

Università degli Studi di Padova
Dipartimento di Scienze Statistiche
Corso di Laurea Magistrale in
Scienze Statistiche



**Identificazione e quantificazione della fragilità negli
anziani: un approccio multivariato basato sull'uso di flussi
sanitari correnti**

Relatore Prof. Giovanna Boccuzzo
Dipartimento di Scienze Statistiche

Correlatori Dott. Luca Benacchio
Dipartimento di Prevenzione, Ulss 15 "Alta Padovana"

Dott. Giulio Caperna
Dipartimento di Scienze Statistiche

Laureando: Nicolò Preo
Matricola N. 1092456

Anno Accademico 2016/2017

Indice

Introduzione.....	5
1 Analisi della fragilità nella letteratura: definizioni, marcatori, fattori di rischio e aspetti critici ..	9
1.1 Definizioni di fragilità	9
1.2 I marcatori di fragilità (outcomes) in letteratura	13
1.3 Le variabili esplicative della fragilità in letteratura	14
1.3 Aspetti salienti nello studio della fragilità.....	21
2 I dati.....	23
2.1 I flussi informativi	24
2.1.1 Anagrafe Sanitaria Regionale.....	24
2.1.2 Schede di Dimissione Ospedaliera	24
2.1.3 Pronto Soccorso.....	25
2.1.4 Assistenza Psichiatrica Territoriale	25
2.1.5 Assistenza Domiciliare Integrata.....	25
2.1.6 Esenzioni ticket	26
2.1.7 Farmaceutica Territoriale	26
2.2 Il record linkage delle fonti	27
2.3 Gli outcomes disponibili nei dati.....	28
2.3.1 Decesso.....	28
2.3.2 Ricovero urgente	29
2.3.3 Disabilità	29
2.3.4 Demenza.....	30
2.3.5 Frattura	31
2.4 Gli outcomes esclusi dall'analisi.....	32
2.5 Le variabili esplicative disponibili nei dati.....	33
2.5.1 Variabili socio-demografiche	33
2.5.2 Variabili relative allo stile di vita	34
2.5.3 Variabili relative a ricoveri precedenti	34
2.5.4 Variabili relative a malattie croniche.....	34
2.5.5 Variabili relative all'assunzione di farmaci.....	35
2.5.6 Variabili relative allo stato fisiologico	36
2.5.7 Variabili relative allo stato psicologico o mentale	36
3 Selezione delle variabili esplicative della fragilità	39
3.1 La codifica delle variabili	39
3.2 Analisi preliminari sui fattori di rischio	41

3.3 Metodi di selezione delle variabili	45
3.4 Selezione Stepwise e Lasso a confronto: analisi e risultati	48
3.4.1 Decesso	50
3.4.2 Ricovero urgente	51
3.4.3 Frattura	52
3.4.4 Demenza	53
3.4.5 Disabilità	54
3.5 Criterio per la selezione delle variabili	55
4 Gli ordinamenti parziali: un metodo per la valutazione della fragilità negli anziani	61
4.1 Definizione di Poset	62
4.1.1 Un esempio con i dati sanitari	62
4.1.2 Definizioni di base	63
4.2 Diagramma di Hasse	65
4.3 Estensioni lineari	65
4.4 Le approssimazioni dell'average rank	66
4.4.1 LPOM: Local Partial Order Model	67
4.4.2 LPOMext: Extended Linear Partial Order Model	67
4.4.3 L'approccio "Mutual Probabilities"	68
4.5 Il poset come strumento operativo per l'Ulss 15	70
4.6 Applicazione del poset sugli utenti anziani dell'Ulss 15: i risultati	70
Conclusioni	79
Bibliografia	81
Sitografia	88

Introduzione

Il presente elaborato è il risultato di un'esperienza di stage svolta nell'Unità di Epidemiologia del Dipartimento di Prevenzione dell'Ulss 15 "Alta Padovana", finalizzata a identificare gli individui più fragili nel territorio di competenza.

La fragilità è un concetto in continua evoluzione da quando venne menzionata per la prima volta, nel 1985, dallo studioso statunitense M.D. Fretwell e da una pubblicazione scientifica curata dal National Institute on Aging degli Stati Uniti d'America (Zucchelli e Silvi, 2001). Inizialmente un individuo veniva definito fragile se presentava condizioni di comorbidità e di disabilità, necessitando quindi di un supporto nello svolgimento delle attività quotidiane (Boers & Jentoft, 2015). Questa definizione si è rivelata nel tempo troppo semplicistica e inadeguata a rappresentare i tanti ambiti in cui la fragilità viene presa in considerazione: sono molti i professionisti (psicologi, medici di base, geriatri, infermieri, assistenti sociali) che vengono a contatto con aspetti diversi della fragilità, che oramai viene riconosciuta come un concetto complesso e multidimensionale (Cavazza e Malvi, 2014). Pertanto non è semplice individuare una definizione univoca e condivisa da tutti: la letteratura abbonda di differenti criteri e modelli per identificare la fragilità, anche se sembra emergere un crescente consenso per l'idea che si possa riassumere come uno stato di ridotta omeostasi¹ e di scarsa resistenza agli stressor² (eventi climatici, ricoveri, etc.), che portano ad un aumento della vulnerabilità e del rischio di conseguenze avverse come le malattie, le cadute, la disabilità e il decesso (Bergman, 2004). Gli studi più recenti hanno cercato di individuare dei fattori di rischio della fragilità non solo fisiologici, ma anche psicologici e sociali. La complessità di fondo è dovuta all'interazione di questi fattori (fisiologici, biologici e sociali) che determinano e influenzano il livello di fragilità dell'individuo. De Vries (2011), tramite un'attenta revisione degli studi più

¹ Omeostasi - In biologia, l'attitudine propria degli organismi viventi, siano essi cellule, individui singoli o comunità, a mantenere in stato di equilibrio le proprie caratteristiche al variare delle condizioni esterne.

² Stressor – Termine che si riferisce a stimoli di diversa natura (fisici, ambientali, metabolici, psicologici, affettivi, alimentari) che causano stress all'organismo e alla psiche di una persona.

recenti, ha stabilito otto marcatori che più caratterizzano una persona fragile: carente stato nutrizionale, ridotta attività fisica, scarsa mobilità e mancanza di forza ed energia nell'ambito fisiologico; limitate capacità cognitive e depressione in ambito psicologico; la mancanza di contatti sociali e supporto umano nella dimensione sociale dell'individuo.

Nonostante i numerosi studi effettuati, non è ancora stato definito un metodo valido e univoco per valutare quali sono i soggetti anziani effettivamente fragili, ovvero i destinatari di eventuali servizi di assistenza (Fried, 2001).

Questo è un tema di ricerca centrale nel sistema sanitario per poter pianificare, programmare e ottimizzare i servizi a cui accede l'utenza. L'importanza di classificare gli utenti secondo il loro livello di fragilità, oltre ad avere uno scopo di programmazione sanitaria, mira anche alla prevenzione delle conseguenze negative della fragilità tramite interventi di gestione proattivi volti a evitare o a ritardare eventi quali il decesso e il ricovero d'urgenza riducendo di conseguenza i costi della sanità (Cavazza e Malvi, 2014). Nel contesto locale dell'Ulss 15 questo obiettivo può essere raggiunto tramite la collaborazione con i servizi sociali che, pur essendo molto presenti sul territorio, non hanno l'intera copertura dei soggetti bisognosi di assistenza, poiché alcuni potenziali utenti ignorano tali servizi e i benefici che ne possono trarre. Grazie alla fiducia che si instaura tra l'utente e il dipendente dei servizi sociali, è inoltre possibile approfondire le problematiche degli utenti tramite delle scale di valutazione sulla salute, permettendo interventi mirati ove necessario. In un contesto di risorse limitate come quello della sanità pubblica, riuscire a individuare con una procedura statisticamente valida un sottoinsieme della popolazione a maggior rischio di eventi negativi porta vantaggi sia alla qualità della vita dell'utente assistito sia al bilancio economico dell'azienda.

L'esigenza di monitorare gli anziani fragili assume ancora più importanza in concomitanza con eventi particolari, come le ondate di calore delle estati 2003 e 2015 o l'epidemia influenzale dell'inverno tra il 2014 e il 2015, in cui la mortalità tra gli anziani ha raggiunto picchi anomali (Il Sole 24 Ore, 2016).

In Italia sono state adottate diverse tipologie di analisi e classificazione dei soggetti fragili. Per esempio in Veneto dal 2014 è stato adottato l'ACG (Adjusted Clinical Groups), un sistema che identifica e raggruppa soggetti confrontabili dal punto di vista del rischio di esiti negativi, in particolar modo focalizzandosi sul consumo di risorse assistenziali, tramite l'informazione presente nei principali flussi informativi già disponibili per scopi gestionali.

Questo strumento si è rivelato utile per la gestione e la programmazione di spese e di interventi nell'Ulss 15, ma non si è dimostrato sufficientemente adeguato nel predire le conseguenze sfavorevoli legate alla condizione di fragilità (Silan, 2015).

Un altro esperimento di classificazione riguarda lo sviluppo e la validazione del Modello Statistico Combinato (MoSaiCo), utilizzato dall'azienda USL di Ravenna per identificare individui ad alto rischio di ricovero non programmato. Il metodo consiste in un modello multivariato di regressione logistica i cui principali fattori predittivi includono variabili demografiche, sociali, cliniche e di uso dei servizi sanitari e sociali (Falasca, 2011). I limiti di questo modello sono l'alto numero di variabili esplicative utilizzate e la scarsa capacità predittiva, dovuta al mancato bilanciamento dei casi e dei controlli.

In definitiva, è necessario quindi rendere utile il concetto di fragilità, passando da una definizione concettuale a una operativa, pur con due limiti principali: il primo risiede nella difficoltà di misurare un fenomeno dinamico e continuo nel tempo (Bergman, 2007) tramite degli strumenti statici come i flussi amministrativi, in cui i dati non vengono aggiornati in tempo reale; il secondo limite riguarda la carenza informativa dei flussi relativamente a fattori di rischio come il consumo di tabacco e alcol, o altre informazioni che sono generalmente raccolte tramite questionari ad hoc negli studi dedicati all'individuazione dei soggetti fragili.

Il primo obiettivo del presente elaborato è quello di identificare gli outcomes e i fattori di rischio più comuni negli studi sulla fragilità, attraverso un'approfondita analisi della letteratura.

Il secondo obiettivo è quello di individuare un sottoinsieme statisticamente rilevante di fattori di rischio, desumibili dagli archivi sanitari, in modo da dare una

visione quanto più completa possibile di un fenomeno così complesso come la fragilità.

Il terzo obiettivo è quello di fornire all'azienda ospedaliera un algoritmo che, tramite le poche variabili selezionate, riesca a creare un rating di fragilità degli individui anziani (dai 65 anni) residenti nel territorio di competenza dell'Ulss 15 "Alta Padovana", in grado di predire efficacemente alcune conseguenze negative della fragilità.

La tesi si sviluppa nei seguenti capitoli:

- Capitolo 1: analisi della fragilità nella letteratura, definendo i fattori di rischio e gli esiti negativi legati ad essa.
- Capitolo 2: presentazione dei dati e dei flussi sanitari utilizzati nelle analisi statistiche.
- Capitolo 3: introduzione teorica alle tecniche di selezione delle variabili utilizzate per evidenziare i fattori di rischio della fragilità e presentazione dei risultati ottenuti.
- Capitolo 4: introduzione teorica dell'algoritmo di ordinamento degli individui anziani e presentazione dei risultati ottenuti in termini di previsione degli esiti negativi della fragilità.
- Conclusioni.

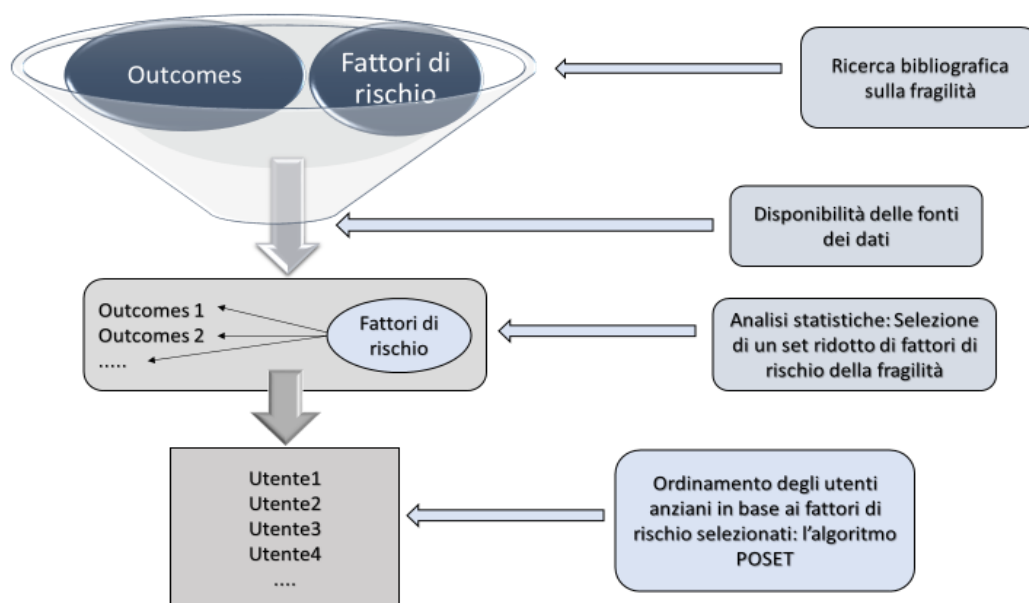


Figura 1.1: Struttura logica dell'elaborato.

Capitolo 1

Analisi della fragilità nella letteratura: definizioni, marcatori, fattori di rischio e aspetti critici

Non esiste in letteratura una definizione univoca di fragilità, ed è per questo che sono numerose le pubblicazioni relative a questo concetto.

In questo capitolo verranno proposte diverse definizioni di fragilità e delle metodologie di analisi, focalizzandosi in particolar modo sugli outcomes in base ai quali una persona è definita fragile e sulle variabili esplicative di tali outcomes, che più frequentemente sono state prese in considerazione. Questo lavoro preliminare è fondamentale per il proseguo dell'elaborato, sia per avere un quadro generale dell'argomento, sia perché in seguito si stabilirà quali outcomes e quali variabili esplicative possono essere studiate tra quelle trovate nella letteratura, sulla base delle fonti informative disponibili.

1.1 Definizioni di fragilità

Per riassumere le molteplici definizioni presenti in letteratura, si può affermare che la fragilità è uno stato di vulnerabilità in cui un soggetto è a rischio di esiti negativi, in particolare se esposto a degli stressor (Walston, 2006).

La fragilità può essere sia fisica che psicologica, o una combinazione delle due componenti, è correlata positivamente con l'età ed è una condizione dinamica che può migliorare o peggiorare nel tempo (Morley, 2013). Due approcci di misura della fragilità sono diventati particolarmente popolari in questo ambito di studio.

Il primo, ideato da Rockwood nel 1994, è un indice di fragilità che conta il numero di deficit che il soggetto possiede. Per deficit non si intendono solo quelli fisici (ovvero sintomi, segni, disabilità, malattie), ma anche aspetti cognitivi e psicosociali (ad esempio la depressione). Infatti Rockwood in una pubblicazione del 2007 ribadisce come la fragilità sia “uno stato di disorganizzazione caotica dei sistemi fisiologici che può essere stimato valutando lo stato funzionale, le malattie,

i deficit fisici e cognitivi, i fattori di rischio psicosociali e le sindromi geriatriche, nell'ottica di costruire un quadro il più possibile completo delle situazioni di rischio di eventi avversi”.

Tra le linee guida per la costruzione di questo indice, Searle (2008) elenca 5 criteri fondamentali che vanno rispettati:

1. I deficit devono essere associati allo stato di salute;
2. La prevalenza di un deficit deve crescere al crescere dell'età, sebbene alcune gravi malattie correlate all'età possono decrescere in prevalenza nell'età più avanzata a causa di un effetto della selezione, ovvero che sono sopravvissuti gli individui più sani.
3. I deficit selezionati non devono arrivare a saturazione troppo presto: per esempio la presbiopia è noto che si manifesta all'età di 55 anni, e pertanto va esclusa;
4. Nel gruppo devono essere presenti deficit provenienti da tutti i campi considerati (cognitivo, geriatrico, sociale);
5. L'indice deve essere confrontabile nel tempo, per cui i deficit considerati devono rimanere gli stessi.

Le variabili riferite ai deficit sono in gran parte dicotomiche, con il valore 0 ad indicare assenza e 1 ad indicare presenza, mentre in alcune possono essere previsti dei valori intermedi. Sommando i valori dei deficit si ricava così una misura in cui è possibile sia stabilire una soglia per definire se un soggetto è fragile o meno, sia studiare la probabilità che i deficit aumentino e che le condizioni iniziali del paziente peggiorino. Questa logica è coerente con l'idea dinamica di fragilità, in cui Rockwood individua un terzo gruppo di soggetti, i “pre-fragili”, ovvero soggetti in situazioni critiche, ma ancora in grado di vivere in una condizione d'indipendenza. Costoro sono i più indicati per strategie di prevenzione sanitarie e sociali.

Il secondo approccio, ugualmente popolare e citato nella grande maggioranza degli articoli, è quello proposto da Fried (2001), in cui si definisce la fragilità come “una sindrome fisiologica caratterizzata dalla riduzione delle riserve funzionali e dalla

diminuita resistenza agli stressor, risultante dal declino cumulativo di sistemi fisiologici multipli che causano vulnerabilità e conseguenze negative”. In pratica Fried definisce fragile un individuo se soddisfa almeno tre dei seguenti criteri, basandosi sui dati del “Cardiovascular Health Study” (1999): perdita di peso non intenzionale (4,5 kg o il 5% del proprio peso corporeo); affaticamento (un soggetto si dichiara stanco almeno tre volte a settimana); riduzione della forza muscolare (misurata come forza di presa e stratificata in base al sesso e al BMI³); ridotta attività fisica (basata sul consumo di chilocalorie a settimana, tenendo conto del sesso); riduzione della velocità del cammino (valutata come tempo impiegato a percorrere 4.5 metri a piedi, con soglia specifica per sesso ed altezza).

Un altro paradigma, definito da Cavazza e Malvi (2014) “Bio-Psico Sociale”, è quello proposto da Gobbens (2010), che descrive la fragilità come “uno stato dinamico che colpisce un individuo che sperimenta perdite in uno o più domini funzionali (fisico, psichico, sociale), causate dall’influenza di più variabili che aumentano il rischio di risultati avversi per la salute”.

Gli autori sono arrivati a questa sintesi tramite una prima fase di ricerca bibliografica che ha prodotto 11 definizioni candidate, che poi sono state sottoposte a un gruppo di esperti, ognuno dei quali ha stilato una classifica assegnando 11 punti alla più adatta fino a 1 alla meno adatta. Cinque definizioni hanno avuto punteggi nettamente superiori alle altre e sono:

1. “La fragilità consiste nella perdita di risorse in diversi domini funzionali, che porta ad una diminuzione della capacità di conservazione e alla ridotta resistenza agli stressor”.
2. “La fragilità è una sindrome che coinvolge l’accumulo di disturbi e la perdita di capacità in più domini, che rendono l’individuo più vulnerabile ai cambiamenti ambientali”.
3. “La fragilità è una sindrome che consiste nella riduzione multisistemica della capacità di conservazione a seguito della quale le attività di una

³ Il Body Mass Index (BMI), in italiano “indice di massa corporea”, è un indicatore che, mettendo in relazione l’altezza e il peso, esprime una valutazione dello stato del peso forma di un individuo.

persona anziana possono essere gravemente compromesse da stress ambientali anche minimi, dando origine alla condizione di disabilità”.

4. “La fragilità è una sindrome biologica caratterizzata da una ridotta capacità di conservazione e di resistenza ai fattori di stress, derivante da declini progressivi su più sistemi fisiologici, che causano una maggiore vulnerabilità a esiti avversi”.
5. “La fragilità è una combinazione di cambiamenti biologici, fisiologici, sociali e ambientali che si verificano con l’avanzare dell’età e di un aumento della vulnerabilità ai cambiamenti ambientali e allo stress”.

Da queste definizioni emergono tre principi: 1. l’aspetto multidimensionale della fragilità, che coinvolge i diversi domini funzionali dell’individuo; 2. la continuità, intesa come una sindrome che può peggiorare o migliorare nel tempo; 3. l’importanza dell’omeostasi, l’equilibrio psico-fisico che permette all’individuo di sopportare gli stressor a cui è sottoposto.

Un'altra questione importante è il rapporto tra fragilità, disabilità e comorbidità⁴, per cui sta emergendo in ambito geriatrico un crescente consenso che queste siano entità cliniche distinte che sono causalmente associate (Fried, 2004).

Basandosi sui dati del Cardiovascular Health Study (1999), Fried cerca di stabilire come queste entità interagiscano tra loro, e arriva a diverse conclusioni: sia la fragilità che la comorbidità sono predittori della disabilità, condizionati l’uno dall’altro; la disabilità può peggiorare le condizioni di fragilità e comorbidità, così come quest’ultima può contribuire allo sviluppo negativo della fragilità; la presenza di disabilità o fragilità può contribuire allo sviluppo e alla progressione di malattie croniche, tramite il basso livello di attività fisica associato alle due sindromi o tramite altri deficit che intaccano diversi meccanismi biologici essenziali per il mantenimento dell’omeostasi.

Queste relazioni causali spiegano almeno in parte la compresenza di queste entità, e pongono in evidenza l’importanza di differenziare gli strumenti di prevenzione in base alla condizione del paziente.

⁴ Comorbidità - In ambito medico indica la coesistenza di più patologie diverse in uno stesso individuo.

