Parkinson disease.

6 Results

6.1 Descriptive statistics

The frailty indicator varies between 0 and 1, as explained in section 5.2. However, it assumes quite low values on average, indeed its mean is 0.104 in the whole population. More than 10% of the population has indicator's values equal to 0 (the lowest level of frailty). These individuals are between 65 and 69 years old and do not present any of the risk factors included in the composite indicator. Around 2% of the whole population has the indicator higher than 0.5, indeed the 99th percentile is equal to 0.586.

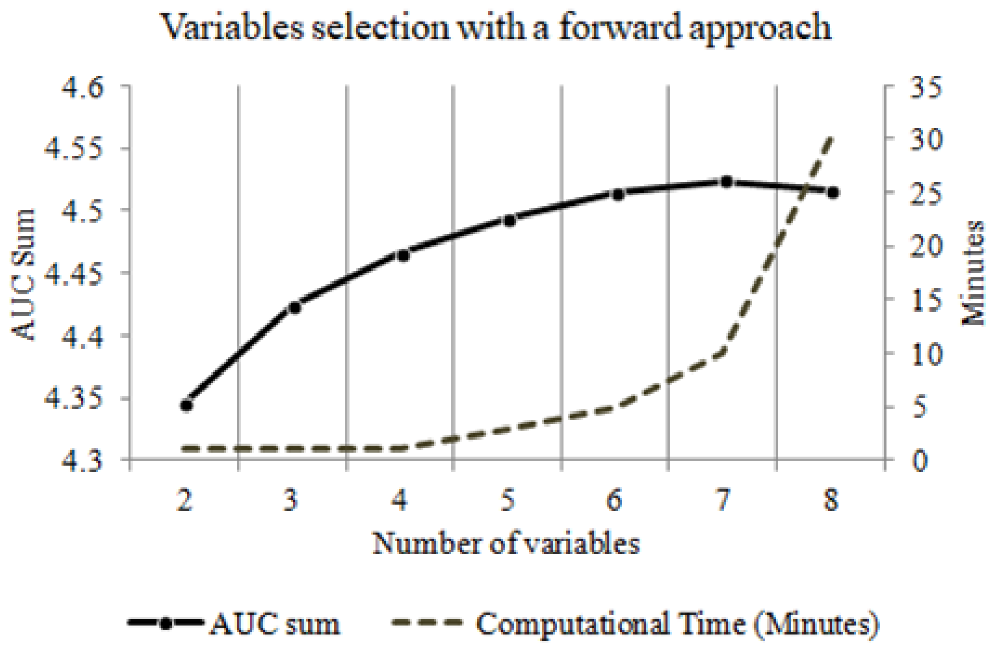


Figure 3: AUC sum and computation time for every step of the variables selection process with a forward approach.

Among women, the mean of the indicator is equal to 0.109, that is quite higher with respect to men (same situation is represented for median values, Figure 4), that is 0.097. However, observing the means standardized by age (assuming an age structure equal to the whole population for both men and women) this relationship is inverted, with standardized means equal to 0.101 for women and 0.107 for men.

Table 5: Distribution of the frailty indicator in the whole population.

Percentiles	0%	1%	5%	10%	25%	50%	75%	90%	95%	99%	100%
Frailty Indicator	0	0	0	0	0.012	0.048	0.145	0.276	0.377	0.586	1

The median of the frailty indicator in the whole population is 0.048, so it is quite low with respect to the total range of the indicator, but we may observe some differences in the median if we consider different subgroups of the population. In Figure 4 are represented medians and interquartile ranges for some variables that, except for gender, are included in the computation of the composite indicator. Variables included in the indicator are ordinal with same direction, thus, higher values of the variables correspond to higher values of the indicator.

Values for individuals with and without disability are quite separated, this means that the presence of disability is highly distinctive in the evaluation of the frailty level of every individual.

The same distinction is present also between values of the indicator for those subjects with and without co-morbidities. Individuals with no co-morbidities had

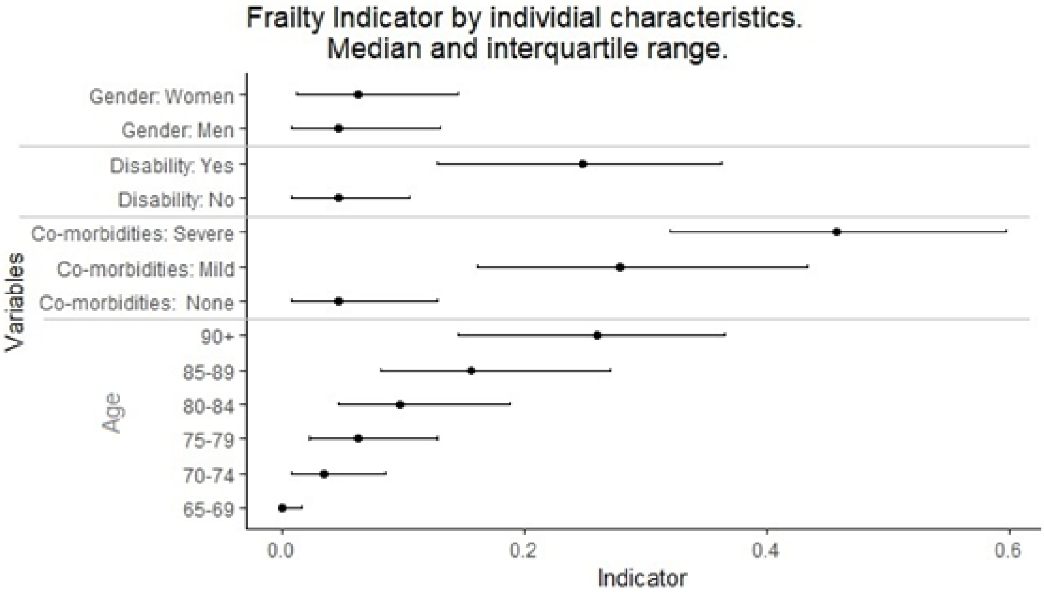


Figure 4: Frailty indicator by individual characteristics (Gender, Disability, Co-morbidities measured by Charlson Index and age). Medians and interquartile ranges.

no hospitalizations in 2014-2015 or had hospitalizations with only the principal diagnosis and no co-morbidities as secondary diagnosis. Mild and severe co-morbidities are assigned in the graph to individuals with at least one hospitalization and with at least one co-morbidity as secondary diagnosis that may be mild (Charlson Index equal to 1 or 2) or severe (Charlson Index higher than 2).

Interquartile ranges of mild and severe co-morbidities are quite large and overlapping. We can observe overlaps also in interquartile ranges of the age variable and medians that increase with age values. Indeed a healthy (without other negative conditions included in the composite indicator) over-80 years old person will have a lower value of the frailty indicator than a 65 years old with other conditions considered in the indicator.

The last aspect that we want to underline is that interquartile ranges of the age variable are larger in correspondence of high values of age. It may be explained by the fact that at older ages it is more common to have combinations of negative conditions included in the composite indicator, that may raise considerably the assigned value of the composite indicator because of implicit interactions among variables.

In conclusion, in Figure 4 we can observe some important characteristics of the frailty indicator, such as the presence or absence of overlapping interquartile ranges according to variables' distinctive power, the threshold relationship between the indicator and co-morbidities, and the presence of interactions among variables.

6.2 Performance of the indicator

As explained before, we built the frailty indicator having in mind that frail individuals have an increased susceptibility to adverse negative outcomes related with frailty condition, thus the first step to evaluate the performance of the indicator is to observe its relationship with the outcomes. The indicator is associated to high AUC with respect to all the six outcomes (0.816 for death, 0.706 for access to the ER with red code, 0.695 for emergency hospitalization, 0.773 for avoidable hospitalization, 0.753 for hip fracture and 0.781 for disability), as shown in Figure 5.

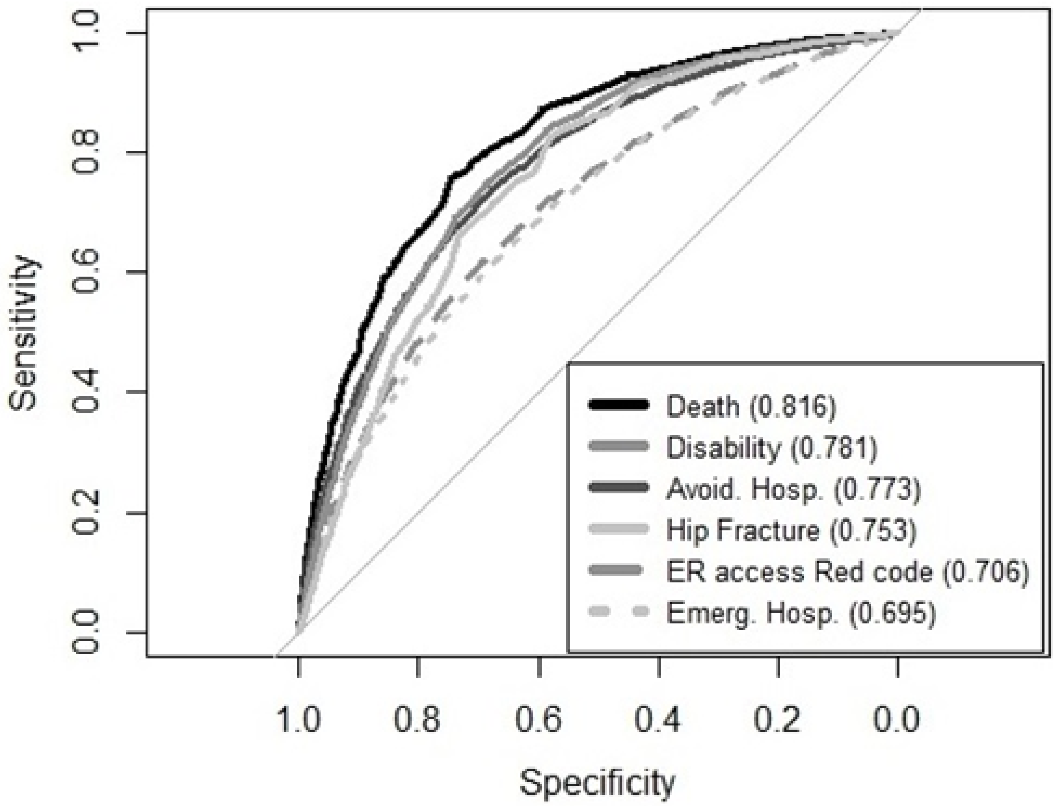


Figure 5: Roc curves for the six outcomes.