1.2 I marcatori di fragilità (*outcomes*) in letteratura

Se sono numerosi gli studi che hanno tentato di dare una definizione vincente di fragilità, altrettanto numerosi sono quelli che indagano sulle conseguenze di essa. Nella tabella 1.1 sono elencati i diversi outcomes trovati in letteratura, che principalmente provengono da tre fonti.

La prima è la bibliografia della tesi di Margherita Silan, “Identificazione dei soggetti fragili nell’Ulss 15: una nuova proposta basata sugli ordinamenti parziali” (2015) di cui questo elaborato è il naturale proseguimento.

La seconda fonte è Sternberg (2011), in cui è stata redatta una revisione sistematica della definizione clinica, degli strumenti di screening e soprattutto delle conseguenze della fragilità. La ricerca è stata svolta tramite il database MEDLINE, limitandola ad articoli pubblicati tra il 1997 e il 2009 e che si riferiscono a persone di età superiore ai 65 anni. Dei 4.334 articoli candidati, solo 22 hanno superato i criteri di inclusione. Gli outcomes più comuni sono risultati il decesso (13,76%), la disabilità (7,41%) e l’istituzionalizzazione (6,35%); un gruppo di outcomes riferito alla macro categoria del declino fisico, comprendente la disabilità, il declino della performance fisica e il declino funzionale, è presente in 11 articoli (il 65% del totale).

La terza e ultima fonte è un'altra revisione sistematica (Wallace, 2014), che però si concentra sul rischio di ricovero urgente, con lo scopo di facilitare interventi mirati di prevenzione sui pazienti. Dei 18.983 articoli, solo 27 hanno superato i criteri di inclusione, mentre la popolazione presa in esame è quella dai 18 anni in su. Gli articoli si differenziano tra quelli che hanno utilizzato un questionario, quelli che hanno utilizzato dati amministrativi, quelli che hanno fatto uno studio retrospettivo o che hanno adottato un metodo prospettico.

Tabella 1.1 Outcomes della fragilità e relative fonti bibliografiche.

Outcome	Fonti della letteratura
Decesso	Paw (1999), Fried (2001), Saliba (2001), Mitnitski (2002), Klein (2005), Mazzaglia (2007), Ravaglia (2008), Rothman (2008), Avila-Funes (2009), Kamaruzzaman (2010), Falasca (2011), Cavazza (2014)
Disabilità	Paw (1999), Woods (2005), Carriere (2005), Ravaglia (2008), Rothman (2008), Avila-Funes (2009)
Declino della performance fisica	Paw (1999), Saliba (2001), Puts (2005)
Istituzionalizzazione	Brody (1997), Brody (2002), Jones (2004), Rockwood (2006), Rothman (2008), Luppà (2009), Kamaruzzaman (2010)
Assistenza Domiciliare	Brody (1997)
Caduta/incidente	Fried (2001), Rothman (2008), De Vries (2013), Tom (2013)
Frattura	Woods (2005), Ravaglia (2008)
Ospedalizzazione	Shelton (2000), Fried (2001), Landi (2004), Damush (2004), Walker (2005), Mazzaglia (2007), Ravaglia (2008), Avila-Funes (2009), Crane (2010), Kamaruzzaman (2010)
Peggioramento della mobilità	Fried (2001)
Peggioramento della disabilità	Fried (2001), Puts (2005), Ravaglia (2008)
Ricovero d'urgenza	Shelton (2000), Damush (2004), Walker (2005), Inouye (2008), Crane (2010), Falasca (2011), Lopez-aguilà (2011), Billings (2013), Hippisley-Cox (2013), Cavazza (2014)
Demenza	Buchman (2007), Avila-Funes (2009), Song (2011), Avila-Funes (2012), Solfrizzi (2013), Gray (2013)
Comorbidità	Tammemagi (2003)
Giorni di ricovero totali	Rockwood (2005), Makary (2010)

1.3 Le variabili esplicative della fragilità in letteratura

Dalle fonti precedentemente descritte, sono state ricavate anche le variabili esplicative utilizzate, che sono elencate con i rispettivi outcomes nella tabella 1.2.

È necessario precisare che non sempre le variabili elencate nella tabella entrano nell'analisi direttamente come predittori, ma a volte i valori che assumono determinano a loro volta il valore complessivo di un indice, che viene a sua volta usato come predittore per verificare il suo impatto su uno degli outcomes precedentemente elencati.

Per sintetizzare, le numerose variabili esplicative possono essere riunite nelle seguenti categorie:

- Variabili socio-demografiche (età, sesso, nazionalità, reddito, condizione familiare...);
- Stili di vita (attività fisica, fumo, alcool...);
- Variabili relative alla storia del paziente (ospedalizzazioni, diagnosi, ricoveri urgenti, numero giorni di ricoveri...);
- Variabili relative a malattie croniche (diabete, anemia...);
- Variabili relative all'uso dei farmaci del paziente;
- Variabili relative alle problematiche fisiche e di mobilità dell'individuo (disabilità, scarsa energia, perdita di peso involontaria...);
- Variabili relative allo stato mentale dell'individuo (depressione, demenza, uso di servizi psichiatrici...).

Un altro aspetto da considerare è che quelle che sono variabili esplicative in alcuni studi, sono invece outcomes in altri: tra queste per esempio ci sono le variabili riferite a ospedalizzazioni e ricoveri urgenti precedenti, disabilità, comorbidità, demenza e frattura.

Tabella 1.2: Variabili esplicative degli outcomes della fragilità nella letteratura.

Variabili Esplicative	Outcomes delle analisi in cui sono presenti le variabili esplicative
Età	Decesso, disabilità, assistenza domiciliare, istituzionalizzazione, caduta, frattura, ospedalizzazione, ricovero d'urgenza, demenza, comorbidità, declino della performance fisica, peggioramento della mobilità, peggioramento della disabilità
Sesso	Decesso, disabilità, assistenza domiciliare, caduta, frattura, ospedalizzazione, ricovero d'urgenza, demenza, comorbidità, istituzionalizzazione, peggioramento della mobilità, peggioramento della disabilità
Nazionalità/etnia	Decesso, caduta, ospedalizzazione, peggioramento della mobilità, peggioramento della disabilità, ricovero d'urgenza, demenza
Titolo di studio	Istituzionalizzazione, decesso, disabilità, caduta, frattura, ospedalizzazione, peggioramento della mobilità, peggioramento della disabilità, ricovero d'urgenza, demenza
Stato civile	Istituzionalizzazione, ospedalizzazione, ricovero d'urgenza
Reddito	Decesso, disabilità, caduta, ospedalizzazione, ricovero d'urgenza, demenza, comorbidità, istituzionalizzazione, peggioramento della mobilità, peggioramento della disabilità, ricovero d'urgenza, demenza
Indice di deprivazione	Decesso, ricovero d'urgenza
Rete sociale	Istituzionalizzazione
Proprietà di una casa	Istituzionalizzazione, decesso, ospedalizzazione, ricovero d'urgenza
Anziano che vive solo	Decesso, istituzionalizzazione, disabilità, caduta, frattura, ospedalizzazione, peggioramento della mobilità, peggioramento della disabilità, ricovero d'urgenza
Charlson Index ⁵	Decesso, ricovero urgenza, demenza
Malattie croniche	Decesso, disabilità, assistenza domiciliare, caduta, frattura, ospedalizzazione, ricovero d'urgenza, demenza, peggioramento della mobilità, peggioramento della disabilità
Auto-valutazione dello stato di salute	Istituzionalizzazione, decesso, disabilità, declino della performance fisica, assistenza domiciliare, caduta,

⁵ Il Charlson Comorbidity Index è un indice che assegna un peso alla comorbidità (considerando le diagnosi assegnate nei ricoveri ospedalieri) sulla base del rischio relativo di morte delle specifiche patologie incluse.

Variabili Esplicative	Outcomes delle analisi in cui sono presenti le variabili esplicative
	frattura, ospedalizzazione, peggioramento della mobilità, ricovero d'urgenza, demenza
Attività fisica	Istituzionalizzazione, decesso, disabilità, declino della performance fisica, caduta, frattura, ospedalizzazione, peggioramento della mobilità, peggioramento della disabilità, ricovero d'urgenza, demenza
Fumatore/ex fumatore	Decesso, disabilità, declino della performance fisica, caduta, frattura, ospedalizzazione, peggioramento della disabilità, ricovero d'urgenza, demenza, comorbidità
Abuso di alcol	Decesso, disabilità, caduta, ospedalizzazione, ricovero d'urgenza, demenza, comorbidità
Cadute nei precedenti 12 mesi	Caduta
Storia di Ricoveri urgenti	Decesso, assistenza domiciliare, ospedalizzazione, ricovero d'urgenza
Parziale/totale invalidità	Decesso
Disabilità	Decesso, demenza
ADL / IADL ⁶	Assistenza domiciliare, ospedalizzazione, demenza
Accesso ai servizi di salute mentale/disturbi psichiatrici	Decesso, disabilità, ricovero d'urgenza, declino della performance fisica
Mini-Mental score	Decesso, disabilità, declino della performance fisica, caduta, frattura, ospedalizzazione, peggioramento della mobilità, peggioramento della disabilità
Assistenza domiciliare	Decesso, ospedalizzazione, ricovero d'urgenza, istituzionalizzazione
Poliprescrizione di farmaci	Decesso, disabilità, fratture, ospedalizzazione, ricovero d'urgenza, istituzionalizzazione, peggioramento della disabilità
Storia di ospedalizzazioni precedenti	Decesso, assistenza domiciliare, ospedalizzazione, ricovero d'urgenza, istituzionalizzazione
Numero di ammissioni totale all'ospedale	Assistenza domiciliare
Accesso al pronto soccorso tramite 118	Decesso, ricovero d'urgenza
Accessi al pronto soccorso	Decesso, ricovero d'urgenza
Numero di diagnosi	Ospedalizzazione
3 o + ricoveri con differenti diagnosi	Decesso, ricovero d'urgenza
Ricovero per malattia mentale	Decesso
CAD: Coronary Artery Disease	Decesso, ospedalizzazione, ricovero d'urgenza, demenza
Infarto miocardio	Decesso, disabilità, caduta, ospedalizzazione, ricovero d'urgenza, declino della performance fisica,

⁶ Le ADL (Activities of Daily Living) sono le attività fondamentali per la cura personale. Le IADL (Instrumental Activities of Daily Living) si riferiscono al livello di indipendenza nelle abilità strumentali della vita quotidiana (fare la spesa, utilizzo dei trasporti, etc.).

Variabili Esplicative	Outcomes delle analisi in cui sono presenti le variabili esplicative
	peggioremento della mobilità, peggioremento della disabilit�
Angina	Decesso, disabilit�, caduta, ospedalizzazione, declino della performance fisica, peggioremento della mobilit�, peggioremento della disabilit�
PVD: Peripheral Vascular Disease	Decesso, caduta, disabilit�, ricovero d'urgenza, peggioremento della mobilit�, peggioremento della disabilit�
VHD: Valvular Heart Disease	Decesso, ricovero d'urgenza
CHF: Congestive Heart Failure	Decesso, disabilit�, caduta, ospedalizzazione, ricovero d'urgenza, declino della performance fisica, peggioremento della mobilit�, peggioremento della disabilit�, demenza
VHF: Vascular Heart Disease	Declino della performance fisica
Cancro	Decesso, disabilit�, caduta, ospedalizzazione, ricovero d'urgenza, declino della performance fisica, peggioremento della mobilit�, peggioremento della disabilit�
Asma	Ricovero d'urgenza
Fibrillazione atriale	Ricovero d'urgenza
Epilessia	Ricovero d'urgenza
Bronchite	Ricovero d'urgenza
Bronchiectasia ⁷	Ricovero d'urgenza
Enfisema	Ricovero d'urgenza
COPD: Chronic Obstructive Pulmonary Disease	Decesso, disabilit�, caduta, ospedalizzazione, ricovero d'urgenza, peggioremento della mobilit�, peggioremento della disabilit�
Artrosi	Decesso, disabilit�, caduta, ospedalizzazione, ricovero d'urgenza, istituzionalizzazione, declino della performance fisica, peggioremento della mobilit�, peggioremento della disabilit�
Diabete	Decesso, disabilit�, caduta, ospedalizzazione, ricovero d'urgenza, istituzionalizzazione, declino della performance fisica, peggioremento della mobilit�, peggioremento della disabilit�, demenza
Ipertensione	Decesso, disabilit�, caduta, ospedalizzazione, ricovero d'urgenza, istituzionalizzazione, peggioremento della mobilit�, peggioremento della disabilit�, demenza
Morbo di Parkinson	Decesso, ricovero d'urgenza
Epatite	Decesso, ricovero d'urgenza
Trombosi venosa profonda	Ricovero d'urgenza
Malattia infiammatoria pelvica	Ricovero d'urgenza
Pielonefrite ⁸	Ricovero d'urgenza

⁷ Bronchiectasia - Malattia in cui si riscontra una dilatazione irreversibile di una porzione dell'albero bronchiale dei polmoni.

Variabili Esplicative	Outcomes delle analisi in cui sono presenti le variabili esplicative
Gastroenterite	Ricovero d'urgenza
Cirrosi	Decesso, ricovero d'urgenza
Demenza	Decesso, disabilità, ospedalizzazione, ricovero d'urgenza, istituzionalizzazione, declino della performance fisica
Insufficienza renale	Decesso, ospedalizzazione, ricovero d'urgenza,
Morbo di Alzheimer	Ricovero d'urgenza
Ictus	Decesso, disabilità, ospedalizzazione, ricovero d'urgenza, istituzionalizzazione, demenza
Anemia	Ricovero d'urgenza
Ipercolesterolemia	Decesso, ricovero d'urgenza, demenza
Incontinenza	Istituzionalizzazione, decesso, declino della performance fisica, peggioramento della disabilità
Cadute/Frattura femore	Decesso, disabilità, ricovero d'urgenza, declino della performance fisica
Sintomi depressivi/depressione	Decesso, disabilità, caduta, frattura, ospedalizzazione, ricovero d'urgenza, demenza, istituzionalizzazione, declino della performance fisica, peggioramento della mobilità, peggioramento della disabilità, peggioramento della disabilità
Essere esausto	Decesso, disabilità, caduta, ospedalizzazione, peggioramento della mobilità, peggioramento della disabilità, demenza
Perdita involontaria di peso	Decesso, disabilità, declino della performance fisica, caduta, ospedalizzazione, peggioramento della mobilità, peggioramento della disabilità, demenza
Lentezza nei movimenti	Decesso, disabilità, caduta, ospedalizzazione, peggioramento della mobilità, peggioramento della disabilità, demenza
Forza nella presa	Decesso, disabilità, caduta, ospedalizzazione, peggioramento della mobilità, peggioramento della disabilità, demenza
Abilità fisica (faccende domestiche, fare le scale, camminare, uscire, lavarsi, artrosi, cadute)	Decesso, ospedalizzazione
Difficoltà visive (cataratta, glaucoma)	Decesso, ospedalizzazione, disabilità, declino della performance fisica, peggioramento della disabilità, ricovero d'urgenza
Difficoltà uditive	Decesso, disabilità, declino della performance fisica, ospedalizzazione, peggioramento della disabilità, ricovero d'urgenza
Problemi di memoria	Ospedalizzazione, ricovero d'urgenza
Sintomi e malattie cardiache	Decesso, ospedalizzazione

⁸ Pielonefrite - Infiammazione del parenchima, dei calici e della pelvi renale.

Variabili Esplicative	Outcomes delle analisi in cui sono presenti le variabili esplicative
(angina, infarto miocardio, dolori al petto)	
Sintomi e malattie respiratorie (asma, bronchite/enfisema, fiato corto, tosse)	Decesso, ospedalizzazione, istituzionalizzazione
Problemi psicologici (ansia, depressione, memoria)	Decesso, ospedalizzazione
Problemi fisiologici (ipertensivo, rapporto anca/vita, BMI, ipotensione posturale, tachicardia sinusale)	Decesso
Disturbi funzionali	Istituzionalizzazione
Disturbi cognitivi	Istituzionalizzazione, decesso, caduta, demenza
Deficit sensoriali	Decesso, disabilità, frattura, ospedalizzazione
Malattie polmonari	Decesso, ricovero d'urgenza
Malattie del fegato	Ospedalizzazione, ricovero d'urgenza
Malattie cerebrovascolari	Decesso, ricovero d'urgenza
Patologie vascolari	Decesso
Malattie gastroenteriche	Decesso, ricovero d'urgenza
Mobilità	Disabilità
Equilibrio	Disabilità
Paura di cadere	Disabilità
Scarsa energia	Decesso, disabilità, declino della performance fisica
BMI (Body Mass Index)	Decesso, disabilità, declino della performance fisica, ospedalizzazione, peggioramento della disabilità, ricovero d'urgenza
Circonferenza del polpaccio	Decesso, disabilità, frattura, ospedalizzazione, peggioramento della disabilità
Umore e cognizione	Decesso, disabilità, frattura, ospedalizzazione, peggioramento della disabilità
Pain score	Ospedalizzazione, ricovero d'urgenza
Energy score	Ospedalizzazione, ricovero d'urgenza
Usare il bagno	Decesso
Fiato corto	Declino della performance fisica, peggioramento della disabilità
Difficoltà nel cucinare	Decesso
Problemi alla pelle	Decesso
Difficoltà ad alzarsi	Decesso
Tremore	Decesso
Lee score (indice di rischio cardiaco)	Giorni di ricovero totali
Asa score (classificazione stato fisico)	Giorni di ricovero totali
Eagle score (indice di rischio di morte in operazioni al cuore)	Giorni di ricovero totali

Variabili Esplicative	Outcomes delle analisi in cui sono presenti le variabili esplicative
Numero di giorni in ospedale	Assistenza domiciliare, ospedalizzazione
Numero di giorni nell'ultima ospedalizzazione	Ricovero d'urgenza
Numero totale di giorni ricoverato d'urgenza	Ricovero d'urgenza
Giorni totali di ospedale	Ricovero d'urgenza
Costo dei ricoveri precedenti	Ricovero d'urgenza
Numero di diversi farmaci negli ultimi 12 mesi	Assistenza domiciliare
Uso di antianemici	Ospedalizzazione, ricovero d'urgenza
Uso di statine	Ricovero d'urgenza
Uso di anti infiammatori non steroideali	Ricovero d'urgenza
Uso di corticosteroidi	Ricovero d'urgenza
Uso di anticoagulanti	Ricovero d'urgenza
Uso di antidepressivi	Ricovero d'urgenza
Uso di farmaci per problemi gastrointestinali	Ricovero d'urgenza
Uso di diuretici	Ricovero d'urgenza
Uso di anti-ipertensivi	Ricovero d'urgenza, demenza
Uso di antiaggreganti	Ricovero d'urgenza
Uso di farmaci per problemi respiratori	Ricovero d'urgenza
Uso di farmaci ipnotici e ansiolitici	Ricovero d'urgenza
Uso di antipsicotici	Ricovero d'urgenza
Uso di analgesici	Ricovero d'urgenza
Uso di antiparkinsoniani	Ricovero d'urgenza
Uso di antidiabetici	Ricovero d'urgenza, demenza
Uso di antiosteoporotici inclusi estrogeni	Caduta, ricovero d'urgenza

1.3 Aspetti salienti nello studio della fragilità

Prima di proseguire con la descrizione dei dati, è utile sottolineare la complessità della questione, in riferimento al tipo di analisi che si sta svolgendo. Innanzitutto la prima questione da affrontare è la gran mole di variabili esplicative presenti in letteratura e di conseguenza le varie tipologie di raccolta dati: alcuni articoli hanno usato dati amministrativi, altri hanno circoscritto l'analisi a pazienti di case di cura o a pazienti di un ospedale, altri ancora si sono basati su indagini ad hoc, con questionari appositi per lo studio della fragilità. Il presente elaborato si vuole porre

all'interno di un contesto di sanità pubblica, dove i tempi e i costi di un'indagine ad hoc sono il più delle volte proibitivi e dove allo stesso tempo sono disponibili molte basi dati di natura amministrativa. Pertanto, pur non potendo disporre di informazioni ricavabili da questionari o scale di valutazione personalizzate per utente, la mole di dati presenti nei flussi sanitari permette di studiare gran parte delle variabili presenti nella tabella 1.2.

Un secondo problema, già accennato nell'introduzione, è il carattere dinamico della fragilità, un concetto poco predisposto ad essere misurato tramite l'uso di soglie in analisi classiche di regressione. Un metodo probabilmente più adeguato è l'analisi di sopravvivenza, implementata per esempio in un lavoro di Rockwood nel 2004 per studiare la prevalenza, le caratteristiche e gli outcomes della fragilità. Un'analisi di questo tipo è difficilmente realizzabile, basti pensare alla necessità di programmare degli interventi di prevenzione, o i posti letto previsti in ospedale: poiché le basi di dati si rendono disponibili di anno in anno, per raggiungere questi obiettivi è necessario riferirsi agli eventi fino a una certa data, e basarsi su questi per predire le risorse necessarie da investire nell'anno successivo.

Infine, un altro elemento di complessità è dato dal fatto che gli stessi outcomes sono variabili esplicative di altri outcomes, ad eccezione del decesso, per ovvie ragioni. Quindi è evidente che le altre variabili esplicative avranno sia un effetto diretto su un dato outcome, sia un effetto indiretto in virtù della presenza di altri outcomes nelle altre variabili esplicative. Inoltre è possibile che un outcome, come per esempio la disabilità, abbia una relazione causale con la comorbidità in entrambe le direzioni, e che questi stessi effetti causali siano diretti o indiretti.