Moreover, the frailty indicator classifies well also individuals that received at least one outcome in 2016, regardless which was the outcome or outcomes. Indeed, the AUC for having at least one outcome is 0.769, but it increases if we consider a higher number of outcomes until the value of 0.845 for individuals that had at least 4 outcomes in 2016 (Figure 6).

The choice of variables to be included in the frailty indicator was driven by their ability to predict the selected outcomes, so the good performance in terms of ability to predict negative outcomes related with frailty comes by construction. However, the frailty indicator assumes different values in different subgroups of the population, for instance, individuals with chronic diseases have higher values of the indicator than other subjects, even for diseases that are not directly included in the computation of the frailty indicator (Figure 7). Parkinson disease is one of the variables that are included in the computation of the composite indicator for frailty, thus medians and interquartile range are quite far and not overlapping with respect to individuals with and without the disease. However, this is not the only

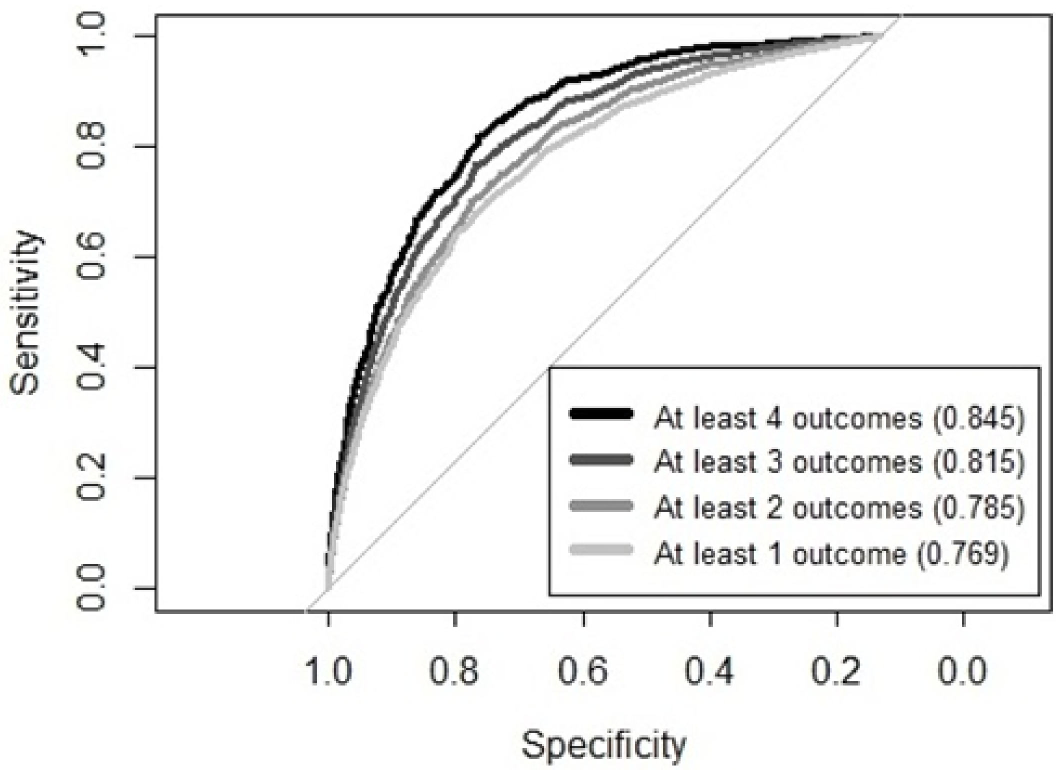


Figure 6: ROC curves for increasing number of outcomes suffered in 2016.

case in which we find such differences between the healthy and the sick groups, it happens also for other chronic conditions not included in the frailty indicator, such as dementia, chronic obstructive pulmonary disease and chronic heart failure.

Even if some socio-economic variables were available from the 2011 census, the variables selection process discarded all of them in favor of variables that represent use of services and chronic conditions. Some concerns about this indicator may regard its ability to represent and take into account also socio-economic aspects of individuals and their conditions. The empirical cumulative distribution functions of indicator's values for three different education levels are represented in Figure 8, distinguishing between individuals with no title, with just primary education and with secondary or higher education levels. Mean values of the indicator are quite different and inversely proportional with respect to the three education levels: 0.076 for secondary or higher degrees, 0.123 for primary degree and 0.177 for individuals with no titles. Some of these differences are due to different age structure in the three groups, thus we computed and compared also the standardized means (assuming an age structure equal to the whole population for all the education levels). The relationship between frailty indicator and education remains, but is slightly weakened.