Capitolo 2

I dati

Nel seguente capitolo vengono presentati e descritte le fonti dei dati utilizzati nell'analisi, ovvero alcuni tra i flussi informativi utilizzati dalle Ulss del Veneto in ambito amministrativo: archivio dell'Anagrafe sanitaria regionale, schede di Dimissione Ospedaliera, flusso del Pronto Soccorso Aziendale, flusso della Psichiatria Territoriale, flusso dell'Assistenza Domiciliare Integrata, archivio delle Esenzioni del ticket, flusso della Farmaceutica Territoriale. I flussi in questione si riferiscono solo al territorio dell'Ulss 15 "Alta Padovana" che ha sede a Camposampiero.

Pur essendo nate per fini amministrativi, queste fonti sono utilizzabili anche in ambito epidemiologico grazie alla mole di informazione deducibile dai dati raccolti: naturalmente qualsiasi studio basato su dati amministrativi ha il limite di non avere a disposizione delle informazioni raccolte *ad hoc*, ma la quantità e la qualità informativa dei flussi sanitari garantiscono la possibilità di sopperire a questo tipo di mancanza.

Per ognuno dei flussi elencati vengono presentati le informazioni più importanti in riferimento alle variabili esplicative e agli outcomes descritti nel capitolo precedente, oltre al numero di record presenti per l'analisi, considerando che sono stati presi in esame solo le persone anziane, con una soglia fissata a 65 anni.

Viene poi descritto il record linkage delle varie fonti informative, che ha portato alla costruzione del dataset utilizzato nelle analisi successive.

Infine vengono esplicitate le scelte relative agli outcomes e alle esplicative che verranno studiate nelle analisi successive, soffermandosi sui motivi di tali scelte e sulle approssimazioni applicate, dovute ai limiti dei flussi informativi.

2.1 I flussi informativi

2.1.1 Anagrafe Sanitaria Regionale

L'archivio dell'anagrafe sanitaria contiene informazioni socio-demografiche, quali età, sesso, residenza, cittadinanza e le informazioni relative al medico di base. Ogni individuo ha un proprio e unico record, un codice identificativo grazie al quale è possibile collegare questa fonte alle altre, e una variabile che indica se il soggetto è deceduto alla fine dell'anno. In totale sono presenti 49132 record per l'anagrafe del 2014, anno in cui viene considerato il decesso ai fini dell'analisi.

2.1.2 Schede di Dimissione Ospedaliera

Il flusso informativo delle Schede di Dimissione Ospedaliera (SDO) è lo strumento di raccolta delle informazioni relative a tutti gli episodi di ricovero erogati nelle strutture ospedaliere pubbliche e private presenti in tutto il territorio nazionale.

In origine le SDO nascono per fini amministrativi, ma grazie alla ricchezza e alla profondità di informazioni anche di carattere clinico, sono diventate uno strumento importante per un'ampia gamma di analisi ed elaborazioni, che comprendono il supporto dell'attività di programmazione sanitaria, il monitoraggio dell'erogazione dell'assistenza ospedaliera e dei Livelli Essenziali di Assistenza, nonché altre analisi di carattere clinico-epidemiologico (Ministero della Salute, 2015).

Le SDO forniscono informazione sulle caratteristiche anagrafiche del paziente (età, sesso, residenza, stato civile), caratteristiche del ricovero (come regime di ricovero, data di entrata e di dimissione, modalità di dimissione, classe di priorità del ricovero) e caratteristiche cliniche (ad esempio diagnosi principale, diagnosi concomitanti, procedure diagnostiche o terapeutiche). Le diagnosi sono codificate tramite la Classificazione delle malattie, dei traumatismi, degli interventi chirurgici e delle procedure diagnostiche e terapeutiche (ICD-9-CM). Ogni record si riferisce a un evento di ricovero, per un totale di 32293 record per gli anni 2012 e 2013, anni in cui vengono rilevate le variabili esplicative nell'analisi.

2.1.3 Pronto Soccorso

Il flusso del Pronto soccorso è lo strumento di monitoraggio delle prestazioni erogate nell'ambito dell'assistenza sanitaria in emergenza-urgenza da parte del pronto Soccorso e del 118. Il flusso si compone di una parte di informazioni di tipo anagrafico (sesso, età, residenza) e di una più cospicua parte di tipo sanitario, che comprende in particolare la modalità di arrivo (chiamata 118, autonomo, polizia...), il problema principale rilevato, le diagnosi secondarie, l'esito della prestazione (decesso, trasferimento in ospedale...). Anche in questo caso le diagnosi sono codificate tramite l'ICD-9-CM e ogni record si riferisce a un evento di ricovero, per un totale di 39408 record per gli anni 2012 e 2013.

2.1.4 Assistenza Psichiatrica Territoriale

Con il flusso informativo dell'Assistenza Psichiatrica Territoriale (APT) vengono rilevate, oltre alle prestazioni erogate, le caratteristiche anagrafiche e cliniche sia degli utenti dei Dipartimenti di Salute Mentale (DSM) delle Aziende Ulss del Veneto, sia degli utenti dei Servizi per i Disturbi del Comportamento Alimentare (DCA) con diagnosi psichiatrica.

Il flusso informativo è costituito da due strutture di dati: la prima contiene informazioni di carattere anagrafico e clinico, come età, residenza, stato civile, diagnosi principali, diagnosi concomitanti, data chiusura cartella clinica e motivo della conclusione; il secondo riguarda la prestazione erogata, ed è composta per esempio dal codice dell'azienda Ulss, dal codice erogatore, data della prestazione e modalità di erogazione (urgente o programmato).

Le diagnosi presenti in questo flusso sono codificati tramite l'ICD-10, che è la versione più recente e dettagliata dell'ICD-9-CM. Ogni record si riferisce a una prestazione, per un totale di 3679 eventi registrati negli anni 2012 e 2013.

2.1.5 Assistenza Domiciliare Integrata

Il flusso dell'Assistenza Domiciliare Integrata è un servizio che “garantisce alle persone non autosufficienti e in condizioni di fragilità, con patologie in atto o esiti

delle stesse, percorsi assistenziali nel proprio domicilio denominati “cure domiciliari” consistenti in un insieme organizzato di trattamenti medici, infermieristici e riabilitativi, necessari per stabilizzare il quadro clinico, limitare il declino funzionale e migliorare la qualità della vita” (Ministero della Salute, 2013). Il flusso è suddiviso in due gruppi d’informazioni ed i tracciati previsti sono due: il primo contiene le informazioni associate all’evento di presa in carico (dati anagrafici del paziente, ASL erogante, soggetto richiedente la presa in carico, valutazione iniziale, etc.); il secondo contiene le informazioni associate agli eventi erogazione (dati riferiti agli accessi e i dati di prestazione), sospensione, rivalutazione (autonomia e bisogni assistenziali) e conclusione. Anche in questo caso ogni record è riferito alla prestazione, per un totale di 4346 record registrati nel biennio 2012-2013.

2.1.6 Esenzioni ticket

Il flusso delle esenzioni ticket riporta i codici delle esenzioni degli individui avente tale diritto. Le diverse tipologie di esenzione sono:

- Esenzioni per patologie croniche
- Esenzioni per malattie rare
- Esenzioni per reddito
- Esenzioni per invalidità
- Esenzioni per diagnosi precoci di tumore
- Esenzioni per gravidanza

Ogni record si riferisce a un’esenzione, e poiché un individuo può avere diverse esenzioni, si ha un totale di 120430 record nell’anno 2013.

2.1.7 Farmaceutica Territoriale

Il flusso della farmaceutica territoriale è finalizzato alla raccolta dei dati sui medicinali dispensati tramite le strutture sanitarie agli assistiti per la somministrazione al proprio domicilio. Oggetto di rilevazione del flusso sono le

informazioni relative alle prestazioni farmaceutiche, destinate al consumo al domicilio, erogate:

- alla dimissione da ricovero o a seguito di visita specialistica, limitatamente al primo ciclo terapeutico completo;
- ai pazienti cronici soggetti a piani terapeutici o presi in carico;
- ai pazienti in assistenza domiciliare, residenziale o semiresidenziale;
- da parte delle farmacie convenzionate, pubbliche o private, per conto delle Aziende sanitarie locali.

I farmaci sono codificati tramite il sistema di classificazione “Anatomical Therapeutic Chemical” (ATC), e sono affiancati da altre informazioni, come quelle anagrafiche o quelle di prestazione erogata (data di erogazione, data prescrizione, tipo di farmaco, quantità).

Il dataset contiene 1375977 record, ed ognuno di questi rappresenta un ordine singolo di un determinato farmaco avvenuto nel 2013.

2.2 Il record linkage delle fonti

Per poter creare il dataset finale per le successive analisi, sono state unite le fonti citate nel precedente paragrafo, sfruttando il codice identificativo univoco per ogni individuo.

I dati a disposizione vanno dal 2012 al 2015: inizialmente è stato creato un dataset contenente gli outcomes unendo i vari flussi relativi all’anno 2014 tramite il codice identificativo della persona. Poiché l’obiettivo della tesi è l’identificazione di soggetti fragili anziani, si è presa come soglia d’età i 65 anni compiuti nel 2014. Per la costruzione di alcune variabili risposta dicotomiche, come demenza, disabilità e frattura, sono state integrate informazioni provenienti da diversi flussi, utilizzando l’operatore logico unione. Inoltre la maggior parte delle fonti non prevede un record per persona, ma per prestazione: pertanto nelle SDO ogni riga rappresenta un ricovero, nel flusso farmaceutico ogni riga rappresenta un acquisto di un dato farmaco e analogamente per gli altri flussi. Pertanto dopo aver ordinato i vari flussi per codice identificativo, gli stessi sono stati aggiornati in modo da

avere una riga per ogni utente, senza perdere le informazioni essenziali. Il dataset è risultato essere composto da 49142 unità, e 5 variabili (gli outcomes).

Allo stesso modo è stato creato il dataset delle variabili esplicative, relativo agli anni 2012 e 2013, utilizzando diverse fonti per le variabili esplicative finali (tutte dicotomiche o categoriali) e ottenendo una riga per ogni soggetto, comprensiva di tutte le informazioni necessarie. Infine si è proceduto con un record linkage tra il dataset degli outcomes e quello delle esplicative, con la logica del left join: il dataset finale ha tante righe quante quelle del dataset di base (quello degli outcomes) e la somma delle colonne dei due dataset, unite tramite il codice identificativo dell'utente. Questo primo dataset è il *“training set”* per le analisi statistiche presentate nel capitolo 3, mentre il *“test set”* è stato costruito analogamente ma riferendosi agli anni 2013-2014 per le variabili esplicative, e all'anno 2015 per le variabili risposta.

2.3 Gli outcomes disponibili nei dati

Nei paragrafi seguenti, in riferimento alla tabella 1.1, viene descritto il processo decisionale che ha portato a scegliere o scartare determinati outcomes, e come questi vengano costruiti in relazione alle fonti dei dati disponibili. Questa è una fase importante e delicata, poiché condiziona le analisi statistiche che verranno presentate nel terzo capitolo.

2.3.1 Decesso

Il decesso è il marcatore di fragilità maggiormente analizzato in letteratura, in primo luogo perché rappresenta l'evento più grave tra le conseguenze della fragilità, e in secondo luogo perché è un evento facile da rilevare.

In questa analisi, il decesso è stato rilevato per gli individui presenti nell'anagrafe relativa al territorio di competenza all'ultimo giorno del 2014; tra questi sono stati esclusi coloro i quali non erano presenti nell'anagrafe relativa al 2012, poiché il 2012 e il 2013 sono gli anni in cui vengono rilevate le variabili esplicative: in questo modo si ha la certezza di avere individui osservati per tutta la durata del periodo di rilevazione.

2.3.2 Ricovero urgente

Il ricovero urgente è, dopo il decesso, l'outcome più importante della fragilità, anche per la sua influenza nella gestione delle risorse sanitarie. Infatti il numero di urgenze, i posti letto, i giorni di ricovero e le conseguenti spese sanitarie per paziente sono temi di fondamentale rilevanza nell'ambiente sanitario.

In letteratura, il ricovero d'urgenza come marcatore della fragilità è stato studiato soprattutto in tre modi differenti: il primo ricovero d'urgenza, il secondo ricovero d'urgenza e infine il generico ricovero d'urgenza, che è la codifica scelta nelle successive analisi: si è preferito studiare il generico ricovero urgente principalmente per utilizzare il maggior numero di individui possibili, senza limitarsi o solo a quelli che non hanno mai avuto un ricovero urgente, o solo a quelli che ne hanno avuto almeno uno.

Per rilevare questo outcome è stata utilizzata una variabile presente nelle SDO, ovvero la tipologia di ricovero, che assume valori da 1 a 4, con questa codifica:

- 1: Ricovero programmato non urgente;
- 2: Ricovero urgente;
- 3: Ricovero per trattamento sanitario obbligatorio TSO;
- 4: Ricovero programmato con pre-ospedalizzazione.

Pertanto l'outcome "Ricovero Urgente" è assente (assume valore 0) quando tipologia di ricovero assume valore 1 o 4, mentre è presente (assume valore 1) nei casi opposti.

2.3.3 Disabilità

La disabilità è definita da Fried (2004) come la difficoltà o la dipendenza nello svolgimento delle attività essenziali della vita quotidiana e della cura di sé. La disabilità può essere anche inclusa nella dimensione sociale (l'incapacità di comunicare o semplicemente di uscire di casa per motivi fisici o psichici), o in quella fisica (problemi motori, difficoltà nel fare pochi passi). Proprio per questa sua dimensione quotidiana, la disabilità è rilevata quasi sempre da questionari o da

scaie di valutazione, composte da items che indagano per esempio le difficoltà nel camminare per due o tre passi, nel prepararsi da mangiare, nell'utilizzare il bagno o nel vestirsi da soli. Poiché non sono a disposizione queste tipologie di dati, è stato necessario approssimare la disabilità utilizzando due fonti amministrative: le esenzioni e l'assistenza domiciliare. Nel primo caso è stata considerata l'esenzione per invalidità, che è assegnata a coloro che hanno difficoltà a svolgere le funzioni tipiche della vita quotidiana o di relazione a causa di una menomazione o di un deficit fisico, psichico o intellettuale, della vista o dell'udito (Ministero della Salute, 2013); nel secondo caso è sufficiente aver ricevuto almeno un giorno di assistenza domiciliare. Pertanto un individuo è stato classificato disabile avendo avuto l'esenzione di invalidità e/o l'assistenza domiciliare nel 2014.

Tra i marcatori indicati in tabella 1.1, quattro di questi sono stati approssimati da un unico outcome, la disabilità, e sono:

- l'assistenza domiciliare;
- il declino della performance fisica;
- il peggioramento della disabilità;
- il peggioramento della mobilità.

A differenza del decesso e del ricovero d'urgenza, la disabilità è studiata solo per gli individui che nei due anni precedenti non hanno avuto assistenza domiciliare o l'esenzione per invalidità, in modo che i fattori di rischio emergenti dall'analisi siano temporalmente collocati prima dell'evento disabilità, nei limiti dell'approssimazione utilizzata.

2.3.4 Demenza

La demenza è un disturbo neurocognitivo caratterizzato dal peggioramento della memoria, del linguaggio e di altre funzioni cognitive, che col tempo riduce al minimo la capacità dell'individuo di svolgere le attività fondamentali del vivere quotidiano. La demenza è tipicamente la maggior causa di disabilità negli anziani e si stima che ne soffrano circa 35 milioni di persone nel mondo, con una prevalenza tra gli ultra sessantenni tra il 5% e il 7% (Kojima, 2016).

Per costruire questo outcome sono state utilizzate cinque fonti:

- Esenzioni: l'esenzione per demenza o per il morbo di Alzheimer;
- SDO: la diagnosi di demenza o Alzheimer, tenendo conto di diagnosi principale e secondarie;
- Pronto Soccorso: la diagnosi di demenza o Alzheimer, tramite la diagnosi principale e le diagnosi secondarie;
- Psichiatria territoriale: la diagnosi di demenza o Alzheimer;
- Farmaceutica territoriale: l'uso di farmaci contro la demenza.

Pertanto l'individuo è valutato come malato di demenza se almeno una di queste variabili risulta presente nelle rispettive fonti di dati. Come per la disabilità, anche in questo caso sono stati esclusi coloro che, tramite lo stesso criterio, risultavano malati di demenza nei due anni precedenti al 2014, anno di rilevazione degli outcomes.

2.3.5 Frattura

Le fratture sono diventate più frequenti con il progressivo invecchiamento della popolazione, tanto che si è stimato che il 50% delle donne e il 20% degli uomini sopra i 50 anni abbiano avuto una frattura nel restante tempo di vita (Van den Bergh, 2011). Le fratture possono avere un impatto negativo non solo fisicamente ma anche mentalmente sulle persone anziane, mentre per le spese pubbliche rappresentano una fonte importante di costi. Un esempio è la frattura dell'anca, che è spesso seguita da mobilità ridotta, disabilità funzionale, aumento della dipendenza e istituzionalizzazione. In letteratura esistono vari outcomes relativi agli incidenti, tra cui la frattura o più specificatamente la frattura del femore come rischio per gli anziani fragili.

Le uniche fonti amministrative a disposizione per rilevare un incidente sono le diagnosi delle SDO e del pronto soccorso: il generico incidente fisico ha un proprio codice nella codifica utilizzata in queste due fonti (l'ICD9-CM) ma è poco utilizzato, poiché viene preferito codificare la frattura, se presente. Pertanto dalla

tabella 1.1, caduta, incidente e frattura sono stati approssimati in un unico outcome, la frattura.

2.4 Gli outcomes esclusi dall'analisi

Tra gli outcomes esclusi dall'analisi c'è l'istituzionalizzazione, ovvero il processo di inserimento, nell'ambito delle procedure assistenziali, di persone in stato di bisogno in strutture a loro dedicate a fini educativi e riabilitativi (Treccani, 2010). Nonostante sia un fenomeno legato alla fragilità largamente studiato in letteratura, non è stato possibile analizzarlo tramite le fonti sanitarie disponibili, e nessuna proxy è stata ritenuta valida. Nelle SDO è presente una variabile denominata "Modalità di dimissione" che tra le altre modalità assume anche quella di "dimissione ordinaria presso Residenza Sanitaria Assistenziale", ma usare solo questa fonte d'informazione sarebbe stato limitante, poiché si sarebbe limitato lo studio solo a coloro che erano stati ricoverati nel 2012 o nel 2013. Una possibile soluzione a questo problema sarebbe una maggior integrazione tra i dati amministrativi sanitari e quelli a disposizione dei servizi sociali, che sono in grado almeno parzialmente di far emergere quel "sommerso" di persone che non utilizzano i servizi sanitari frequentemente, ma che hanno caratteristiche tipiche degli individui fragili.

L'ospedalizzazione è stata esclusa in favore del ricovero d'urgenza: infatti quest'ultimo è chiaramente un sottoinsieme della prima. Si è preferito il ricovero d'urgenza, perchè è un evento più grave dell'ospedalizzazione programmata, e quindi è ipotizzabile che determini più facilmente una persona fragile.

Un altro possibile outcome che è stato scartato è la comorbidità, per tre principali motivi: il primo è dato dal fatto che per la maggior parte dei fattori di rischio che vengono usati per gli altri outcomes (malattie, ricoveri, poliprescrizione) non ha senso studiare una relazione causale con la comorbidità, poiché sono già deterministicamente legati ad essa. In aggiunta a ciò, non è possibile indagare i tipici fattori di rischio della comorbidità quali abitudini al fumo o all'alcool e la componente genetica o ereditarietà, che non sono riscontrabili nelle fonti sanitarie citate precedentemente. In terzo luogo in letteratura non sono molte le fonti in cui

la comorbidità viene trattata come variabile risposta piuttosto che come esplicativa.

Infine i giorni di ricovero totali sono stati esclusi in virtù della loro forte correlazione con il ricovero d'urgenza e a causa della loro altrettanto forte dipendenza con il decesso: infatti è evidente che un individuo deceduto nei primi giorni dell'anno può potenzialmente avere meno giorni di ricovero, e quindi essere valutato meno fragile, di una persona che viene ricoverata più volte nel corso dell'anno, pur senza morire. Una soluzione poteva essere quella di studiare i giorni di ricovero solo delle persone non decedute nel 2014: ma così facendo si sarebbero escluse le persone probabilmente più fragili, allontanandosi così dagli obiettivi di questo elaborato.

2.5 Le variabili esplicative disponibili nei dati

Dopo aver illustrato quali outcomes sono idonei all'analisi, in questo paragrafo viene svolto un primo analogo processo di selezione per le variabili esplicative.

2.5.1 Variabili socio-demografiche

Tra le variabili socio-demografiche sono state considerate l'età, il sesso e il reddito. Le prime due sono facilmente ricavabili dall'Anagrafe Sanitaria, mentre per il reddito è stata utilizzata come proxy l'esenzione ticket per reddito basso. Questa proxy ha un potenziale limite dovuto al fatto che dal momento che un soggetto ha già ottenuto l'esenzione per un altro motivo (malattia cronica, invalidità etc.) può decidere di non richiederla per il reddito, pur avendolo sotto la soglia stabilita dalla legge.

Le altre variabili relative alla sfera economico-finanziaria dell'individuo (indice di deprivazione e proprietà di una casa) non sono ricavabili dalle fonti a disposizione, così come il titolo di studio e quelle relative all'ambito sociale. Infine lo stato civile è stato scartato poiché, essendo disponibile solo tramite le SDO, sarebbe risultato mancante per tutti i cittadini non ospedalizzati nel 2012 o 2013.

2.5.2 Variabili relative allo stile di vita

Le variabili relative al comportamento passato e presente dell'individuo come l'attività fisica, l'abitudine al fumo e l'abuso di alcol non sono rintracciabili nei flussi sanitari correnti. Questo è un limite oggettivo, poiché queste variabili possono avere un forte effetto, sia diretto che indiretto (per esempio tramite i tumori), sulla fragilità dell'anziano.