In conclusion, observing differences in the indicator's values for subgroups that are not directly represented as variables in the computation of the indicator means that the small set of 7 chosen variables in the two described selection processes is extremely powerful. Even if the variables used for the computation of the frailty indicator are just 7, our composite indicator for frailty is able to predict the six outcomes and describe several individual characteristics (such as chronic conditions and education level).

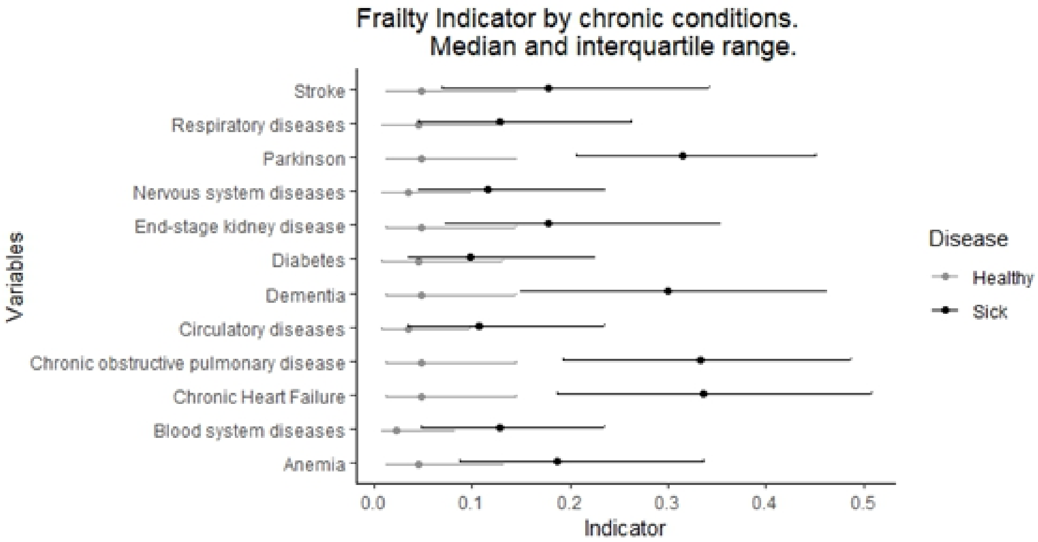


Figure 7: Frailty indicator by chronic conditions. Medians and interquartile ranges.

6.3 Success of our proposal

Given that a unique definition for frailty does not exist and that it is not possible to observe and measure the real value of frailty for every individual, the validation of the frailty indicator is not trivial. However, we may get some indications by Rockwood's work (2005), that enlists some characteristics that an operative definition of frailty should have in order to be successful. These criteria for a successful definition of frailty are grouped into three categories: content validity, construct validity and criterion validity.

Content validity refers to whether the definition makes sense on first principles (Rockwood, 2005). Among the criteria to assess content validity, there are the facts that the definition should include multiple determinants, but should also be dynamic and computationally tractable at the same time. Our proposed frailty indicator meets these characteristics. Indeed, it includes multiple determinants in three different ways: it includes 7 variables that represent different aspects of elderly, such as age, use of services, chronic conditions and health characteristics; the variables are selected according to six different outcomes that represent different shades of frailty; and, as pointed out in section 6.2, the indicator indirectly includes also some other aspects such as chronic conditions and education level. The dynamicity of the indicator is given by the poset methodology, indeed, even if no assumptions are required for its computation, it is able to catch interactions and non linearity in the data. Moreover, assuming that the same set of variables (that were chosen in a