2.5.3 Variabili relative a ricoveri precedenti

Le variabili relative ai ricoveri del paziente si riferiscono sia alle ospedalizzazioni programmate, sia a quelle d'urgenza che a quelle relative al pronto soccorso. Esempi di questa tipologia di variabili sono: numero giorni di ricovero, aver avuto almeno un ricovero urgente, numero di ricoveri con differenti diagnosi, aver avuto accesso al pronto soccorso e molte altre. Tutte le possibili variabili presenti in letteratura hanno trovato un facile riscontro con i dati a disposizione in questo elaborato, confermando la qualità delle fonti amministrative.

2.5.4 Variabili relative a malattie croniche

Vista la natura del fenomeno, è evidente che siano molteplici le malattie studiate come fattori di rischio della fragilità. Come mostrato nella tabella 1.2, nella letteratura si sono indagate queste cause entrando più o meno nello specifico: alcuni studi aggregano malattie in categorie più generiche come malattie cardiovascolari o malattie gastrointestinali, altre restano sullo specifico (asma, angina pectoris etc.). Generalmente in quest'analisi si è preferito un approccio più specifico nel selezionare un primo gruppo di variabili, che poi sono state ulteriormente scremate tramite l'analisi della prevalenza e delle misure di associazione con i diversi outcomes.

I flussi sanitari utilizzati variano a seconda della malattia: tutte le malattie sono state individuate tramite le diagnosi delle SDO e del pronto soccorso, che però, riferendosi solo agli ultimi due anni, possono essere limitanti; per alcune malattie sono stati integrati i flussi dell'esenzione e dei farmaci, che si sono rivelati

particolarmente utili su malattie di lunga durata e non ad alto rischio di ricovero, come per esempio il diabete.

La comorbidità è misurata dal Charlson Index, un indice che utilizza le diagnosi secondarie (dalla seconda alla sesta) dei ricoveri ospedalieri, pesandole secondo un criterio legato alla mortalità della patologia diagnosticata. I pesi sono stati scelti secondo il criterio proposto da Deyo (1992):

1 per infarto del miocardio, insufficienza cardiaca congestizia, malattie vascolari periferiche, demenza, disturbi cerebrovascolari, malattie polmonari croniche, malattie del tessuto connettivo, ulcera, malattie croniche del fegato e diabete;

2 per emiplegia, malattie renali moderate o gravi, diabete con danno d'organo, tumori, leucemia, linfomi;

3 per malattia epatica moderata o grave;

6 per comorbidità più gravi quali tumori maligni, metastasi e AIDS.

Per i pazienti che presentano più di una di queste malattie i pesi vengono sommati per farne un indice unico di comorbidità. Per esempio una persona sofferente di insufficienza cardiaca congestizia, demenza, ulcera ed emiplegia avrebbe un punteggio di 5 sulla scala definita dal Charlson Index.

2.5.5 Variabili relative all'assunzione di farmaci

Le ultime variabili della tabella 1.2 sono relative all'assunzione di determinati farmaci, come diuretici, antiaggreganti, antianemici eccetera. Nonostante la completezza e la pertinenza del flusso sanitario dei farmaci prescritti, non sono state inserite questo tipo di variabili esplicative per due motivi: in primo luogo perché una certa tipologia di farmaco può essere assunto da due pazienti con malattie specifiche differenti (per esempio i Beta-bloccanti possono essere presi sia per l'angina pectoris che per la fibrillazione atriale); in secondo luogo perché si è preferito integrare l'assunzione di un farmaco con altre fonti (come le diagnosi nelle SDO) laddove possibile, per determinare la presenza di una data malattia. Pertanto il flusso della farmaceutica non viene utilizzato per creare variabili relative solo all'assunzione di un determinato farmaco, ma viene utilizzato per

integrare le altre fonti a disposizione allo scopo di determinare se un individuo soffre di una certa patologia.

2.5.6 Variabili relative allo stato fisiologico

Come già accennato in precedenza, questo tipo di variabili si prestano a una particolare tipologia di raccolta dati, come il questionario o la scala di valutazione, e quindi a un certo tipo di informazione, che non può essere ricavato dai flussi sanitari correnti ma che richiede l'integrazione dei flussi sociali, in grado di definire con accuratezza il livello di autonomia di una persona. Pertanto tutte le variabili riguardanti la capacità dell'individuo nello svolgere azioni quotidiane come lavarsi, vestirsi, cucinare in maniera indipendente, non possono entrare nell'analisi di questo elaborato, così come la salute percepita, l'energia, la forza di presa etc. Le uniche due variabili a disposizione restano pertanto l'utilizzo di assistenza domiciliare integrata e l'esenzione di invalidità.

2.5.7 Variabili relative allo stato psicologico o mentale

Le variabili relative allo stato psicologico e mentale, come demenza, depressione/ansia, ritardo mentale sono state inserite nell'analisi, e sono quelle che maggiormente sfruttano l'integrazione tra più fonti sanitarie. Infatti oltre alle diagnosi delle SDO e del pronto soccorso, ai farmaci e alle esenzioni, il flusso della psichiatria territoriale permette di ricavare tramite il codice ICD-10 le diagnosi per i disturbi psicologici e mentali che sono presenti nella tabella 1.2.

Nella tabella 2.1 sono elencate le variabili esplicative che hanno “superato” questa prima fase di selezione, e le fonti sanitarie da cui vengono ricavate. Per quanto riguarda le malattie, è necessario specificare che sono state utilizzate sia le diagnosi principali rilevate al momento del ricovero in ospedale o al pronto soccorso, sia quelle secondarie, ovvero quelle che non sono state il motivo del ricovero ma di cui comunque il paziente è affetto in quel momento. Inoltre il flusso della Farmaceutica è stato integrato alle altre fonti solo per determinare la presenza

di malattie che hanno degli specifici farmaci di riferimento, come gli antidiabetici e gli antidepressivi.

Tabella 2.1 Variabili esplicative della fragilità selezionate tramite la ricerca bibliografica e relative fonti dati.

Variabile	Fonte dati
Età	Anagrafe sanitaria
Sesso	Anagrafe sanitaria
Reddito	Esenzioni ticket
Charlson index	SDO
Cadute nei precedenti 12 mesi/ Frattura femore	SDO, Pronto soccorso
Storia di Ricoveri urgenti	SDO
Parziale/totale invalidità	Esenzioni ticket
Accesso ai servizi di salute mentale/disturbi psichiatrici	Psichiatria territoriale
Assistenza domiciliare integrata	Assistenza domiciliare integrata
Poliprescrizione	Farmaceutica territoriale
Storia di ospedalizzazione precedenti	SDO, Pronto soccorso
Ricoveri con differenti diagnosi	SDO
Malattia mentale	SDO, Pronto soccorso, Psichiatria territoriale, Esenzioni ticket
CAD: Coronary Artery Disease	SDO, Pronto soccorso, Esenzioni ticket
Infarto miocardio	SDO, Pronto soccorso, Esenzioni ticket
Angina pectoris	SDO, Pronto soccorso, Esenzioni ticket
PVD: Peripheral Vascular Disease	SDO, Pronto soccorso
VHD: Vascular Heart Disease	SDO, Pronto soccorso, Esenzioni ticket
CHF: Congestive Heart Failure	SDO, Pronto soccorso, Esenzioni ticket
Asma	SDO, Pronto soccorso, Esenzioni ticket
Fibrillazione atriale	SDO, Pronto soccorso, Esenzioni ticket
Epilessia	SDO, Pronto soccorso, Esenzioni ticket, Farmaceutica territoriale
Bronchite	SDO, Pronto soccorso
Bronchiectasia	SDO, Pronto soccorso
Enfisema	SDO, Pronto soccorso
COPD: Chronic Obstructive Pulmonary Disease	SDO, Pronto soccorso
Artrosi	SDO, Pronto soccorso, Esenzioni ticket
Diabete	SDO, Pronto soccorso, Esenzioni ticket, Farmaceutica territoriale
Ipertensione	SDO, Pronto soccorso, Esenzioni ticket
Morbo di Parkinson	SDO, Pronto soccorso, Esenzioni ticket, Farmaceutica territoriale
Epatite	SDO, Pronto soccorso, Esenzioni ticket
Trombosi Venosa	SDO, Pronto soccorso
Malattia infiammatoria pelvica	SDO, Pronto soccorso
Pielonefrite	SDO, Pronto soccorso

Variabile	Fonte dati
Gastroenterite	SDO, Pronto soccorso
Cirrosi	SDO, Pronto soccorso, Esenzioni ticket
Demenza	SDO, Pronto soccorso, Esenzioni ticket, Psichiatria territoriale, Farmaceutica territoriale
Insufficienza renale	SDO, Pronto soccorso, Esenzioni ticket
Morbo di Alzheimer	SDO, Pronto soccorso, Esenzioni ticket, Psichiatria territoriale
Ictus	SDO, Pronto soccorso, Esenzioni ticket
Anemia	SDO, Pronto soccorso, Farmaceutica territoriale
Ipercolesterolemia	SDO, Pronto soccorso
Depressione	SDO, Pronto soccorso, Esenzioni ticket, Psichiatria territoriale, Farmaceutica territoriale
Difficoltà visive (cataratta, glaucoma)	SDO, Pronto soccorso, Esenzioni ticket
Numero di giorni in ospedale	SDO
Numero di giorni dall'ultima ospedalizzazione	SDO
Numero totale di giorni ricoverato d'urgenza	SDO
Numero totale di giorni di ospedalizzazione	SDO

Capitolo 3

Selezione delle variabili esplicative della fragilità

Dopo una prima selezione delle variabili tramite un'approfondita ricerca bibliografica, sono stati applicati dei metodi statistici per selezionare un numero ristretto di variabili in grado di spiegare un complesso fenomeno come la fragilità. Tale riduzione è necessaria per fornire un algoritmo basato su poche variabili in grado di quantificare in modo soddisfacente la fragilità negli anziani.

Inizialmente è stata stabilita la codifica delle variabili continue che sono state riorganizzate in classi tramite degli alberi decisionali.

In seguito sono state escluse alcune variabili tramite delle analisi descrittive preliminari: sono state escluse alcune variabili a prevalenza molto bassa, altre perché troppo correlate con altri fattori di rischio e altre perché non significativamente associate, in termini di odds ratio, ai cinque outcomes studiati.

Terminata l'analisi preliminare, sono state svolte cinque analisi statistiche distinte per i rispettivi outcomes, ovvero il decesso, il ricovero urgente, la frattura, la demenza e la disabilità. Per ognuna di queste variabili dipendenti sono state implementate due tecniche per la selezione di un sottoinsieme di esplicative, la regressione logistica Stepwise e la regressione logistica Lasso, e per ogni outcome è stato scelto il sottoinsieme di covariate stabilito dal metodo migliore per capacità predittiva, scelto sulla base di 100 diversi campioni bilanciati. A questo punto sono state selezionate solo le variabili esplicative che in media, nelle cinque analisi, sono state selezionate almeno in 60 campioni su 100, in modo da creare un sottoinsieme ridotto di variabili informative.

3.1 La codifica delle variabili

La maggior parte delle variabili sono dicotomiche, con il valore 0 ad indicare l'assenza e 1 la presenza. Alcune variabili di natura continua sono state ricodificate in variabili categoriali, costruendo le modalità tramite degli alberi decisionali.

Ogni outcome suggerisce delle codifiche diverse, ma poiché è necessario averne solo una, l'albero è stato sviluppato sulla variabile risposta più grave, ovvero il decesso nell'anno 2014, sul totale della popolazione di quell'anno.

In particolar modo sono 6 le variabili rese categoriali:

- l'età;
- il Charlson Index;
- il numero di farmaci;
- il numero di accessi al pronto soccorso tramite il 118;
- il numero di ricoveri con differenti diagnosi;
- il numero di ospedalizzazioni precedenti;
- il numero di giorni di ricovero d'urgenza.

Nella tabella 3.1 sono descritte tutte le codifiche utilizzate.

Tabella 3.1 Variabili esplicative della fragilità e relativa codifica utilizzata nei modelli di regressione.

Variabile	Codifica
Età	65-69; 70-74; 75-79; 80-84; 85-89; 90+
Sesso	0=Femminile; 1=Maschile
Esenzione per reddito basso	0=Non esente; 1=Esente
Esenzione per invalidità	0=Non esente; 1=Esente
Frattura nei 2 anni precedenti	0= No; 1=Si
Storia di Ricoveri urgenti – Almeno un ricovero urgente nell'anno precedente	0=No; 1=Si
Charlson index	0; 1; 2-4; 5+
Accesso ai servizi di salute mentale/disturbi psichiatrici	0= No; 1=Si
Assistenza domiciliare integrata (ADI)	0= No; 1=Si
Poliprescrizione	0-5; 6-10; 11+
Storia di ospedalizzazione precedenti – Numero di ospedalizzazioni precedenti	0; 1; 2; 3+
Accesso al Pronto Soccorso tramite 118	0; 1; 2-3; 4+
Ricoveri con differenti diagnosi	0; 1; 2; 3+
Malattia mentale	0= No; 1=Si
CAD: Coronary Artery Disease	0= No; 1=Si
Infarto miocardio	0= No; 1=Si
Angina pectoris	0= No; 1=Si
PVD: Peripheral Vascular Disease	0= No; 1=Si
VHD: Vascular Heart Disease	0= No; 1=Si
CHF: Congestive Heart Failure	0= No; 1=Si
Asma	0= No; 1=Si

Variabile	Codifica
Fibrillazione atriale	0= No; 1=Si
Epilessia	0= No; 1=Si
Bronchite	0= No; 1=Si
Bronchiectasia	0= No; 1=Si
Enfisema	0= No; 1=Si
COPD: Chronic Obstructive Pulmonary Disease	0= No; 1=Si
Artrosi	0= No; 1=Si
Diabete	0= No; 1=Si
Ipertensione	0= No; 1=Si
Morbo di Parkinson	0= No; 1=Si
Epatite	0= No; 1=Si
Trombosi Venosa	0= No; 1=Si
Malattia infiammatoria pelvica	0= No; 1=Si
Pielonefrite	0= No; 1=Si
Gastroenterite	0= No; 1=Si
Cirrosi	0= No; 1=Si
Demenza	0= No; 1=Si
Insufficienza renale	0= No; 1=Si
Morbo di Alzheimer	0= No; 1=Si
Ictus	0= No; 1=Si
Anemia	0= No; 1=Si
Ipercolesterolemia	0= No; 1=Si
Depressione	0= No; 1=Si
Difficoltà visive (cataratta, glaucoma)	0= No; 1=Si
Numero totale di giorni ricoverato d'urgenza	0-2; 3-11; 12-39; 40+

3.2 Analisi preliminari sui fattori di rischio

Prima di procedere con i metodi statistici citati, è stata svolta un'ulteriore pulizia dei dati. In primo luogo è stato deciso di escludere delle malattie specifiche in favore di gruppi di malattie che hanno un proprio codice diagnostico nelle SDO: pertanto l'angina pectoris e l'infarto del miocardio sono stati esclusi poiché la presenza di una delle due condizioni o di un'altra cardiopatia ischemica determina il CAD (Coronary Artery Disease). Allo stesso modo bronchite cronica, bronchiectasie e enfisema sono stati scartati in favore del gruppo di malattie COPD (Chronic Obstructive Pulmonary Disease).

Analogamente la demenza e il morbo di Alzheimer appartengono al gruppo di malattie a cui si riferisce la variabile "Malattia mentale", che comprende anche disturbi nevrotici, psichici, della personalità e ritardo mentale. Vista l'importanza

della demenza e del morbo di Alzheimer nel determinare la fragilità, è stato deciso di creare due variabili complementari, separando queste due malattie dalle restanti e creando così due variabili dicotomiche distinte.

Per quanto riguarda le due variabili che si riferiscono al pronto soccorso, è stato scelto di tenere quella con annessa chiamata al 118, ritenendo la caratteristica intrinseca di urgenza in questa variabile più adatta al contesto della fragilità, così come si era preferito l'outcome "ricovero urgente" rispetto a "ospedalizzazione". Analogamente si è preferito considerare i giorni di ricovero urgente, piuttosto che quelli per un'ospedalizzazione generica.

L'analisi della prevalenza ha portato ad escludere le malattie meno frequenti (con prevalenza sotto l'1%) nel dataset, come cirrosi, epatite, gastroenterite, osteoporosi, trombosi venosa, pielonefrite e malattia infiammatoria pelvica, in modo da evitare stime poco precise nelle successive analisi. È necessario specificare che la prevalenza di queste malattie è condizionata dalle fonti sanitarie a disposizione e non è sempre attendibile: per esempio secondo un recente studio epidemiologico multicentrico nazionale, ESOPO, l'osteoporosi in Italia ha una prevalenza del 23% tra le donne di oltre 40 anni e del 14% tra gli uomini con più di 60 anni (Epicentro, 2015); ciononostante è una condizione che non comporta frequenti ricoveri ospedalieri come altre malattie, e i farmaci che vengono prescritti per gli osteoporotici possono essere associati anche a condizioni meno gravi (per esempio gli antinfiammatori).

L'analisi sulla correlazione tra i fattori di rischio è stata svolta tramite il calcolo del coefficiente di Kendall, una statistica che misura il grado di covariazione tra due variabili ordinali basandosi sul numero di coppie concordanti e discordanti:

$$\tau = \frac{2 (\text{numero di coppie concordanti} - \text{numero di coppie discordanti})}{n(n-1)}$$

Date due variabili X e Y e una coppia di osservazioni i e j, (x_i, y_i) e (x_j, y_j) , con $i \neq j$, si dice che la coppia è concordante se $x_i > x_j$ e $y_i > y_j$ oppure se $x_i < x_j$ e $y_i < y_j$. Viceversa la coppia è discordante se $x_i < x_j$ e $y_i > y_j$ oppure se $x_i > x_j$ e $y_i < y_j$. Se $x_i = x_j$ o $y_i = y_j$ allora la coppia non è né concordante né discordante. Con n si intendono il numero di osservazioni, in modo che il denominatore sia il numero totale di coppie

possibili. Il coefficiente di Kendall varia tra -1 e 1, indicando con 0 l'assenza di covariazione.

Sono emerse delle covariazioni particolarmente alte per queste variabili:

Tabella 3.2: Variabili esplicative con la più alta correlazione di Kendall.

	Ricoveri con differenti diagnosi	Almeno un ricovero nell'anno precedente	Numero di ospedalizzazioni precedenti	Numero di giorni ricoverato d'urgenza
Ricoveri con differenti diagnosi	1	0.52377 (<.0001)	0.75391 (<.0001)	0.66228 (<.0001)
Almeno un ricovero nell'anno precedente		1	0.15042 (<.0001)	0.71843 (<.0001)
Numero di ospedalizzazioni precedenti			1	0.19735 (<.0001)
Numero di giorni ricoverato d'urgenza				1

Tra le 4 variabili presenti nella tabella, è stato scelto di escludere dall'analisi il numero di giorni di ricovero, principalmente perché, oltre all'alta correlazione con le altre variabili, è citato più raramente nella letteratura come fattore di rischio, rispetto alle altre tre variabili. Nonostante l'alta correlazione tra l'aver avuto un ricovero urgente e differenti diagnosi, si è preferito lasciare che i successivi metodi statistici selezionassero la variabile informativa più idonea nel contesto della fragilità.

L'odds ratio è una misura di associazione tra due fattori, come un fattore di rischio e una malattia, ed è definito come la quota dei "malati" tra i soggetti esposti diviso la quota dei "malati" tra i soggetti non esposti al fattore di rischio. Data Y una delle cinque variabili risposta dell'analisi, e X una delle esplicative dicotomiche della tabella 3.1:

$$OR = \frac{Pr(Y = 1|X = 1)/Pr(Y = 0|X = 1)}{Pr(Y = 1|X = 0)/Pr(Y = 0|X = 0)}$$

Sono stati calcolati gli odds ratio per valutare l'associazione tra i cinque outcomes selezionati e le variabili dicotomiche, e solo l'asma e l'ipercolesterolemia non sono risultati significativi per tutti gli outcomes (o in alcuni casi hanno avuto odds ratio minore di 1), e pertanto sono state escluse dalle successive analisi.

Tabella 3.3: Odds ratio tra i cinque outcomes e le variabili esplicative dicotomiche.

	Decesso	Ricovero urgente	Frattura	Disabilità	Demenza
Sesso (Maschio=1 vs Femmina=0)	n.s (0.93-1.13)	1.1 (1.04-1.16)	0.56 (0.5-0.62)	0.85 (0.76-0.95)	0.65 (0.55-0.76)
ADI: assistenza domiciliare integrata	9.07 (8.2-10.04)	4.67 (4.35-5.02)	2.12 (1.85-2.43)		6.04 (5.15-7.07)
Anemia	4.99 (4.36-5.72)	3.73 (3.38-4.12)	2.25 (1.88-2.70)	3.20 (2.58-3.95)	2.28 (1.75-2.97)
Artrosi	n.s (0.99-1.52)	1.42 (1.25-1.61)	1.40 (1.13-1.73)	1.61 (1.26-2.03)	1.44 (1.05-1.95)
Asma	n.s (0.34-1.03)	n.s (0.88-1.44)	n.s (0.82-1.86)	n.s (0.71-1.84)	n.s (0.51-1.93)
CAD: cardiovascular artery disease	1.86 (1.6-2.14)	2.23 (2.04-2.43)	1.22 (1.02-1.46)	1.60 (1.33-1.93)	n.s (0.94-1.58)
Cancro	2.20 (1.98-2.46)	1.44 (1.34-1.55)	1.19 (1.05-1.36)	2.40 (2.11-2.72)	n.s (0.77-1.17)
CHF: congestive heart failure	6.75 (5.86-7.78)	5.53 (4.97-6.16)	1.93 (1.54-2.4)	5.36 (4.28-6.7)	2.86 (2.16-3.77)
COPD: chronic obstructive pulmonary disease	6.04 (4.96-7.35)	5.68 (4.9-6.59)	2.20 (1.65-2.95)	5.34 (3.8-7.51)	2.25 (1.45-3.45)
Demenza	5.49 (4.62-6.5)	3.46 (3.03-3.96)	1.69 (1.28-2.21)	13.55 (10.56-17.37)	
Depressione	2.26 (2.01-2.54)	1.96 (1.82-2.11)	2.08 (1.83-2.35)	3.09 (2.69-3.53)	4.63 (3.96-5.41)
Diabete	1.80 (1.61-2.01)	1.85 (1.73-1.98)	1.17 (1.02-1.32)	1.62 (1.44-1.89)	1.38 (1.15-1.65)
Epilessia	2.27 (1.91-2.7)	2.10 (1.87-2.34)	1.60 (1.31-1.97)	2.93 (2.36-3.64)	2.59 (2.01-3.33)

	Decesso	Ricovero urgente	Frattura	Disabilità	Demenza
Esenzione reddito	n.s (0.83-1.06)	1.23 (1.13-1.33)	1.23 (1.07-1.41)	1.55 (1.31-1.82)	1.28 (1.03-1.58)
Fibrillazione atriale	2.97 (2.63-3.35)	2.44 (2.25-2.65)	1.55 (1.32-1.80)	2.11 (1.79-2.48)	1.64 (1.32-2.04)
Frattura	2.13 (1.81-2.51)	1.84 (1.65-2.05)	2.71 (2.31-3.16)	2.02 (1.63-2.50)	2.95 (2.36-3.68)
Glaucoma	n.s (0.94-1.37)	1.31 (1.17-1.46)	n.s (0.99-1.46)	1.32 (1.06-1.62)	n.s (0.98-1.7)
Ictus	3.29 (2.81-3.84)	2.62 (2.35-2.92)	1.54 (1.24-1.90)	3.59 (2.93-4.39)	2.65 (2.06-3.43)
Insufficienza renale	4.31 (3.50-5.31)	4.08 (3.52-4.74)	1.45 (1.03-2.01)	3.82 (2.83-5.15)	n.s (0.92-2.44)
Invalidità	5.97 (5.41-6.58)	3.33 (3.13-3.54)	2.08 (1.86-2.32)		4.95 (4.27-5.74)
Iperensione	1.71 (1.55-1.88)	1.72 (1.62-1.82)	1.32 (1.18-1.45)	1.80 (1.6-2)	1.45 (1.25-1.68)
Ipercolesterolemia	0.68 (0.53-0.86)	n.s (0.87-1.11)	n.s (0.65-1.03)	n.s (0.78-1.25)	n.s (0.66-1.28)
Malattie mentali	1.41 (1.1-1.83)	2.02 (1.76-2.32)	1.62 (1.22-2.08)	2.45 (1.85-3.23)	4.01 (3.05-5.27)
Parkinson	2.43 (1.93-3.07)	2.07 (1.77-2.43)	1.37 (1.01-1.86)	5.44 (4.11-6.92)	4.34 (3.24-5.79)
Ricovero nell'ultimo anno	6.47 (5.82-7.2)	4.36 (4.05-4.70)	2.14 (1.85-2.45)	5.49 (4.76-6.31)	3.18 (2.63-3.83)
Disturbi tiroidei	1.17 (1.01-1.36)	1.20 (1.1-1.32)	1.23 (1.04-1.43)	n.s (0.94-1.36)	n.s (0.87-1.41)
VHD: vascular heart disease	2.39 (1.87-3.07)	2.23 (1.89-2.63)	n.s (0.8-1.63)	2.10 (1.53-2.89)	n.s (0.63-1.85)

3.3 Metodi di selezione delle variabili

A questo punto dell'analisi, il problema si è ridotto a 5 variabili risposta dicotomiche, e 31 variabili indipendenti, di natura categoriale. L'obiettivo di

questo capitolo è determinare un sottoinsieme delle 31 variabili, tenendo in considerazione le loro relazioni causali con tutte le variabili dipendenti. Per ottenere ciò sono state considerate delle procedure che garantissero sia accuratezza predittiva che interpretabilità del modello scelto. In particolare ci sono due classi di metodi per la selezione delle variabili:

1. Subset Selection: questo approccio identifica un sottoinsieme di predittori che possa spiegare adeguatamente la variabile risposta. Tra i metodi più popolari ci sono la “Best Subset Selection” e la “Stepwise Selection” (James, 2013).
2. Shrinkage: questi metodi stimano modelli a partire da tutti i predittori, ma “schiacciano” verso lo zero i coefficienti delle variabili meno informative. Questo metodo funziona particolarmente bene quando il numero di predittori p è maggiore del numero di osservazioni n , riducendo la varianza delle stime a costo di un leggero aumento della distorsione. Un esempio di metodo di shrinkage è la “Ridge Regression” e la sua generalizzazione per la selezione delle variabili è il “Lasso” (James, 2013).

Un terzo approccio è quello di riduzione della dimensionalità, rappresentato dall’analisi fattoriale, che però non è adeguato allo scopo del seguente capitolo, poiché riduce l’analisi a dei fattori che sono combinazioni, lineari o non lineari, delle variabili originali, che, data la loro difficile interpretazione, non sono presentabili all’azienda Ulss.

Il metodo della Best Subset Selection è stato escluso per il suo costo computazionale: infatti esso prevede la stima di ogni possibile combinazione dei p predittori, scegliendo il miglior modello prima tra tutti i modelli con un certo numero k di predittori, tramite l’ R^2 per un modello lineare o la devianza per un modello logistico, poi tra tutti i modelli migliori tramite l’AIC o un altro criterio d’informazione o con il calcolo dell’errore predittivo della convalida incrociata. Questo comporta che al primo passo vengano valutati 2^p possibili modelli e, scelti i migliori, al secondo passo $p+1$ modelli. Nel caso di questo elaborato vorrebbe

dire valutare più di due miliardi di modelli statistici, un onere computazionale eccessivo.

Un metodo computazionalmente meno oneroso è quello della selezione Stepwise, che consiste nel partire dal modello con la sola intercetta, e a questa aggiungere la variabile che abbassa maggiormente la devianza. Si procede poi allo stesso modo inserendo un nuovo predittore o eliminandone uno già presente nel modello al fine di ridurre significativamente la variabilità della variabile risposta. La procedura itera finché il modello con un certo numero k di predittori non è migliorabile tramite l'inserimento o l'eliminazione di una variabile esplicativa (Fabbris, 1997). Naturalmente questo approccio è migliore in termini di costo computazionale rispetto alla Best Subset Selection, ma non garantisce come risultato il modello migliore possibile, poiché non esplora tutti i possibili modelli, ma solo quelli condizionati al percorso dell'algoritmo.

La Regressione Ridge è, in ambito di modelli generalizzati, molto simile alla log-verosimiglianza, ma in questo caso la stima dei parametri viene ricavata dalla massimizzazione della log-verosimiglianza con una penalità sulla norma di β :

$$l(\beta, \lambda) = l(\beta) - \lambda \sum_{j=1}^p \beta_j^2$$

Il termine sottratto alla log-verosimiglianza è chiamato “shrinkage penalty”, ed è tanto più piccolo quanto i coefficienti si avvicinano allo zero, e questo ha l'effetto di “schiacciare” le stime verso lo zero. Il parametro λ regola l'impatto dello shrinkage: se $\lambda=0$, la penalità non ha effetto e si ottiene la stima di massima verosimiglianza; se $\lambda \rightarrow \infty$ l'impatto della penalità cresce fino ad avere tutte le stime vicine allo 0. Naturalmente il parametro λ va scelto in modo che minimizzi l'errore predittivo, per esempio tramite convalida incrociata.

La Regressione Ridge non è adeguata a rispondere all'esigenza di selezionare solo le variabili più informative, ma tra le sue generalizzazioni ha il Lasso, che introducendo una penalità più restrittiva, “obbliga” le stime dei coefficienti meno informativi ad essere esattamente uguali a 0, introducendo una forma di selezione delle variabili:

$$l(\beta, \lambda) = l(\beta) - \lambda \sum_{j=1}^p |\beta_j|$$

Il numero di variabili selezionate dipenderà dal parametro λ , che sarà scelto in funzione dell'errore di previsione.

3.4 Selezione Stepwise e Lasso a confronto: analisi e risultati

I metodi più adeguati al problema di riferimento sono indubbiamente la regressione Stepwise e il Lasso: per determinare il parametro di regolazione del Lasso è stata implementata una procedura di convalida incrociata con 10 partizioni del dataset, in modo che ad ogni passo della procedura una partizione è usata come dataset di verifica e le restanti come un unico dataset di stima del parametro λ . Le variabili risposta sono tutte dicotomiche, pertanto sono state implementati i metodi in ambito logistico, con funzione legame Logit.

Per ogni outcome è stata costruita la matrice delle esplicative formata dalle 31 variabili selezionate precedentemente, ovvero le variabili della tabella 3.3 ad esclusione di asma e ipercolesterolemia, ed è stato valutato quale dei due metodi fosse il migliore in termini di previsione tramite l'AUC, ovvero l'area sotto la curva ROC.

La curva ROC è uno strumento grafico per valutare la bontà di un metodo di classificazione: nell'asse delle ascisse viene posta la proporzione dei “falsi positivi”, ovvero la proporzione dei soggetti che vengono erroneamente predetti positivamente, mentre nell'asse delle ordinate viene rappresentata la proporzione di “veri positivi”, ovvero coloro i quali sono stati valutati correttamente positivi. L'area sotto la curva ROC è una misura della bontà predittiva del modello, in particolar modo è ritenuta soddisfacente se è compresa tra 0,7 e 1.

Un problema che è emerso dalle cinque diverse classificazioni è il bilanciamento della variabile risposta: infatti solo il 3,4% dei soggetti sperimenta il decesso, il 10,49% il ricovero d'urgenza, 3,13% la frattura, 3,22% la disabilità e 1,46% la demenza. Il problema che ne deriva è che un classificatore non riesce a discriminare gli osservati “positivi”, ottenendo così un'elevata percentuale di falsi

negativi. È quindi necessario procedere con un campionamento casuale semplice per ogni outcome tra i soggetti che non hanno sperimentato l'evento, ottenendo dei campioni equilibrati. Affinché i risultati non dipendano da un solo campione bilanciato, per ogni outcome si sono costruiti 100 diversi campioni bilanciati e per ogni campione si è valutato in termini di auc quale tecnica tra la Stepwise e il Lasso risultasse migliore: per far ciò è stato usato il dataset di stima (bilanciato) avente le covariate rilevate nel biennio 2012-2013 e le risposte nel 2014, e il dataset di verifica avente le stesse covariate relative al biennio 2013-2014 e le stesse variabili risposta relative 2015.

Tabella 3.4: Struttura dei dataset di stima e di validazione.

	2012	2013	2014	2015
Stima	Covariate	Covariate	Risposta	
Verifica		Covariate	Covariate	Risposta

Un altro aspetto da chiarire è quello riguardante la composizione dei dataset per i diversi outcomes, in relazione alla tipologia di evento che rappresentano: il decesso, il ricovero urgente e la frattura sono eventi per cui tutta la popolazione è potenzialmente a rischio, compresi coloro i quali hanno già avuto uno di questi eventi (ad eccezione del decesso ovviamente) negli anni precedenti all'anno di rilevazione; la disabilità e la demenza invece sono eventi irreversibili, tali per cui solo chi non ha una delle due condizioni può definirsi a rischio. Poiché la regressione logistica studia relazioni causali unidirezionali, dai fattori di rischio alla variabile risposta, per la demenza sono stati scartati, sia dal dataset di stima sia da quello di verifica, tutti coloro che avevano la variabile demenza=1 negli anni in cui sono state rilevate le covariate, e analogamente per la disabilità sono stati esclusi tutti coloro che avevano la variabile Assistenza domiciliare integrata o l'esenzione di invalidità uguale alla modalità 1, poiché in questo modo è stata costruita la variabile risposta disabilità. Questo procedimento è fondamentale altrimenti la relazione causale tra i fattori di rischio e l'outcome (disabilità/demenza) non sarebbe stata rispettata, ottenendo risultati non credibili.

Tabella 3.5: Numerosità dei dataset di stima e di verifica per ogni variabile risposta.

	Decesso	Ricovero urgente	Frattura	Disabilità	Demenza
Stima (2014)	49132	49132	49132	38866	48018
Verifica (2015)	50733	50733	50733	40710	49542

3.4.1 Decesso

La procedura descritta nel paragrafo precedente porta a preferire la selezione Stepwise in 71 dei 100 campioni bilanciati, con una auc mediamente pari a 0,855. La frequenza di ingresso nelle 100 Stepwise per ogni variabile esplicativa è descritta dal seguente grafico:

Frequenze d'ingresso nelle Stepwise: Decesso

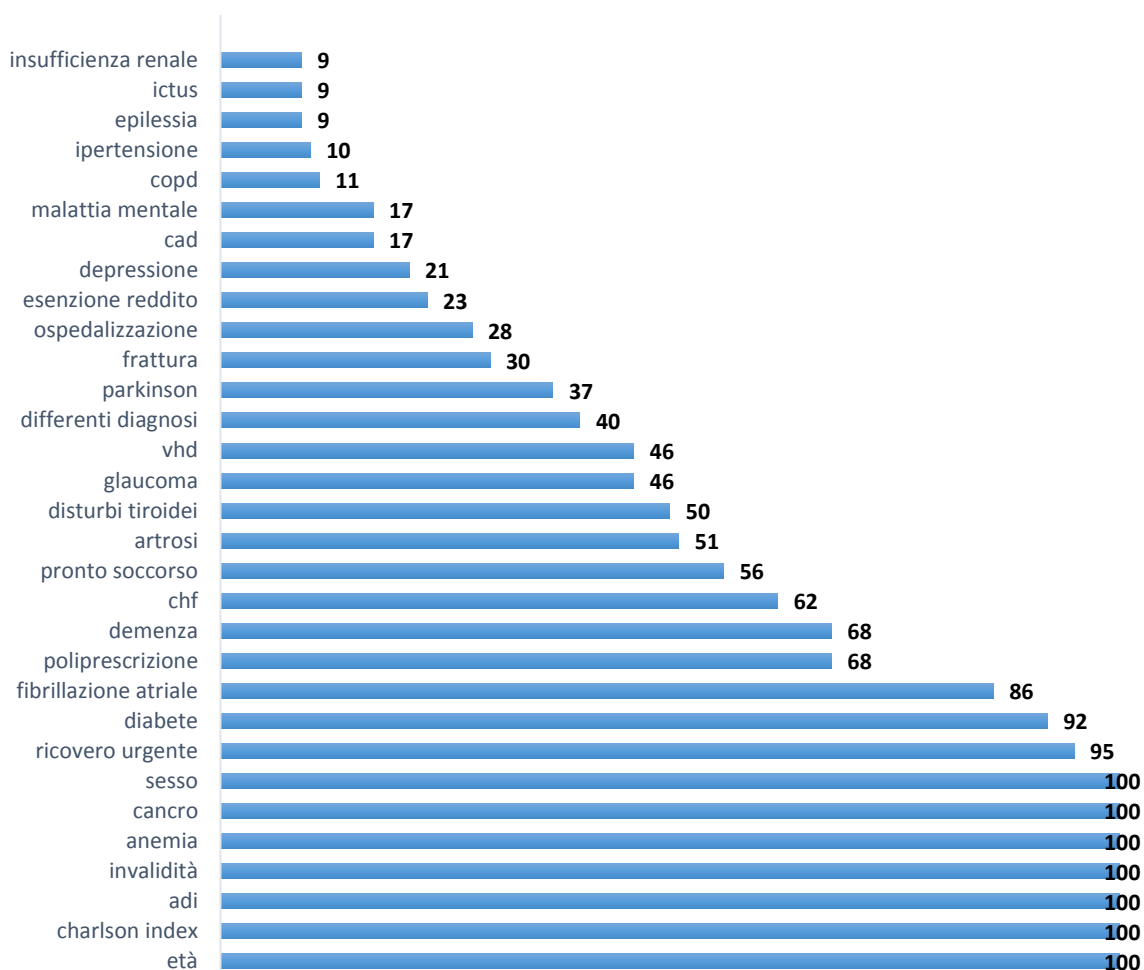


Figura 3.1: Frequenze d'ingresso delle variabili esplicative per l'outcome decesso.

Le variabili che eguagliano o superano la soglia dei 95 “ingressi” sono: ricovero urgente, sesso, cancro, anemia, invalidità, ADI, Charlson Index ed età. Implementando una regressione logistica con queste variabili per valutare la “direzione” del rischio, si è notato che un aumento di qualsiasi variabile categoriale comporta un progressivo aumento del rischio di decesso, così come per tutte le variabili dicotomiche, la modalità “1” (presenza del fattore di rischio) aumenta il rischio: questo significa che il sesso maschile è più a rischio di morte rispetto a quello femminile.

3.4.2 Ricovero urgente

Per il ricovero urgente, la selezione Stepwise viene preferita al Lasso nel 97% dei campioni bilanciati, con una auc media pari a 0,759. Le variabili che superano la soglia dei 95 ingressi sono la fibrillazione atriale, il cancro, il sesso, la malattia mentale, il CAD, l’anemia, il diabete, gli accessi al pronto soccorso, la poliprescrizione, l’invalidità, l’ADI e l’età.

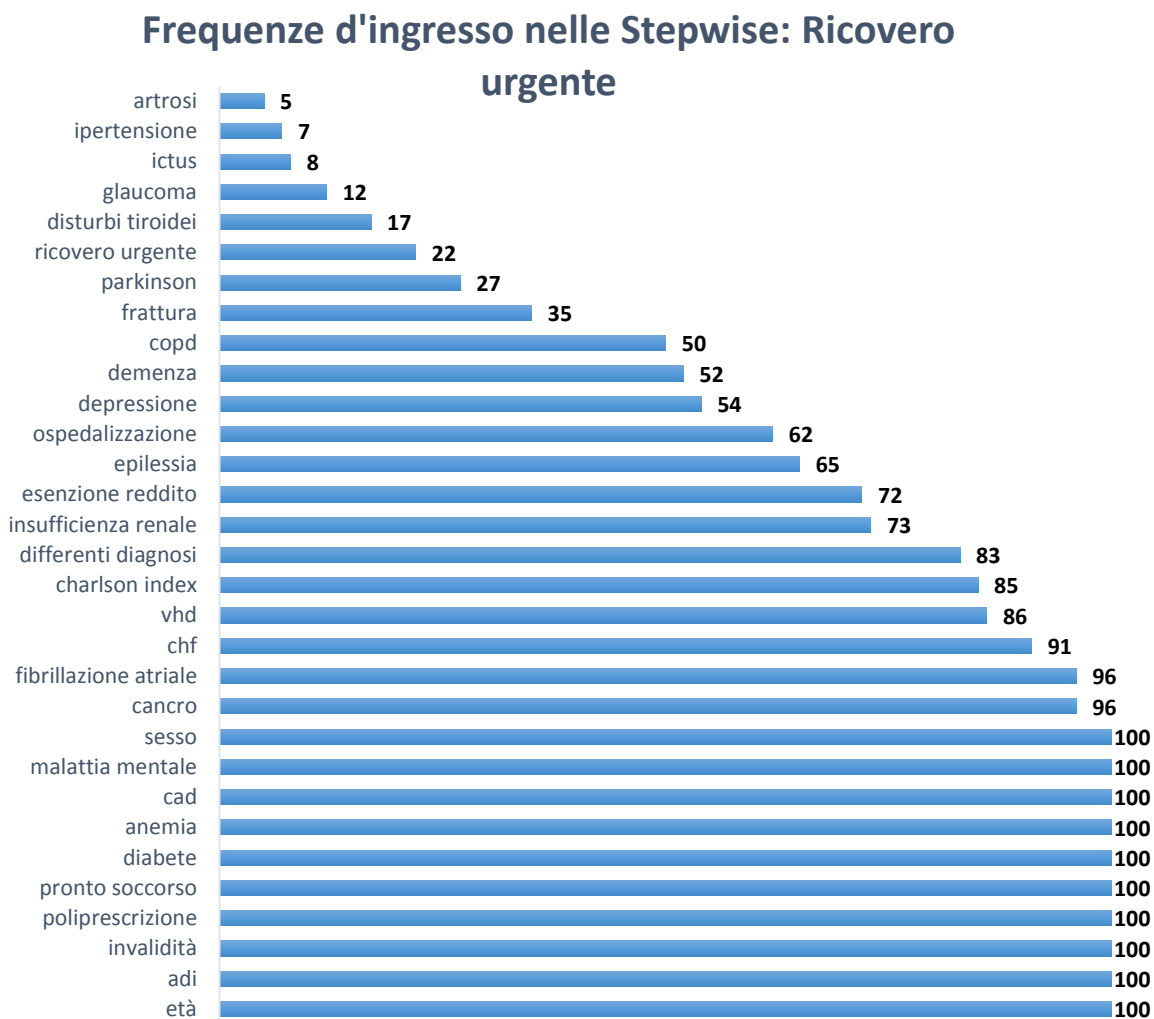


Figura 3.2: Frequenze d'ingresso delle variabili esplicative per l'outcome ricovero urgente.

Anche questo caso le variabili relative alle malattie e ai ricoveri assumono la direzione di rischio più ovvia e inoltre il sesso maschile è più a rischio di ricovero rispetto a quello femminile, come nel caso del decesso.