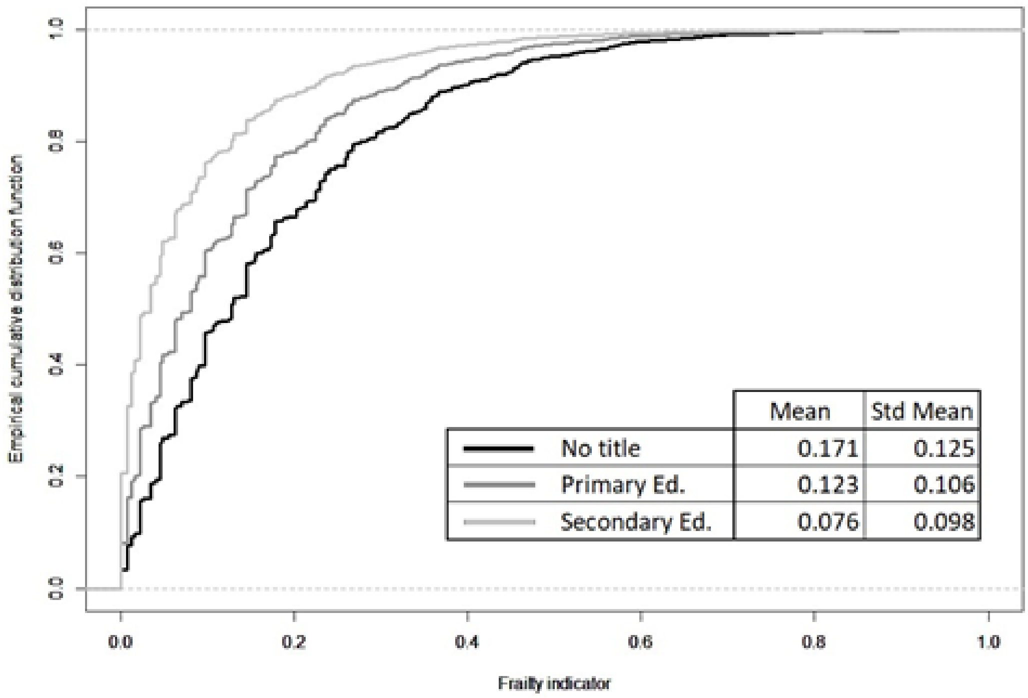


Figure 8: Empirical cumulative distribution functions of the frailty indicator by education level.

very robust way) is important to measure frailty on any population, it is possible to compute our frailty indicator even in different populations or in the same population in different years, without any needs to repeat the variables selection. Thanks to the R function (Caperna, 2019) it is possible to compute values of the indicator in around 10 minutes (on a population of 1095613 subjects), thus the indicator is also computationally tractable.

Construct validity exists when the operational definition coheres with other measures of the phenomenon: frailty should be more common among women than men, higher frailty indicator values at higher ages, related to disability and related to co-morbidity (Rockwood, 2005). As shown in section 6, our frailty indicator is on average higher for women than for men. Moreover, by construction, its values are strictly related with age, disability and co-morbidities because these are variables included in the computation of the indicator.

Criterion validity deals with the ability of the frailty definition to predict adverse outcomes, especially death (Rockwood, 2005). In section 6.2 are reported AUCs of our indicator with respect to six outcomes related with frailty condition. Our indicator has the best performance with respect to death (AUC equal to 0.816), but still good performances doe the other outcomes, thus we may say that it meets also criterion validity.

In conclusion, our frailty indicator has good properties in general and meets many of the requirements that in literature are asked for a successful definition of frailty.

7 Discussion

In the context of the regional implementation of the National Plan for Chronic Diseases (NPCD), there was a need to identify multimorbid patients at higher risk of clinical worsening, for whom preventive actions could be taken and who may benefit from better disease management. We developed a frailty indicator that provides a score for all over-65 years old in Piedmont Region, thus a graduated classification of individuals useful for this purpose. This indicator provides a simplification of a complex and multidimensional concept as frailty is. It can be particularly useful for the implementation of the NPCD at the local level, because it is built with a small set of variables obtained from current administrative health care data flows, usually available in all Italian Local Health Units. Even if used methods are quite sophisticated, the indicator is easy to calculate, thanks to the construction of a user-friendly application that returns the individual frailty indicator after a guided loading of a suitable input data file. Moreover, the same frailty indicator might become also an individual variable to use for further analysis, since it represents individual health condition and predicts negative outcomes related with frailty condition. As mentioned before, the validation of the frailty indicator is not an easy task. As a first step towards a more thorough validation analysis, we compared our results with the subjective judgement of a small sample of 10 general practitioners (GPs), who joined the project on a voluntary basis. We therefore selected all the patients aged 65+ and asked the GPs to classify them, according to their knowledge of the clinical and social conditions of their patients, into three classes: not frail, at risk of frailty, and definitely frail. We then analyzed the performance of our frail indicator in the 3 classes of patients identified by the GPs and found statistically significant differences among the groups: the mean score was 0.09 (95%CI: 0.08-0.10) in the first class of not frail people (n=1402); 0.14 (95%CI: 0.13-0.15) in the second class of patients at risk of frailty (n=742); and 0.23 (95%CI: 022-0.24) in the class of people identified as definitely frail (n=839). Although based on a small and selected sample of GPs, these results go in the expected direction of increasing score with increasing frailty, and there is no reason to think that the algorithm should perform differently among other groups of patients. Next steps of validation will involve both the validation of the indicator in different populations and in time and a sensitivity analysis with respect to construction and selection of variables. In conclusion, this frailty indicator appears to be a promising tool for the identification of patients to whom preventive actions should be addressed as a priority. Once its validity and applicability have been further evaluated, it could become a useful tool for the clinical management of multichronic patients and be transferred also to other Italian regions.

References

Ahn, E., Kim, J., Rahman, K., Baldacchino, T., and Baird, C. (2018). Development of a risk predictive scoring system to identify patients at risk of representation to emergency department: a retrospective population-based analysis in australia. *BMJ open*, 8(9):e021323.

Ávila-Funes, J. A., Amieva, H., Barberger-Gateau, P., Le Goff, M., Raoux, N., Ritchie, K., Carriere, I., Tavernier, B., Tzourio, C., Gutiérrez-Robledo, L. M., et al. (2009). Cognitive impairment improves the predictive validity of the phenotype of frailty for adverse health outcomes: the three-city study. *Journal of the American Geriatrics Society*, 57(3):453–461.

Avila-Funes, J. A., Carcaillon, L., Helmer, C., Carrière, I., Ritchie, K., Rouaud, O., Tzourio, C., Dartigues, J.-F., and Amieva, H. (2012). Is frailty a prodromal stage of vascular dementia? results from the three-city study. *Journal of the American Geriatrics Society*, 60(9):1708–1712.

Bergman, H., François, B., and Perrault, A. (2002). The global challenge of understanding and meeting the needs of the frail older population. *Aging clinical and experimental research*, 14(4):223–225.

Billings, J., Georghiou, T., Blunt, I., and Bardsley, M. (2013). Choosing a model to predict hospital admission: an observational study of new variants of predictive models for case finding. *BMJ open*, 3(8):e003352.

Boccuzzo, G. and G., C. (2017). *Complexity in society: From indicators construction to their synthesis*, chapter Evaluation of life satisfaction in Italy: proposal of a synthetic measure based on poset theory, page 291–321. Cham, Switzerland: Springer International Publishing.

Bortz, W. M. (2002). A conceptual framework of frailty: a review. *The Journals of Gerontology Series A: Biological Sciences and Medical Sciences*, 57(5):M283–M288.

Brüggemann, R. and Carlsen, L. (2011). An improved estimation of averaged ranks of partial orders. *MATCH Communications in Mathematical and in Computer Chemistry*, 65:383–414.

Brüggemann, R. and Patil, G. P. (2011). *Ranking and Prioritization with Multiple Indicators - Introduction to Partial Order Applications*. New-York: Springer-Verlag.

Brody, K. K., Johnson, R. E., and Ried, L. D. (1997). Evaluation of a self-report screening instrument to predict frailty outcomes in aging populations. *The Gerontologist*, 37(2):182–191.

Brody, K. K., Johnson, R. E., Ried, L. D., Carder, P. C., and Perrin, N. (2002). A comparison of two methods for identifying frail medicare-aged persons. *Journal of the American Geriatrics Society*, 50(3):562–569.

Buchman, A. S., Boyle, P. A., Wilson, R. S., Tang, Y., and Bennett, D. A. (2007). Frailty is associated with incident alzheimer's disease and cognitive decline in the elderly. *Psychosomatic medicine*, 69(5):483–489.

Buchner, D. and Wagner, E. (1992). Preventing frail health. *Clin Geriatr Med*, 8:1–17.

Campbell, A. and Buchner, D. (1997). Unstable disability and the fluctuations of frailty. *Age Ageing*, 26:315–318.

Caperna, G. (2016). *Partial Order Theory for Synthetic Indicators*. PhD thesis, University of Padova. (Available from http://paduaresearch.cab.unipd.it/9588/).

Caperna, G. (2019). Approximation of averagerank by means of a formula.

Caperna, G. and Boccuzzo, G. (2018). Use of poset theory with big datasets: A new proposal applied to the analysis of life satisfaction in Italy. *Social Indicators Research*, 136(3):1071–1088.

Carrière, I., Colvez, A., Favier, F., Jeandel, C., Blain, H., study group, E., et al. (2005). Hierarchical components of physical frailty predicted incidence of dependency in a cohort of elderly women. *Journal of clinical epidemiology*, 58(11):1180–1187.

Cavazza, G. and Malvi, C. (2014). *La fragilità degli anziani. Strategie, progetti, strumenti per invecchiare bene.* Santarcangelo di Romagna: Maggioli Editore.

Charlson, M. E., Pompei, P., Ales, K. L., and MacKenzie, C. R. (1987). A new method of classifying prognostic comorbidity in longitudinal studies: development and validation. *Journal of chronic diseases*, 40(5):373–383.

Crane, S. J., Tung, E. E., Hanson, G. J., Cha, S., Chaudhry, R., and Takahashi, P. Y. (2010). Use of an electronic administrative database to identify older community dwelling adults at high-risk for hospitalization or emergency department visits: the elders risk assessment index. *BMC health services research*, 10(1):338.

Damush, T. M., Smith, D. M., Perkins, A. J., Dexter, P. R., and Smith, F. (2004). Risk factors for nonelective hospitalization in frail and older adult, inner-city outpatients. *The Gerontologist*, 44(1):68–75.

Davey, B. A. and Priestley, H. A. (2002). *Introduction to lattices and order*. Cambridge university press.

De Groot, V., Beckerman, H., Lankhorst, G. J., and Bouter, L. M. (2003). How to measure comorbidity: a critical review of available methods. *J. Clin. Epidemiol.*, 56:221–229.

De Loof, K., De Baets, B., and De Meyer, H. (2011). Approximation of average ranks in posets. *MATCH Communications in Mathematical and in Computer Chemistry*, 66:219–229.

De Vries, O., Peeters, G., Lips, P., and Deeg, D. (2013). Does frailty predict increased risk of falls and fractures? a prospective population-based study. *Osteoporosis international*, 24(9):2397–2403.

Deyo, R., Cherkin, D. C., and Ciol, M. A. (1993). Adapting a clinical comorbidity index for use with icd-9-cm administrative data: a response. *J. Clin. Epidemiol.*, 45:613–619.