3.4.3 Frattura

Nella frattura la selezione Stepwise viene preferita al lasso in 66 campioni bilanciati, con una auc media di 0,646. Le variabili più selezionate dalla procedura sono state il sesso, aver avuto una frattura precedente, la depressione, gli accessi al pronto soccorso e l'età. In questo caso è risultato che le donne hanno un maggior rischio di frattura rispetto agli uomini: questo significa che il sesso ha effetti

significativi ma opposti tra i vari outcomes, e quindi verrà escluso nelle successive analisi.

Frequenze d'ingresso nelle Stepwise: Frattura

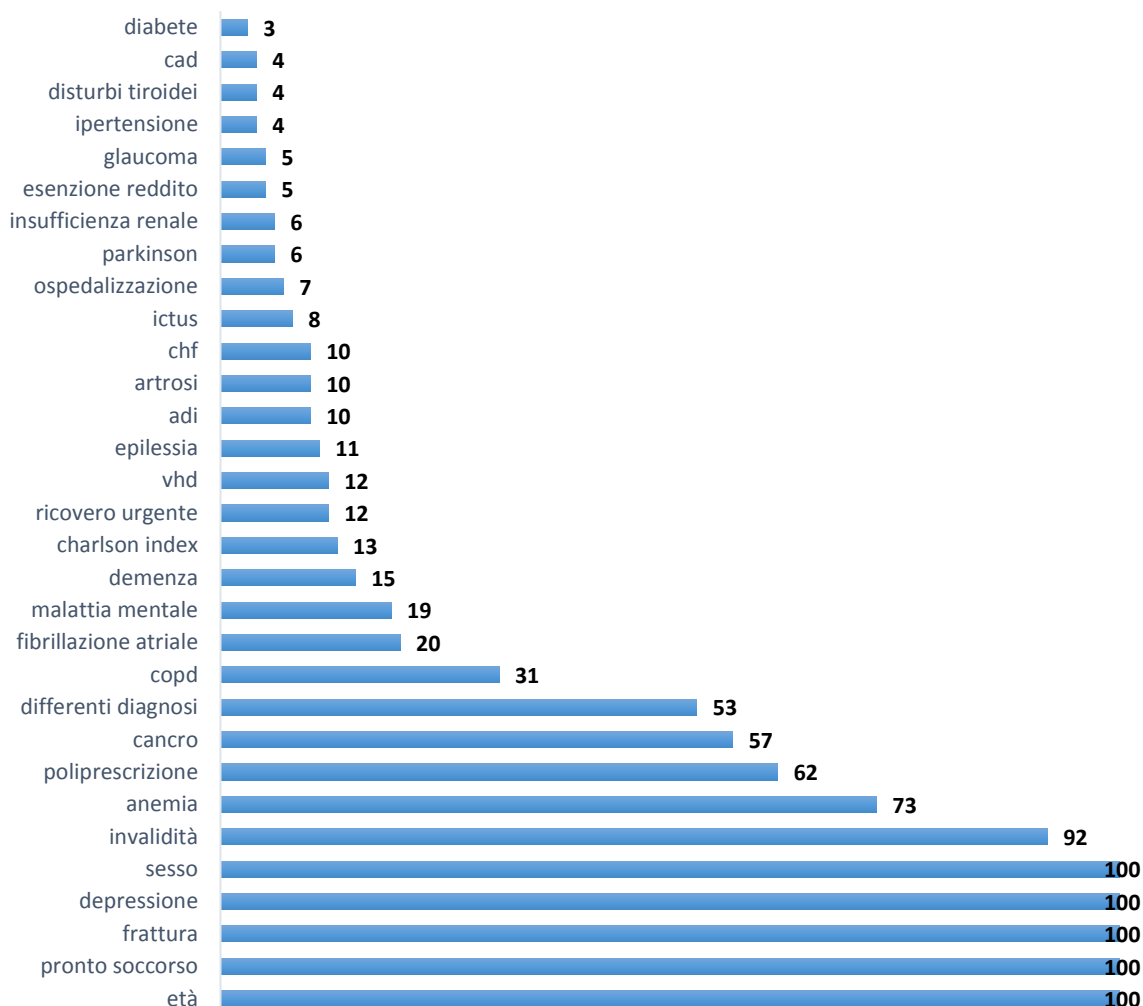


Figura 3.3: Frequenze d'ingresso delle variabili esplicative per l'outcome frattura.

3.4.4 Demenza

Per l'outcome demenza la selezione Stepwise si è confermata essere il predittore migliore, avendo un'auc mediamente pari a 0,826 e più alto nel 72% dei campioni. Le variabili selezionate sono ADI, essere affetto da un'altra malattia mentale, depressione, invalidità ed età. Anche in questo caso la presenza di una delle prime

4 variabili comporta un aumento del rischio di avere la demenza, e con il passare dell'età questo rischio aumenta esponenzialmente.

Frequenze d'ingresso nelle Stepwise: Demenza

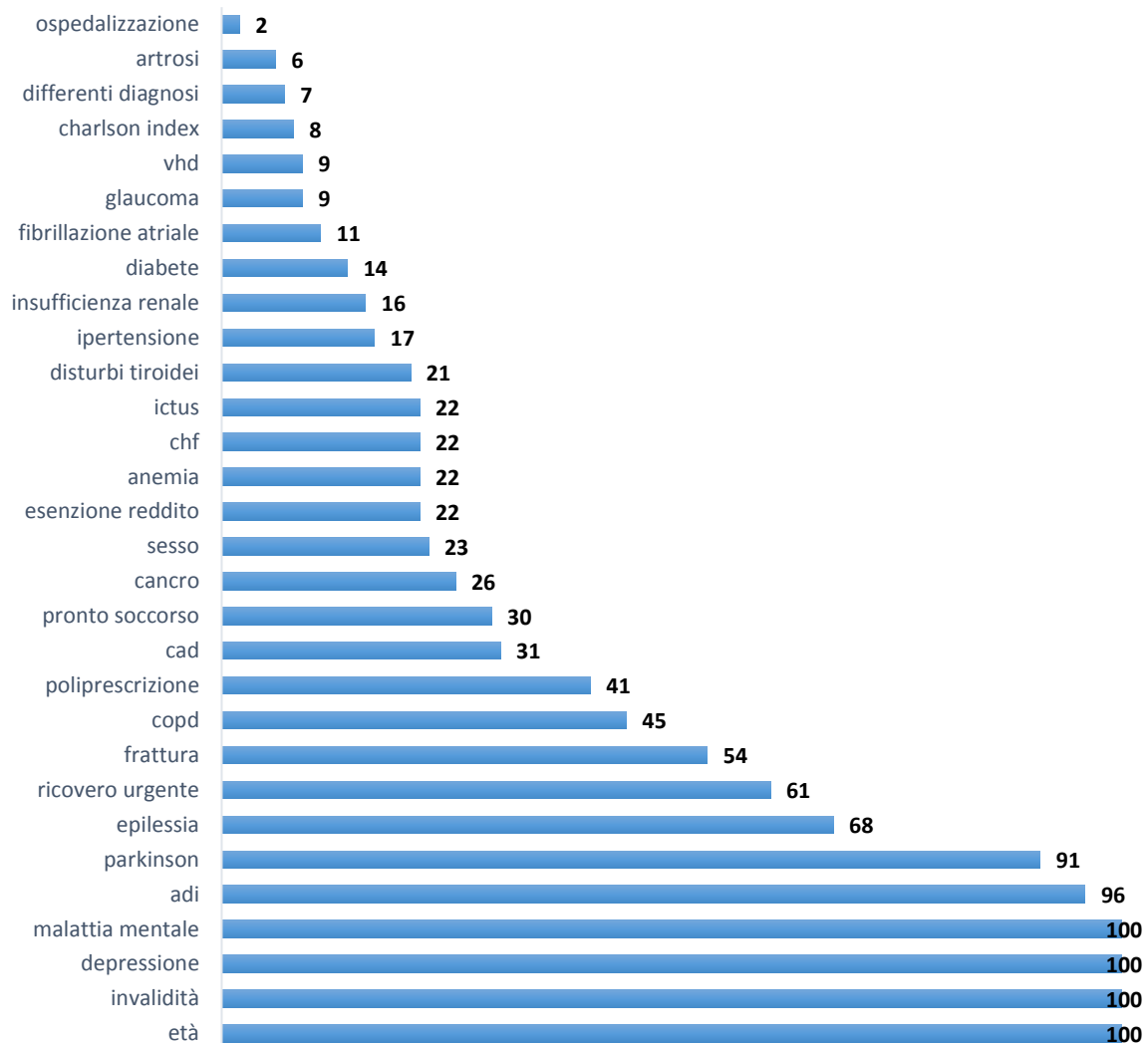


Figura 3.4: *Frequenze d'ingresso delle variabili esplicative per l'outcome demenza.*

3.4.5 Disabilità

Infine anche per questa variabile risposta è stata preferita la selezione Stepwise nell'82% dei casi. Le variabili selezionate dal classificatore sono state l'essere ricoverato con differenti diagnosi, la demenza, il cancro, il morbo di parkinson, la depressione, la poliprescrizione e l'età.

Anche in questo caso le variabili sono tutte effettivamente di rischio e non protettive, pertanto il sesso è rimasta l'unica variabile esclusa. In questo caso particolarmente rilevante è il fattore di rischio “demenza” che aumenta la probabilità di disabilità di undici volte.

Frequenze d'ingresso nelle Stepwise: Disabilità

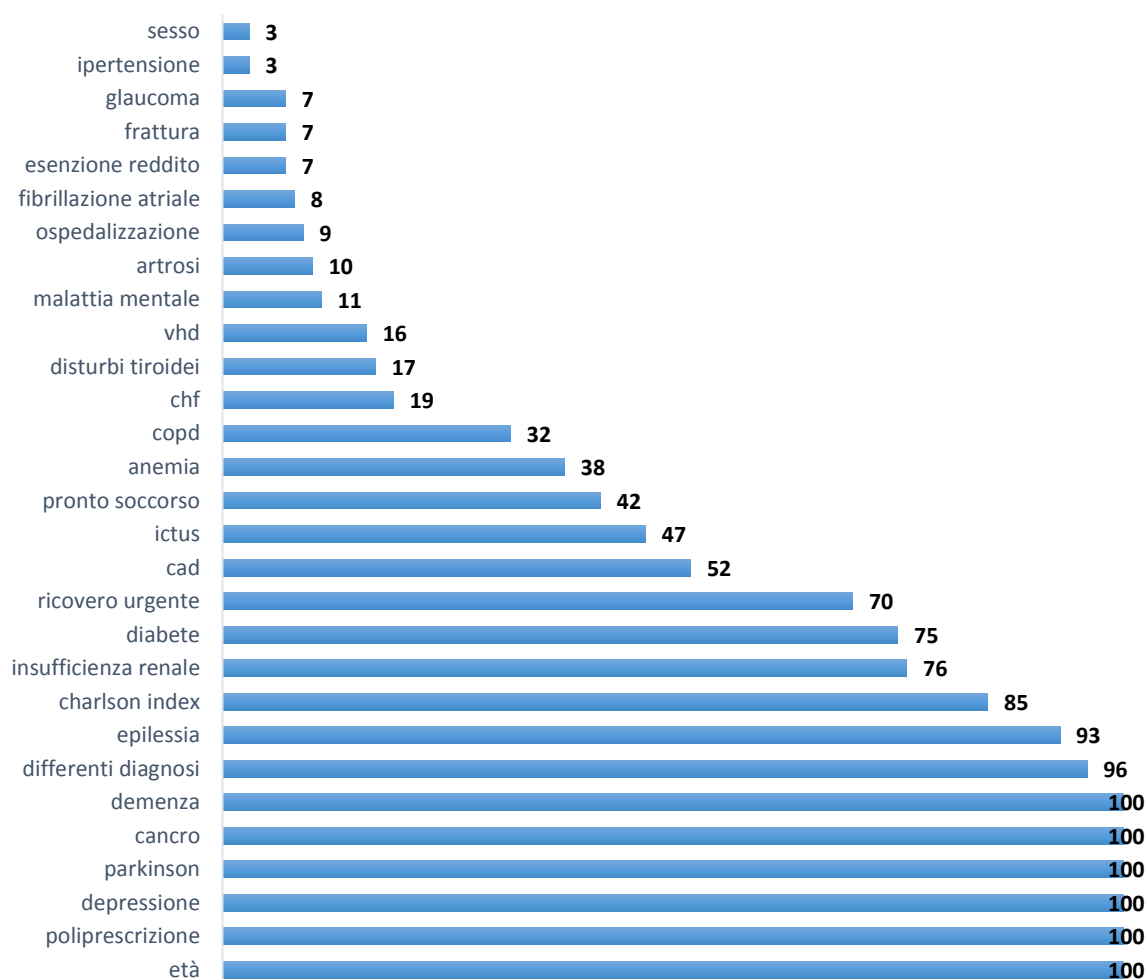


Figura 3.5: Frequenze d'ingresso delle variabili esplicative per l'outcome disabilità.

3.5 Criterio per la selezione delle variabili

Terminate le singole analisi, è stato necessario trovare un criterio che selezionasse le variabili esplicative più importanti a partire dalle distinte analisi appena descritte. Per ottenere ciò è stato scelto di prendere in considerazione per ogni singolo outcome le variabili che compaiono almeno nel 95% dei campioni, e in

seguito tenere solo quelle che in media sono state selezionate in più di 60 campioni nelle cinque diverse analisi. Di seguito sono rappresentate la media delle frequenze d'ingresso per tutte le variabili esplicative considerate:

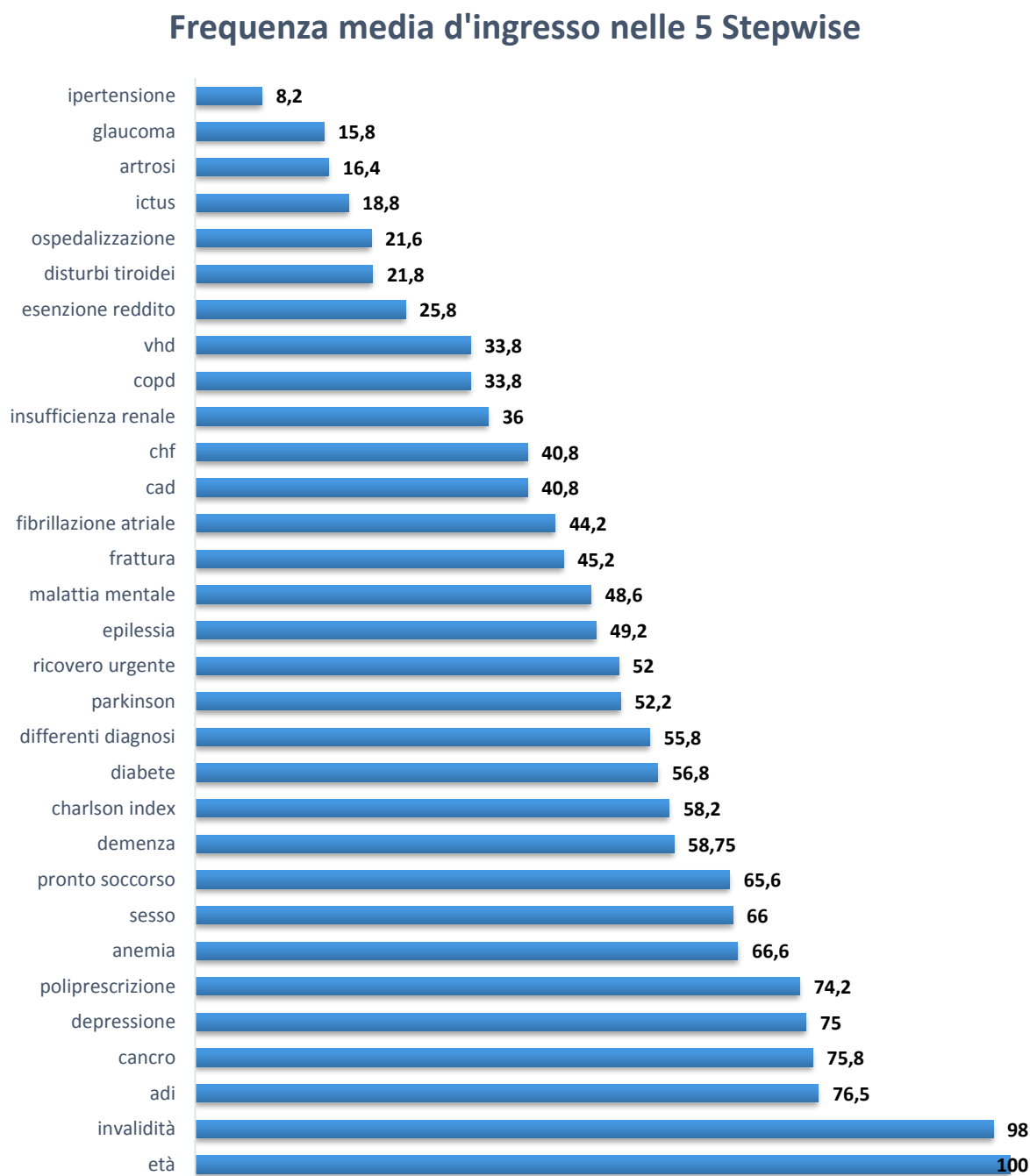


Figura 3.6: Frequenze medie d'ingresso delle variabili esplicative per i cinque outcomes.

Come si poteva intuire, le variabili sopra la soglia del 60% sono tutte covariate che in almeno una delle 5 analisi hanno superato il 95% di “presenza” nei campioni bilanciati. In questo modo si è selezionato un numero ristretto di variabili esplicative, tenendo conto di tutti i modelli statistici implementati.

A partire quindi da età, invalidità, ADI, cancro, depressione, poliprescrizione, anemia e accessi al pronto soccorso si costruirà un poset, e si andrà a valutare la sua capacità previsiva per ogni outcome.

Di seguito sono presentati gli odds ratio delle 8 covariate selezionate per ogni outcome e la capacità predittiva di ogni modello logistico rispetto al dataset di verifica.

Per ottenere le tabelle 3.6 e 3.7 si sono utilizzati dei dataset di stima in cui le variabili esplicative sono rilevate al biennio 2012-2013 e le variabili risposta al 2014. I cinque dataset sono stati bilanciati tramite campionamento casuale semplice dei soggetti appartenenti alla classe più numerosa della variabile risposta. Ogni modello è stato poi validato nel rispettivo dataset di verifica, in cui le esplicative si riferiscono al biennio 2013-2014 e le variabili risposta al 2015.

Tabella 3.6: Odds ratio dei fattori di rischio rispetto ai 5 outcomes.

	Decesso	Ricovero urgente	Frattura	Demenza	Disabilità
Età 70-74 vs 65-69	n.s (0.995- 1.853)	1.260 (1.098- 1.445)	n.s (0.795- 1.257)	2.346 (1.411-3.9)	n.s (0.851- 1.567)
Età 75-79 vs 65-69	2.097 (1.557- 1.853)	1.673 (1.460- 1.918)	1.458 (1.156- 1.840)	5.795 (3.548- 9.465)	2.185 (1.627- 2.933)
Età 80-84 vs 65-69	3.873 (2.855- 5.253)	2.159 (1.873- 2.490)	1.771 (1.385- 2.264)	8.831 (5.398- 14.448)	3.392 (2.500- 4.601)
Età 85-89 vs 65-69	6.890 (5.047- 9.406)	3.145 (2.686- 3.683)	2.179 (1.658- 2.864)	13.052 (7.844- 21.718)	10.539 (7.393- 15.024)
Età 90+ vs 65- 69	8.891 (6.287- 12.575)	4.177 (3.433- 5.081)	2.025 (1.477- 2.778)	13.382 (7.478- 23.946)	26.338 (15.567- 44.307)
Invalidità 1 vs 0	1.949 (1.594- 2.383)	1.612 (1.441- 1.802)	1.400 (1.143- 1.715)	1.876 (1.388- 2.535)	
Adi 1 vs 0	2.499 (1.958- 3.190)	1.761 (1.522- 2.038)	n.s (0.902- 1.552)	2.168 (1.485- 3.166)	
Cancro 1 vs 0	2.405 (1.954- 2.959)	1.163 (1.038- 1.304)	n.s (0.880- 1.312)	0.646 (0.455- 0.915)	2.975 (2.349- 3.768)
Depressione 1 vs 0	n.s (0.878- 1.412)	n.s (0.903- 1.161)	1.449 (1.178- 1.784)	3.592 (2.581- 5.001)	2.439 (1.849- 3.217)
Anemia 1 vs 0	3.090 (2.160- 4.420)	1.932 (1.586- 2.353)	n.s (0.949- 1.849)	n.s (0.795- 2.415)	n.s (0.872- 2.175)
Poliprescrizione 6-10 vs 0-5	n.s. (0.775- 1.157)	1.332 (1.206- 1.471)	n.s (0.899- 1.268)	n.s (0.675- 1.232)	1.546 (1.250- 1.911)
Poliprescrizione 11+ vs 0-5	1.475 (1.161- 1.875)	2.281 (2.011- 2.588)	1.288 (1.026- 1.617)	0.642 (0.437- 0.944)	3.223 (2.446- 4.247)
Pronto soccorso 1 vs 0	1.723 (1.391- 2.134)	1.532 (1.369- 1.714)	1.399 (1.153- 1.698)	1.500 (1.079- 2.085)	1.811 (1.414- 2.319)
Pronto soccorso 2-3 vs 0	2.122 (1.563- 2.881)	1.887 (1.595- 2.233)	1.577 (1.184- 2.101)	2.103 (1.298- 3.408)	2.022 (1.374- 2.976)
Pronto soccorso 4 vs 0	2.224 (1.213- 4.079)	5.024 (3.129- 8.066)	4.422 (1.931- 10.122)	3.934 (1.193- 12.967)	n.s (0.799- 5.758)

Tabella 3.7: Auc calcolato per le cinque diverse logistiche rispetto al dataset di verifica.

	Decesso	Ricovero urgente	Frattura	Demenza	Disabilità
AUC	0.844	0.747	0.641	0.823	0.815

Dalla lettura dei risultati emerge come il rischio di sperimentare qualsiasi outcome sia significativo soprattutto dopo i 75 anni rispetto alla categoria base, e che l'età ha un forte impatto soprattutto su demenza e disabilità. Le proxy delle disabilità sono significative per tutti gli outcomes, mentre la depressione ha un effetto molto forte sulla demenza, con un rischio più che triplicato. L'anemia è un fattore di rischio soprattutto per il decesso, mentre la poliprescrizione ha un effetto importante su tutti gli outcomes solo per i soggetti più gravi (più di 11 farmaci distinti assunti durante l'anno). Infine il pronto soccorso è un fattore di rischio omogeneo per tutti gli outcomes.

Le auc risultano soddisfacenti, tranne nel caso della frattura, in cui è ipotizzabile che "soffra" della mancanza di variabili legate alla mobilità della persona o a malattie ossee come l'osteoporosi.

Capitolo 4

Gli ordinamenti parziali: un metodo per la valutazione della fragilità negli anziani

La fragilità è un concetto complesso e multidimensionale, che è caratterizzato sia da una grande varietà di fattori di rischio, sia da molteplici esiti negativi. Nel capitolo precedente questo carattere multidimensionale è stato affrontato tramite distinti modelli statistici, ciascuno volto a spiegare un particolare aspetto della fragilità negli anziani. I risultati ottenuti in termini predittivi sono risultati quasi sempre soddisfacenti, ma così facendo non è stato possibile fornire un unico output predittivo per l'azienda ospedaliera, che tenga conto di tutti i possibili esiti negativi della fragilità.

Generalmente, nelle tante fonti citate nel primo capitolo, il problema viene risolto tramite l'utilizzo di indici compositi, costruiti in modo da riferirsi a diverse dimensioni del fenomeno. Tuttavia l'utilizzo di questi indici introduce soggettività all'analisi aggregando e sintetizzando informazioni diverse in unico indice tramite un sistema di pesi non oggettivi (Fattore, 2013). Una soluzione al problema è un algoritmo che permette di ordinare i soggetti considerando le loro caratteristiche tramite la matrice dei dati senza introdurre alcuna forma di soggettività: il "Partially Ordered Set" o più semplicemente poset, è un metodo che permette di ordinare i soggetti in modo oggettivo, basandosi sul concetto teorico di sequenza o disposizione di oggetti.

In questo capitolo verranno descritte le basi teoriche dell'algoritmo, le approssimazioni proposte da Bruggemann e De Loof per il calcolo dell'average rank, e infine la sua applicazione con i dati dell'Ulss 15, in modo da creare un ordinamento dei soggetti analizzati e studiarne le caratteristiche in base al loro "grado" di fragilità.

4.1 Definizione di Poset

Una relazione binaria \leq_P tra due oggetti in un *set* P è chiamata relazione d'ordine se rispetta le seguenti proprietà:

- Riflessività: se $x \in P$: $x \leq x$, ogni elemento è confrontabile con sé stesso;
- Antisimmetria: se $x \leq y$ e $y \geq x$ implica che $x=y$;
- Transitività: se $x \leq y$ e $y \leq z$ allora $x \leq z$.

Dato un *set* P , ogni oggetto $x \in P$ è descritto da un vettore di attributi $q(x) = (q_1(x), q_2(x), q_3(x) \dots q_k(x))$, dove $q_i(x) \in Q_i$ per ogni $i \in \{1, \dots, k\}$. Ogni spazio Q_i è dotato di una relazione d'ordine lineare \leq_i , in modo che è possibile dire che se $q_i(x) \leq_i q_i(y)$ allora x è più piccolo o uguale all'oggetto y in riferimento al solo attributo i . Generalizzando il concetto, si può affermare che $x \leq_P y$ se $q_i(x) \leq_i q_i(y)$ per ogni $i \in \{1, \dots, k\}$. Se non vale né la relazione $x \leq_P y$ né la relazione $x \geq_P y$ allora x e y sono incomparabili, ovvero $x \parallel_P y$. Se x e y sono oggetti distinti ma $q(x) = q(y)$, cioè $q_i(x) = q_i(y)$ per ogni i , allora gli oggetti x e y sono detti equivalenti e la relazione di equivalenza ($x \cong y$) è l'uguaglianza. Con *set* molto grandi generalmente si sceglie arbitrariamente un oggetto tra gli oggetti della classe di equivalenza, e si attuano in seguito i confronti con gli altri oggetti.

Una coppia (P, \leq_P) , dove P è un *set* di oggetti e \leq_P è una relazione d'ordine in P , è chiamato un insieme parzialmente ordinato, in breve poset (De Loof, 2011).

4.1.1 Un esempio con i dati sanitari

Per tornare ai dati presentati nei capitoli precedenti, si mostra un esempio applicativo di come il poset sia utile nel classificare gli individui in base alle loro caratteristiche: nella tabella 4.1 sono elencati 6 soggetti con 3 caratteristiche:

Tabella 4.1: Un esempio di matrice dei dati per la classificazione della fragilità negli anziani.

Individuo	Variabile		
	Invaldit�	Assistenza domiciliare	Depressione
A	0	0	0
B	1	0	0
C	0	0	1
D	1	0	1
E	1	1	0
F	1	1	1

Le tre variabili rappresentano fattori di rischio per gli outcomes della fragilit , quindi il valore 1 (presenza del fattore di rischio)   una situazione peggiore rispetto al valore 0 (assenza del fattore di rischio).   facilmente intuibile che in questo caso il soggetto A sia quello in condizioni migliori, cos  come il soggetto D sia in condizioni peggiori rispetto al soggetto B, perch  oltre ad aver in comune l'esenzione per invalidit , soffre anche di depressione. Allo stesso modo   facilmente intuibile che il soggetto F sia quello con le condizioni peggiori, avendo tutti i fattori di rischio. Gli individui D ed E sono un esempio di oggetti incomparabili in un set, poich    impossibile stabilire se, a parit  di esenzione per invalidit , sia peggio avere l'assistenza domiciliare o la condizione di depressione. Una soluzione potrebbe essere assegnare dei pesi diversi alle variabili stabilendo un ordine di "gravit " dei fattori di rischio, ma cos  facendo si introdurrebbe soggettivit  all'analisi. Come si vedr  in seguito, grazie alle metodologie introdotte da Bruggemann (2011) e da De Loof (2011),   possibile invece attuare dei confronti basandosi unicamente sulla matrice dei dati.

4.1.2 Definizioni di base

1. Se tutti gli n oggetti $x \in X$ sono comparabili tra loro, ovvero possono essere ordinati in modo che risulti $x_1 < x_2 < \dots < x_n$, allora si parla di **ordinamento completo**, totale o lineare.

2. L'insieme di variabili o attributi presenti nella matrice dei dati è detta **base informativa**. Due oggetti risultano comparabili, $x \perp y$, o incomparabili, $x \parallel y$, a seconda della base informativa.
3. Se esiste un sottoinsieme $X' \in X$ tale che per ogni coppia di oggetti $(x, y) \in X' \times X'$ si riesce a stabilire un ordinamento completo, allora questo sottoinsieme è detto **catena**.
4. Se esiste un sottoinsieme $X' \in X$ tale per cui nessuna coppia $(x, y) \in X' \times X'$ risulta comparabile ($x \perp y$), allora tale sottoinsieme è detto **anticatena**.
5. Quando tutti gli elementi di un insieme X formano una catena, allora l'insieme di tali oggetti è detto **ordine lineare**.
6. Quando una catena è composta da oggetti appartenenti a una stessa classe di equivalenza, allora si parla di **ordine debole**.
7. In matrici di dati finite, le catene e le anticate ne contengono un numero finito di oggetti: il numero di elementi definisce la **lunghezza** delle catene o delle anticate ne.
8. Il numero di elementi della catena più lunga di un poset è detta **altezza** del poset. Al contrario il numero di oggetti dell'anticatena più lunga è chiamata **larghezza** del poset.
9. Un elemento $x \in X$ per cui non vale la relazione \leq con alcun elemento y , è detto elemento **massimale**. Al contrario, un elemento $x \in X$ per cui non vale la relazione \geq con alcun elemento y , è detto elemento **minimale**.
10. Si definisce un elemento **massimo/minimo** se nel poset è presente un solo massimale/minimale.
11. Un elemento $x \in X$ si dice **isolato** se è allo stesso tempo massimale e minimale. Un elemento è massimale o minimale solo se non è isolato.
12. Si definisce una **relazione di copertura** tra y e x quando, dato $x < y$, non esiste un elemento z tale per cui $x < z$ e $z < y$: in questo caso si dice che **y copre x**, e si scrive $x \leq y$.
13. Dato un elemento $x \in X$ si definisce **Down Set** l'insieme $O(x) = \{y \in X : y \leq x\}$, con rispettivo **Successor Set** $S(x) = O(x) - \{x\}$.

14. Dato un elemento $x \in X$ si definisce **Up Set** l'insieme $F(x) = \{y \in X : y \geq x\}$, con rispettivo **Predecessor Set** $P(x) = F(x) - \{x\}$.
15. Dato l'insieme degli oggetti **incomparabili** per un elemento $x \in X$, definito come $U(x) = \{y \in X : y \parallel x\}$, l'insieme di tutti gli elementi **comparabili** con x è definito come $C(x) = F(x) \cup O(x)$.

4.2 Diagramma di Hasse

Una rappresentazione grafica utile per i poset è stata proposta dal matematico tedesco Helmut Hasse: l'idea di base è che gli oggetti siano dei nodi orientati a seconda delle relazioni d'ordine che caratterizzano il poset X , in modo che, dato un certo elemento $x \in X$, sotto di esso si posizioni il suo Down Set, e viceversa sopra di esso il suo Up set. Le linee che collegano i vari nodi indicano una relazione di copertura tra essi.

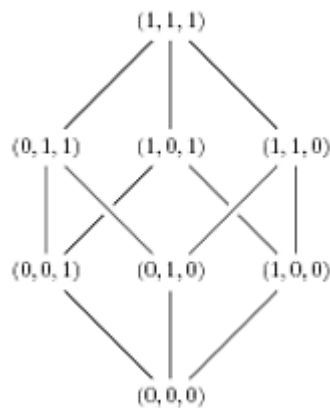


Figura 4.1: Esempio di diagramma di Hasse per la rappresentazione di un poset

4.3 Estensioni lineari

Un'estensione lineare di un poset è un ordinamento completo che è compatibile con le relazioni d'ordine del poset stesso. Dato n il numero di oggetti in un poset, ogni poset ha LT estensioni lineari (con $LT < n!$), dovuto al fatto che un elemento z incomparabile con x può essere compreso sia nel Down Set sia nell'Up set, senza violare alcuna condizione d'ordine.

Utilizzando i dati della tabella 4.1, queste sono le cinque possibili estensioni lineari:

Tabella 4.2: Estensioni lineari in riferimento alla matrice di dati della tabella 4.1.

F	F	F	F	F
D	E	D	E	D
E	D	E	D	C
B	B	C	C	E
C	C	B	B	B
A	A	A	A	A

Per definire un ordinamento tra i sei soggetti è possibile calcolare l'average rank tramite la seguente formula:

$$avr(x) = \frac{\sum_{k=1}^n H(x, le(k))}{LT}$$

Dove $H(x, le(k))$ è l'altezza dell'oggetto x nella k -esima estensione lineare $le(k)$.

Gli average rank per i sei individui nella tabella 4.2 sono i seguenti:

F	D	E	C	B	A
6	4.6	4.2	2.8	2.4	1

In questo modo si è stabilita una classifica tra gli individui, basandosi solo sui dati a disposizione, senza vincolarli a un modello parametrico o assegnando dei pesi in maniera soggettiva.

4.4 Le approssimazioni dell'average rank

L'average rank è calcolato in modo esatto quando i dati lo permettono, ovvero quando le osservazioni sono poche, così come gli oggetti incomparabili. Nel contesto di questo elaborato, in cui si hanno a che fare con circa 50000 osservazioni, è computazionalmente impossibile calcolare l'average rank esatto. Nei successivi paragrafi verranno descritte tre possibili approssimazioni di questa misura.

4.4.1 LPOM: Local Partial Order Model

Il Local Partial Order Model, sviluppato da Bruggemann (2004), permette di ottenere una rapida approssimazione dell'average rank di un elemento del poset, sfruttando alcuni concetti fin qui già visti. Il procedimento sottostante a questa approssimazione dell'average rank consiste nel trattare l'insieme degli elementi incomparabili $U(x)$ come elementi isolati, mentre gli elementi del Successor Set e quelli del Predecessor Set vanno a formare un ordinamento lineare con l'oggetto x . Considerando che gli oggetti in $U(x)$ possono inserirsi in qualsiasi posizione dell'ordinamento lineare, si ottiene la stima dell'average rank tramite la seguente formula:

$$\widehat{Rkav}(x) = \frac{(\|S(x)\| + 1)(n + 1)}{n + 1 - \|U(x)\|}$$

Questo metodo funziona bene quando sono poche le osservazioni e di conseguenza gli oggetti incomparabili. Infatti il punto critico di questo metodo è il considerare isolati tutti gli elementi incomparabili: così facendo non si tiene conto che gli oggetti incomparabili possono formare delle catene con il successor e il predecessor set, non avendo così accesso a tutte le posizioni a cui hanno accesso da isolati. Inoltre non si tiene conto delle relazioni tra gli oggetti incomparabili, che possono formare tra loro delle catene di cui il metodo LPOM non tiene conto. Pertanto si tengono conto di relazioni d'ordine illecite e che provocano quindi distorsioni sistematiche della stima dell'average rank. Una soluzione è stata fornita dallo stesso Bruggemann (2011), che ha modificato il modello per ottenere risultati migliori, pur in parte sacrificando il risparmio computazionale ottenuto col metodo appena descritto.

4.4.2 LPOMext: Extended Linear Partial Order Model

Il miglioramento rispetto al modello precedente è dato dal fatto che gli elementi incomparabili di un oggetto x non vengono più trattati come elementi isolati, ma si tiene conto delle relazioni d'ordine con Up Set e Down Set. Dato $y \in U(x)$, si definiscono i due insiemi $P_y^< = O(x) \cap U(y)$ e $P_y^> = F(x) \cap U(y)$ che rappresentano

l'intersezione tra l'insieme degli elementi incomparabili di y rispettivamente con Down Set ed Up Set di x .

La formula per il calcolo dell'average rank è:

$$\hat{H}(x) = \|O(x)\| + \sum_{y \in U(x)} \frac{\|P_y^<\|}{\|P_y^<\| + \|P_y^>\|}$$

Il numero di elementi del Down Set di x stabilisce l'altezza minima da cui parte l'oggetto x , mentre la seconda parte dell'equazioni rappresenta la probabilità che un oggetto incomparabile a x sia inferiore ad esso, aumentando così l'average rank di x .

4.4.3 L'approccio "Mutual Probabilities"

Nel 2011 De Loof propose un nuovo metodo per l'approssimazione dell'average rank basato sui seguenti concetti:

- La "Rank Probability" ovvero $p(x) = \sum_{i=1}^n i \cdot P(rank(x) = i)$, definita come il rapporto tra le estensioni lineare in cui il rango di x è uguale a i e il numero di possibili estensioni lineari.
- La "Mutual Rank Probability" ovvero la probabilità $P(x > y)$ definita come la frazione di estensioni lineari in cui x è più grande di y .

Il teorema alla base del metodo proposto da De Loof si struttura dalla relazione tra questi due concetti nella relazione:

$$\hat{p}(x) = \sum_{i=1}^n i \cdot P(rank(x) = i) = 1 + \sum_{y \in X} P(x > y)$$

Sulla base di questa formula, l'average rank può essere scritto come:

$$\hat{p}(x) = 1 + \sum_{y \in X} P(x > y) = \|O(x)\| + 1 + \sum_{y \in U(x)} P(x > y)$$

Un' approssimazione della “mutual rank probability” è data da Bruggemann (2004):

$$\hat{P}(x > y) = \frac{(o(x) + 1)(f(y) + 1)}{(o(x) + 1)(f(y) + 1) + (f(x) + 1)(o(y) + 1)}$$

dove $o(x)$ indica il numero di elementi del Down Set di x mentre $f(x)$ il numero di elementi dell'Up Set. Per discostarsi dall'approssimazione di Bruggemann e avere una stima più accurata, De Loof ha deciso di comprendere nella stima non solo gli elementi del Down Set e dell'Up Set ma anche quelli incomparabili. Pertanto tramite le seguenti approssimazioni:

$$\begin{aligned}\tilde{o}(x) &= o(x) + \sum_{z \in U(x)} \hat{P}(x > z) \\ \tilde{f}(x) &= f(x) + \sum_{z \in U(x)} \hat{P}(x < z)\end{aligned}$$

è stata ottenuta la stima finale dell'average rank:

$$\hat{p}(x) = o(x) + 1 + \sum_{y \in U(x)} \frac{(\tilde{o}(x) + 1)(\tilde{f}(y) + 1)}{(\tilde{o}(x) + 1)(\tilde{f}(y) + 1) + (\tilde{o}(y) + 1)(\tilde{f}(x) + 1)}$$

In pratica invece di contare il numero di elementi più piccolo di x , ovvero $o(x)$, $\tilde{o}(x)$ approssima la probabilità totale degli elementi di apparire prima di x in una estensione lineare, tramite il calcolo approssimato delle “mutual rank probabilities”.

Questo metodo ha lo stesso costo computazionale di LPOMext, ma De Loof ha dimostrato che tendenzialmente con numerosità più alte si ottengono risultati migliori in termini di errore quadratico medio.

Pertanto per le prossime elaborazioni è stato scelto questo metodo di approssimazione per il calcolo degli average rank degli anziani residenti nel territorio di competenza dell'Ulss 15.

4.5 Il poset come strumento operativo per l'Ulss 15

La fragilità è un concetto complesso da definire e in continua evoluzione: un soggetto non diventa fragile da un momento preciso della sua vita e di conseguenza non è semplice trovare un metodo operativo per stabilire se un individuo è fragile o meno.

Proprio per questa ragione il poset non si pone l'obiettivo di identificare una persona fragile, ma ordina gli individui studiati sulla base delle loro caratteristiche più importanti in questo contesto, stabilite nel capitolo 3. Questa è la ragione per cui alcuni outcomes (pensati come conseguenze della fragilità, e quindi come marcatori di essa) rientrano nelle caratteristiche studiate nel poset: è chiaro che una persona che soffre di demenza dovrebbe essere classificata fragile in base al fatto che la demenza è un marcatore di fragilità; ma l'Ulss, avendo delle risorse limitate, deve limitare i suoi piani assistenziali a un certo numero di individui, i più gravi. Il poset, proprio per la sua capacità di ordinare gli individui in modo oggettivo, basandosi solo sulle loro caratteristiche, è lo strumento più adatto a questo scopo.

4.6 Applicazione del poset sugli utenti anziani dell'Ulss 15: i risultati

Il metodo di De Loof (2011) è stato applicato agli utenti anziani dell'Ulss 15 in modo da stilare una "classifica" di fragilità basata sui fattori di rischio emersi dalle analisi descritte nel capitolo 3. La funzione per il calcolo dell'average rank approssimato è stata implementata in R (Caperna, 2015). Le variabili utilizzate sono l'età (in classi quinquennali), l'usufruire di assistenza domiciliare, l'esenzione per invalidità, l'essere malato di cancro, di depressione, di anemia, prendere distinti farmaci e il numero di accessi al pronto soccorso. Il numero limitato di variabili permette un tempo computazionalmente meno oneroso e risultati più affidabili: infatti con l'aumento di variabili aumenta la possibilità di avere più oggetti incomparabili, andando a complicare il processo di stima dell'average rank.

Nonostante nel capitolo 3 siano stati utilizzati dei dataset differenti per demenza e disabilità, poiché eventi irreversibili e in relazione solo a un sottoinsieme della

popolazione realmente a rischio, in questo caso è stato studiato un unico poset, comprese le persone disabili e malate di demenza. Ciò è dovuto al fatto che il poset non richiede una variabile risposta, e pertanto non nasce per prevedere un singolo evento quale la disabilità o il ricovero urgente, ma è una misura della fragilità latente basata sui fattori di rischio degli individui. Al contrario, le regressioni logistiche sono state utilizzate allo scopo di trovare dei fattori di rischio dei diversi outcomes, pertanto in quel caso è stato necessario utilizzare dataset ridotti per la demenza e la disabilità, in modo da studiare solo i soggetti effettivamente a rischio di sperimentare l'evento.

Un altro aspetto fondamentale da sottolineare è che le regressioni logistiche necessitano di un meccanismo di validazione, poiché i risultati sono vincolati a delle stime dei parametri che vanno valutate su osservazioni diverse da quelle utilizzate per la stima del modello; al contrario il poset stima l'average rank senza vincolarsi a dei parametri, pertanto è aggiornabile ogni anno con i nuovi dati raccolti.