Eurostat (2019). Ageing europe — looking at the lives of older people in the eu. Last Accessed: 31 October 2019.

Falasca, P., Berardo, A., and Di Tommaso, F. (2011). Development and validation of predictive mosaico (modello statistico combinato) on emergency admissions: can it also identify patients at high risk of frailty? *Ann. Ist. Super. Sanità*, 47:220–228.

Fattore, M. (2016). Partially ordered sets and the measurement of multidimensional ordinal deprivation. *Social Indicators Research*, 128(2):835–858.

Fried, L. P., Tangen, C. M., Walston, J., Newman, A. B., Hirsch, C., Gottdiener, J., Seeman, T., Tracy, R., Kop, W. J., Burke, G., et al. (2001). Frailty in older adults: evidence for a phenotype. *The Journals of Gerontology Series A: Biological Sciences and Medical Sciences*, 56(3):M146–M157.

Gao, J., Moran, E., Li, Y. F., and Almenoff, P. L. (2014). Predicting potentially avoidable hospitalizations. *Medical care*, 52(2):164–171.

Gillick, M. (2001). Guest editorial: pinning down frailty. *The Journals of Gerontology Series A: Biological Sciences and Medical Sciences*, 56(3):M134–M135.

Gobbens, R., Luijkx, K., Wijnen-Sponselee, M. T., and Schols, J. (2010). In search of an integral conceptual definition of frailty: opinions of experts. *J. Am. Med. Dir. Assoc.*, 11:338–343.

Gray, S. L., Anderson, M. L., Hubbard, R. A., LaCroix, A., Crane, P. K., McCormick, W., Bowen, J. D., McCurry, S. M., and Larson, E. B. (2013). Frailty and incident dementia. *The Journals of Gerontology Series A: Biological Sciences and Medical Sciences*.

Hibbard, J. H., Greene, J., Sacks, R. M., Overton, V., and Parrotta, C. .-. (2017). Improving population health management strategies: identifying patients who are more likely to be users of avoidable costly care and those more likely to develop a new chronic disease. *Health services research*, 52(4):1297–1309.

Hippisley-Cox, J. and Coupland, C. (2013). Predicting risk of emergency admission to hospital using primary care data: derivation and validation of qadmissions score. *BMJ open*, 3(8).

Huang, Y., Gou, R., Diao, Y., Yin, Q., Fan, W., and Liang, Y. (2014). Charlson comorbidity index helps predict the risk of mortality for patients with type 2 diabetic nephropathy. *J. Biomed. Biotechnol.*, 15:58–66.

Inouye, S. K., Zhang, Y., Jones, R. N., Shi, P., Cupples, L. A., Calderon, H. N., and Marcantonio, E. R. (2008). Risk factors for hospitalization among community-dwelling primary care older patients: development and validation of a predictive model. *Medical care*, 46(7):726.

Jones, D. M., Song, X., and Rockwood, K. (2004). Operationalizing a frailty index from a standardized comprehensive geriatric assessment. *Journal of the American Geriatrics Society*, 52(11):1929–1933.

Kamaruzzaman, S., Ploubidis, G. B., Fletcher, A., and Ebrahim, S. (2010). A reliable measure of frailty for a community dwelling older population. *Health and quality of life outcomes*, 8(1):123.

Klein, B. E., Klein, R., Knudtson, M. D., and Lee, K. E. (2005). Frailty, morbidity and survival. *Archives of gerontology and geriatrics*, 41(2):141–149.

Landi, F., Onder, G., Cesari, M., Barillaro, C., Lattanzio, F., Carbonin, P. U., and Bernabei, R. (2004). Comorbidity and social factors predicted hospitalization in frail elderly patients. *Journal of clinical epidemiology*, 57(8):832–836.

Lerche, D. and Sorensen, P. (2003). Evaluation of the ranking probabilities for partial orders based on random linear extensions. *Chemosphere*, 53:981–992.

López-Aguilà, S., Contel, J., Farre, J., Campuzano, J., and Rajmil, L. (2011). Predictive model for emergency hospital admission and 6-month readmission. *The American journal of managed care*, 17(9):e348–57.

Louis, D. Z., Robeson, M., McAna, J., Maio, V., Keith, S. W., Liu, M., Gonnella, J. S., and Grilli, R. (2014). Predicting risk of hospitalisation or death: a retrospective population-based analysis. *BMJ open*, 4(9):p.e005223.

Luppa, M., Luck, T., Weyerer, S., König, H. H., Brähler, E., and Riedel-Heller, S. G. (2009). Prediction of institutionalization in the elderly. a systematic review. *Age and ageing*.

Makary, M. A., Segev, D. L., Pronovost, P. J., Syin, D., Bandeen-Roche, K., Patel, P., Takenaga, R., Devgan, L., Holzmueller, C. G., Tian, J., and others901 908 (2010). Frailty as a predictor of surgical outcomes in older patients. *Journal of the American College of Surgeons*, 2010(6):901–908.

Mazzaglia, G., Roti, L., Corsini, G., Colombini, A., Maciocco, G., Marchionni, N., Buiatti, E., Ferrucci, L., and Di Bari, M. (2007). Screening of older community-dwelling people at risk for death and hospitalization: The assistenza socio-sanitaria in italia project. *Journal of the American Geriatrics Society*, 55(12):1955–1960.

Ministero della Salute (2016). Piano nazionale della cronicità. Last access: 31 October 2019.

Mitnitski, A., J., M. A., C., M., and K., R. (2002). The mortality rate as a function of accumulated deficits in a frailty index. *Mech. Ageing Dev.*, 123:1457–1460.

Morley, J. E., Perry III, H. M., and Miller, D. K. (2002). Something about frailty. *The Journals of Gerontology Series A: Biological Sciences and Medical Sciences*, 57(11).

Paw, M. J. C. A., Dekker, J. M., Feskens, E. J., Schouten, E. G., and Kromhout, D. (1999). How to select a frail elderly population? a comparison of three working definitions. *Journal of clinical epidemiology*, 52(11):1015–1021.

Ravaglia, G., Forti, P., Lucicesare, A., Pisacane, N., Rietti, E., and Patterson, C. (2008). Development of an easy prognostic score for frailty outcomes in the aged. *Age and Ageing*, 37(2):161–166.

Rockwood, K. (2005). What would make a definition of frailty successful? *Age and ageing*, 34(5):432–434.

Rockwood, K., Mitnitski, A., Song, X., Steen, B., and Skoog, I. (2006). Long-term risks of death and institutionalization of elderly people in relation to deficit accumulation at age 70. *Journal of the American Geriatrics Society*, 54(6):975–979.

Roppolo, M., Mulasso, A., Gobbens, R. J., Mosso, C. O., and Rabaglietti, E. (2015). A comparison between uni-and multidimensional frailty measures: prevalence, functional status, and relationships with disability. *Clinical interventions in aging*, 10:1669.

Rothman, M. D., Leo-Summers, L., and Gill, T. M. (2008). Prognostic significance of potential frailty criteria. *Journal of the American Geriatrics Society*, 56(12):2211–2216.

Saliba, D., Elliott, M., Rubenstein, L. Z., Solomon, D. H., Young, R. T., Kamberg, C. J., Carol Roth, R., MacLean, C. H., Shekelle, P. G., Sloss, E. M., et al. (2001). The vulnerable elders survey: a tool for identifying vulnerable older people in the community. *Journal of the American Geriatrics Society*, 49(12):1691–1699.

Schuurmans, H., Steverink, N., Lindenberg, S., Frieswijk, N., and Slaets, J. P. (2004). Old or frail: what tells us more? *The Journals of Gerontology Series A: Biological Sciences and Medical Sciences*, 59(9):M962–M965.

Shelton, P., Sager, M. A., and Schraeder, C. (2000). The community assessment risk screen (cars): identifying elderly persons at risk for hospitalization or emergency department visit. *Am J Manag Care*, 6(8):925–33.

Silan, M., Caperna, G., and Boccuzzo, G. (2019). Quantifying frailty in older people at an italian local health unit: A proposal based on partially ordered sets. *Social Indicators Research*, 146(3):757–782.

Solfrizzi, V., Scafato, E., Frisardi, V., Seripa, D., Logroscino, G., Maggi, S., Imbimbo, B. P., Galluzzo, L., Baldereschi, M., Gandin, C., et al. (2013). Frailty syndrome and the risk of vascular dementia: the italian longitudinal study on aging. *Alzheimer's & Dementia*, 9(2):113–122.

Song, X., Mitnitski, A., and Rockwood, K. (2014). Age-related deficit accumulation and the risk of late-life dementia. *Alzheimer's research & therapy*, 6(5):54.

Strawbridge, W. J., Shema, S. J., Balfour, J. L., Higby, H. R., and Kaplan, G. A. (1998). Antecedents of frailty over three decades in an older cohort. *The Journals of Gerontology Series B: Psychological Sciences and Social Sciences*, 53(1):S9–S16.

Tammemagi, C. M., Neslund-Dudas, C., Simoff, M., and Kvale, P. (2004). In lung cancer patients, age, race-ethnicity, gender and smoking predict adverse comorbidity, which in turn predicts treatment and survival. *Journal of clinical epidemiology*, 57(6):597–609.

Walker, L., Jamrozik, K., and Wingfield, D. (2005). The sherbrooke questionnaire predicts use of emergency services. *Age and ageing*, 34(3):233–237.

Woods, N. F., LaCroix, A. Z., Gray, S. L., Aragaki, A., Cochrane, B. B., Brunner, R. L., Masaki, K., Murray, A., and Newman, A. B. (2005). Frailty: emergence and consequences in women aged 65 and older in the women's health initiative observational study. *Journal of the American Geriatrics Society*, 53(8):1321–1330.

Working Paper Series
Department of Statistical Sciences, University of Padua

You may order paper copies of the working papers by emailing wp@stat.unipd.it

Most of the working papers can also be found at the following url: http://wp.stat.unipd.it