Pur non essendo il poset uno strumento nato con l'esigenza di predire una variabile risposta, in questa analisi ci si pone l'obiettivo di fornire uno strumento concreto alle Ulss per poter pianificare delle campagne assistenziali nei territori di competenza. Pertanto, dopo aver standardizzato l'average rank, in modo da avere una misura continua tra 0 e 1, è stata valutata la sua capacità di predire i 5 differenti outcomes nell'anno successivo, in termini di area sotto la curva ROC. Nell'interesse di valutare la capacità predittiva del poset rispetto a quella della regressione logistica, in riferimento agli esiti negativi rilevati nel 2015, sono stati implementati tre differenti poset, le cui basi informative si riferiscono al biennio 2013-2014: il primo, contenente tutti gli utenti, è stato valutato per l'accuratezza nel predire distintamente il decesso, il ricovero urgente e la frattura nel 2015; il secondo è stato implementato per predire l'insorgere della demenza nel 2015, e si è basato solo sugli individui non affetti da demenza nel 2013-2014; il terzo poset è stato implementato per valutare la capacità di predire la disabilità nel 2015, basandosi solo su individui non disabili nel 2013 e nel 2014, e di conseguenza su individui che non hanno assistenza domiciliare o esenzione di invalidità in quei

due anni, che sono due delle otto variabili che costituiscono la base informativa del poset. Questi due ultimi poset sono stati implementati sugli stessi dataset ridotti alla base delle regressioni logistiche sulla demenza e sulla disabilità, in modo da rendere credibile il confronto fra i due approcci in termini di auc. Questo è particolarmente rilevante per l'outcome disabilità, che assume valore 1 quando l'utente ha l'assistenza domiciliare o l'esenzione di validità, due variabili che però fanno parte della base informativa del poset. Se, nel voler valutare la capacità del poset di predire la disabilità nel 2015, si fossero tenuti in considerazione anche gli individui disabili nel 2013-2014, è chiaro che a questi sarebbe stato assegnato tendenzialmente un average rank più alto, e pertanto sarebbe risultata una maggior capacità predittiva del poset rispetto alla regressione logistica per quanto riguarda l'outcome disabilità. Non considerando gli individui disabili (o malati di demenza nell'altro caso) si valuta la capacità del poset di predire l'insorgere della malattia, rendendolo confrontabile con la regressione logistica.

Questo però non deve confondere sull'obiettivo dell'analisi: le auc in questo caso hanno il solo scopo di effettuare il confronto tra poset e regressione logistica, e per rendere credibile tale confronto è necessario partire dagli stessi dataset. Una volta accertata la bontà di questo metodo, all'azienda sanitaria viene fornito un unico poset basato sugli otto fattori di rischio selezionati, con un unico vettore di average rank risultante, da cui ricavare un numero di individui più fragili (con average rank più alti).

Nella tabella 4.3 sono mostrati i risultati: per quanto riguarda la regressione logistica, le auc sono le stesse della tabella 3.7. Si ricorda che i dataset di stima, resi bilanciati tramite campionamento casuale semplice dei soggetti appartenenti alla classe più numerosa della variabile risposta, si riferiscono agli anni 2012-2014, in cui nei primi due sono rilevate le variabili esplicative, mentre nel 2014 le variabili risposta. Una volta stimati i parametri, la validazione dei modelli logistici è avvenuta basandosi su dataset in cui le variabili esplicative si riferiscono al biennio 2013-2014, e le variabili risposta all'anno 2015.

I poset non prevedendo parametri di stima, non hanno bisogno di un dataset di stima e uno di validazione: pertanto sono stati implementati i tre differenti poset

sulla base informativa relativa al biennio 2013-2014, ed è stata valutata la loro capacità predittiva sugli outcomes del 2015.

Tabella 4.3: Auc dei poset e delle regressioni logistiche per ogni outcome rilevato nel 2015.

	Outcome				
	Decesso	Ricovero	Disabilità	Frattura	Demenza
Logistica	0.844	0.747	0.815	0.641	0.823
Poset	0.835	0.743	0.807	0.637	0.784

La regressione logistica predice generalmente meglio dei poset, essendo un metodo che si prefigge l'obiettivo di predire adeguatamente una variabile risposta, ma le auc relative ai poset si avvicinano alle prestazioni delle regressioni logistiche e sono tutte soddisfacenti. Il vantaggio di utilizzare il poset risiede nel fatto che all'azienda si fornisce un unico algoritmo che ha una bontà predittiva paragonabile a quella di cinque regressioni logistiche.

Un'ulteriore conferma della bontà previsiva del poset è mostrata in tabella 4.4 e in figura 4.3, in cui si evidenzia la capacità del poset di predire l'evento decesso o l'evento ricovero urgente nel 2015, a partire dalla stima dell'average rank sui dati del 2013-2014.

Tabella 4.4: Frequenze di decessi o ricoveri urgenti nel 2015 in relazione alla stima dell'average rank standardizzato dal poset relativo al biennio 2013-2014.

Average rank	(Decesso o Ricovero urgente) = 0	(Decesso o Ricovero urgente) = 1
[0 - 0.1)	27186 (95%)	1437 (5%)
[0.1 - 0.2)	7791 (88,5%)	1010 (11,5%)
[0.2 - 0.3)	4175 (82,4%)	893 (17,6%)
[0.3 - 0.4)	2183 (76,4%)	673 (23,6%)
[0.4 - 0.5)	1436 (71%)	586 (29%)
[0.5 - 0.6)	919 (65,1%)	492 (34,9%)
[0.6 - 0.7)	554 (58,4%)	395 (41,6%)
[0.7 - 0.8)	296 (52,1%)	272 (47,9%)
[0.8 - 0.9)	118 (36,9%)	202 (63,1%)
[0.9 - 1]	37 (32,2%)	78 (67,8%)

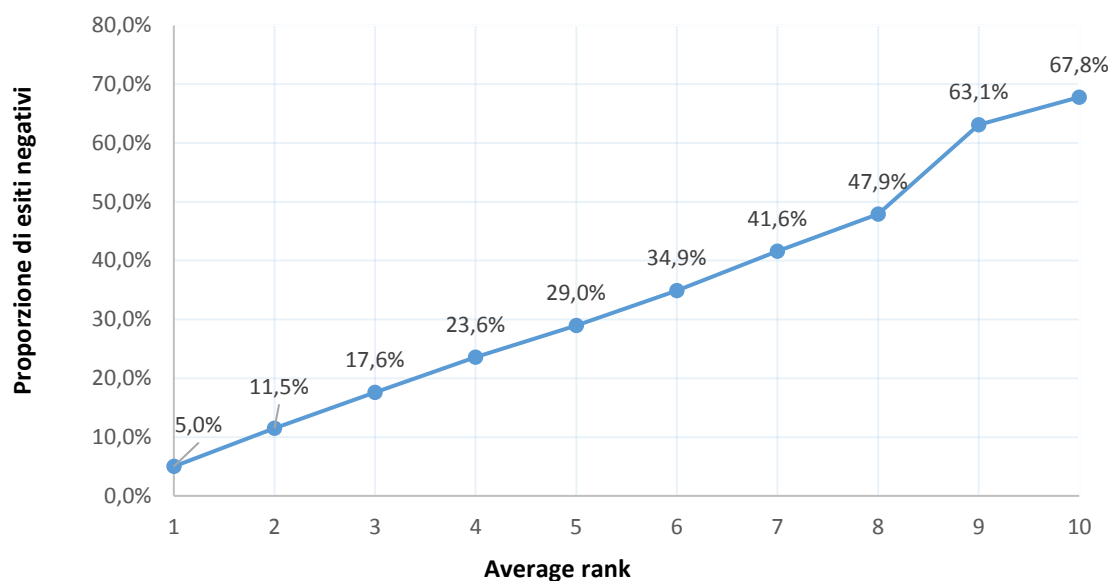


Figura 4.3: Proporzione di decessi o ricoveri urgenti nel 2015 in relazione alla stima dell'average rank standardizzato dal poset relativo al biennio 2013-2014.

Come si può notare nella fascia di soggetti meno fragili, in cui l'average rank (calcolato sul biennio 2013-2014) è compreso tra 0 e 0,3, si arriva a un massimo del 17,6% del totale che sperimenta nell'anno 2015 il decesso o il ricovero urgente; nella fascia intermedia di fragilità, tra 0,3 e 0,7, la percentuale varia tra 23,6% e 41,6% di anziani deceduti o ricoverati d'urgenza nel 2015; infine nella fascia di fragilità più alta si verifica dal 47,9% al 67,8% di anziani che sperimentano almeno uno dei due eventi considerati.

Chiaramente quest'ultima fascia di individui è quella più indicata agli interventi assistenziali, pertanto vengono di seguito proposte le caratteristiche di questa parte di popolazione anziana, caratterizzata da un average rank superiore a 0,7, per un totale di 1003 utenti. La distribuzione dell'average rank con relativa soglia che divide gli individui più fragili dai restanti è mostrata in figura 4.4.

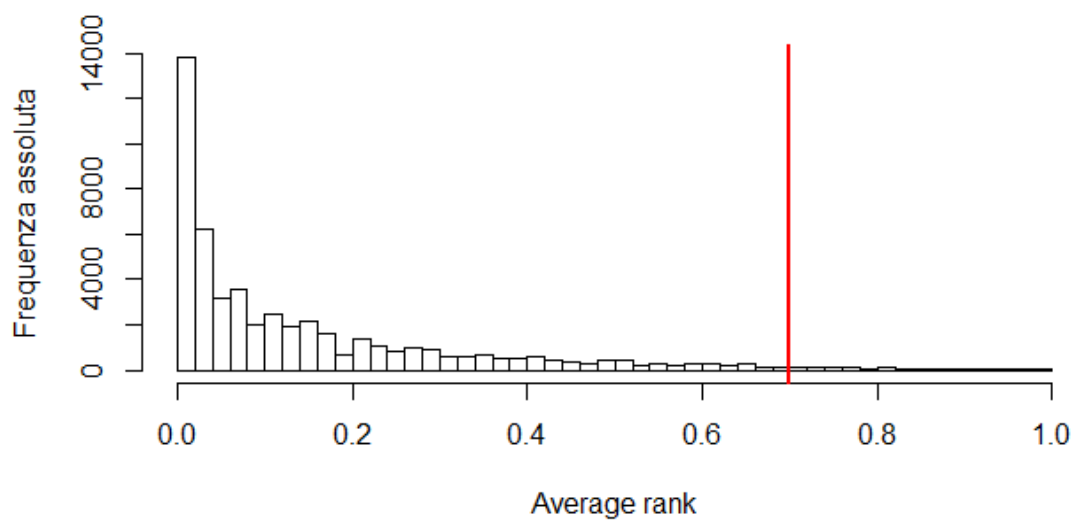


Figura 4.4: Distribuzione di frequenza dell'average rank in tutta la popolazione anziana. In rosso la soglia sul valore 0,7 dell'average rank.

Nella figura 4.5 si mostra come si distribuiscono i soggetti nelle varie classi di alcuni dei fattori di rischio emersi nelle analisi statistiche del capitolo 3.

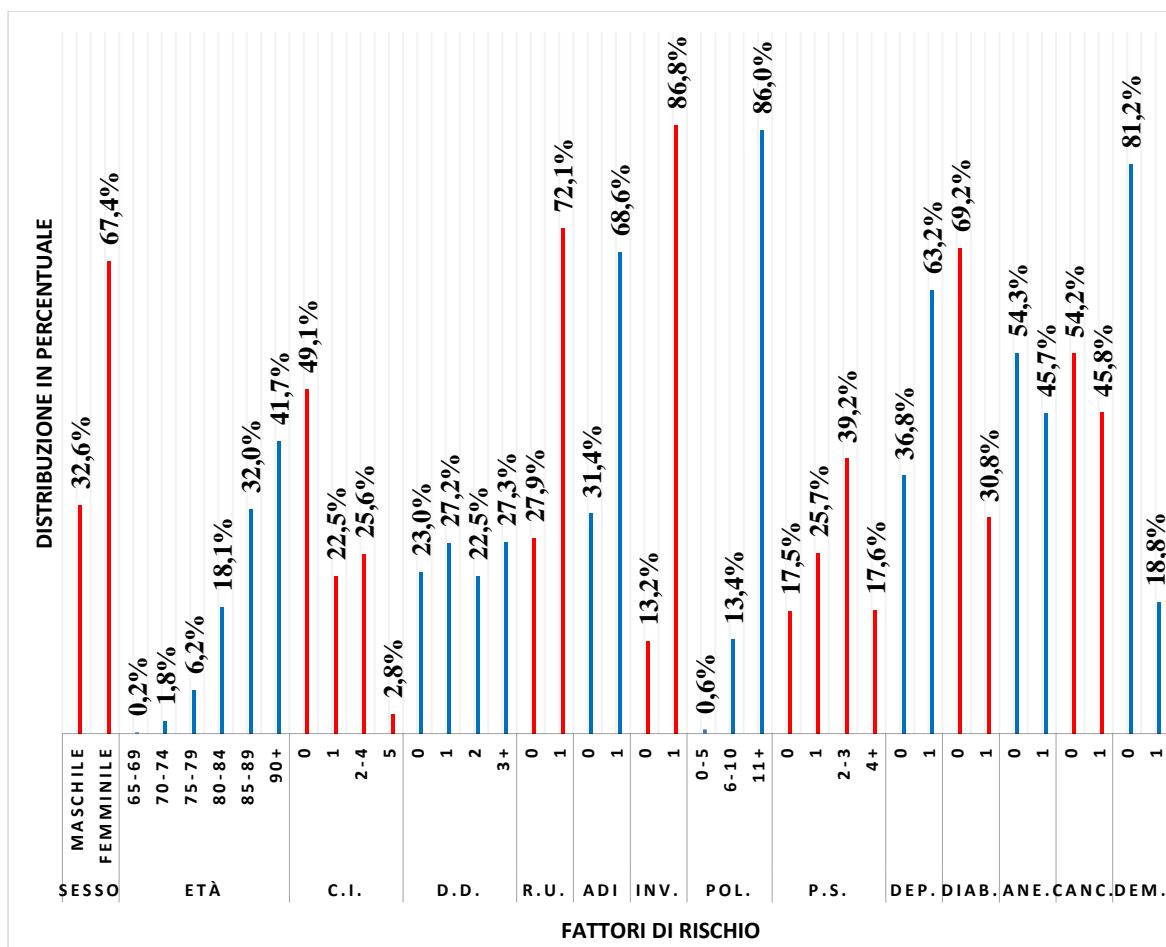


Figura 4.5: Distribuzione in percentuale dei soggetti con alto livello di fragilità (average rank superiore a 0.7) per alcuni fattori di rischio.

Legenda: C.I. Charlson Index; D.D. Numero di differenti diagnosi; R.U. Ricovero urgente; Inv. Invalidità; Pol. Poliprescrizione di farmaci; P.S. Pronto soccorso; Dep. Depressione; Diab. Diabete; Ane. Anemia; Canc. Cancro; Dem. Demenza.

Si nota che gli individui più fragili sono donne (67,4%), con un'età superiore agli 80 anni (91,8%), prevalentemente invalidi (86,8%), con almeno un ricovero urgente nell'ultimo anno (72,1%), che usufruiscono di assistenza domiciliare (68,6%) e che assumono più di 10 farmaci differenti (86%). Inoltre l'82,5% ha avuto almeno un ricovero al pronto soccorso tramite il 118, mentre non emerge una modalità prevalente nel numero di differenti diagnosi e quasi un quinto dei più fragili non ha avuto più ricoveri con differenti diagnosi principali. Un'altra patologia prevalente tra gli anziani fragili è la depressione (63,2%), mentre la patologia meno prevalente tra i fattori di rischio è la demenza (18,8%), che però ha anche una bassa prevalenza in tutta la popolazione anziana, ovvero il 2,27%.

Tra gli individui più fragili con meno di 80 anni (8,2%) vi sono soprattutto pazienti con almeno due ricoveri urgenti con differenti diagnosi (71,1%), con esenzione di invalidità (81,7%), con almeno un ricovero in pronto soccorso (95,8%) e malati di cancro (67,6%). I soggetti con il più alto indice di Charlson Index sono il 2,8% e tra questi circa il 65% muore o viene ricoverato urgentemente nel 2015. Infine i pochi soggetti che, nonostante figurino tra i più fragili, prendono meno di 6 farmaci distinti (lo 0,6%), sono in maggioranza pazienti invalidi (92%), malati di cancro (69,2%) e di depressione (69,2%).

Conclusioni

La fragilità è un fenomeno molto attuale nel mondo di oggi, soprattutto nei paesi sviluppati che denotano un invecchiamento progressivo della popolazione. Nonostante ciò, la fragilità è ancora un concetto complesso da definire, sia a livello concettuale che operativo.

Sul piano concettuale, un punto di vista comune spiega la fragilità come una condizione che coinvolge molteplici campi della vita della persona (psicologica, sociale, fisica), legata a un peggioramento della salute complessiva, a una perdita di autonomia nelle funzioni principali della vita quotidiana e a una vulnerabilità maggiore a eventi avversi quali il decesso o il ricovero d'urgenza.

Dal punto di vista operativo, molti studi hanno analizzato il fenomeno proponendo degli indici compositi basati su scale di valutazione qualitative della fragilità, validandoli in seguito tramite regressioni logistiche su una delle conseguenze della fragilità, spesso individuata nel decesso.

Il presente elaborato è nato dall'idea di fornire uno strumento pratico ed affidabile all'Ulss 15, per quantificare la fragilità negli utenti anziani ordinandoli tramite un criterio oggettivo, basato unicamente sulla matrice dei dati. Per ottenere ciò, è stata svolta un'accurata e attenta analisi della letteratura, per trovare i fattori di rischio ma soprattutto i diversi marcatori della fragilità: in tal modo si è voluta fare un'analisi quanto più completa ed esaustiva possibile, per dare più accuratezza e affidabilità ai risultati finali.

La grande mole di dati a disposizione (SDO, flusso farmaceutico, etc.) ha permesso di studiare cinque outcomes (decesso, ricovero urgente, frattura, demenza e disabilità) e più di 30 fattori di rischio emersi dalla ricerca bibliografica. Per farlo sono state svolte cinque analisi sui distinti outcomes, volte a selezionare un set di variabili esplicative con maggiore capacità predittiva tra le 31 considerate inizialmente. Data la necessità di utilizzare campioni equilibrati, è stato utilizzato il principio del campionamento ripetuto per selezionare le variabili che più di frequente sono state selezionate dalla procedura Stepwise.

Con un approccio euristico, sono state selezionate 8 variabili esplicative, tenendo conto di tutti i diversi outcomes. Queste hanno rappresentato la base informativa per il poset, un algoritmo che permette l'ordinamento dei soggetti tramite la stima del livello di fragilità (average rank).

Dopo aver standardizzato l'average rank ottenuto, è stata valutata la sua capacità predittiva per ogni outcome in termini di area sotto la curva ROC, e si è visto che i risultati sono in linea con quelli delle regressioni logistiche. A parità di capacità predittiva il poset ha indubbiamente dei vantaggi:

- Fornisce come output un unico vettore di average rank, pur tenendo conto dell'aspetto multivariato della fragilità.
- Nonostante non sia vincolato a una variabile dipendente, è in grado di prevedere accuratamente alcune conseguenze della fragilità, come il decesso, il ricovero urgente, la frattura, la demenza e la disabilità.
- Non soffre di soggettività poiché si basa solo sulla matrice dei dati a disposizione.

Pertanto il poset è uno strumento pratico e semplice per l'azienda: è necessario solo inserire i dati degli utenti e implementare l'algoritmo. A quel punto, a seconda delle logiche aziendali, è possibile selezionare un certo numero di utenti con l'average rank più alto, e attuare delle campagne di assistenza sul territorio, al fine di ridurre il rischio di esiti negativi come il ricovero ospedaliero, e di conseguenza ridurre il potenziale costo sanitario dell'utente fragile.

Bibliografia

- Avila-Funes, J. A., Amieva, H., Barberger-Gateau, P., Le Goff, M., Raoux, N., Ritchie, K., ... & Dartigues, J. F. (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., ... & 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., Beland, F., Karunanathan, S., Hummel, S., Hogan, D., & Wolfson, C. (2004). *Developing a working framework for understanding frailty*. *Gérontologie et société*, 109, 15-29.
- Bergman, H., Ferrucci, L., Guralnik, J., Hogan, D. B., Hummel, S., Karunanathan, S., & Wolfson, C. (2007). *Frailty: an emerging research and clinical paradigm—issues and controversies*. *The Journals of Gerontology Series A: Biological Sciences and Medical Sciences*, 62(7), 731-737.
- Billings, J., Georghiou, T., Blunt, I., & 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.
- Boers, M. & Jentoft, A.C. (2015). *A New Concept of Health Can Improve the Definition of Frailty*. *Calcified Tissue International*, 97.5, 429-431.
- Brody, K. K., Johnson, R. E., & 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., & Perrin, N. (2002). *A Comparison of Two Methods for Identifying Frail Medicare-Aged Persons*. *Journal of the American Geriatrics Society*, 50(3), 562-569.

- Brüggemann, R., Sørensen, P. B., Lerche, D., & Carlsen, L. (2004). *Estimation of averaged ranks by a local partial order model*. Journal of chemical information and computer sciences, 44(2), 618-625.
- Bruggemann, R., & Carlsen, L. (2011). *An improved estimation of averaged ranks of partial orders*. Match Commun Math Comput Chem, 65(2), 383-414.
- Buchman, A. S., Boyle, P. A., Wilson, R. S., Tang, Y., & Bennett, D. A. (2007). *Frailty is associated with incident Alzheimer's disease and cognitive decline in the elderly*. Psychosomatic medicine, 69(5), 483-489.
- Caperna, G. (2016). *Partial order theory for synthetic indicators*, Tesi di Dottorato in Statistica, Università degli Studi di Padova.
- Carrière, I., Colvez, A., Favier, F., Jeandel, C., Blain, H., & EPIDOS study group. (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. & Malvi, C. (2014). *La fragilità degli anziani. Strategie, progetti, strumenti per invecchiare bene*. Santarcangelo di Romagna: Maggioli Editore.
- Consiglio Sanitario Regionale della Regione Toscana (2013). *La fragilità nell'anziano*. A cura di Regione Toscana.
- Crane, S. J., Tung, E. E., Hanson, G. J., Cha, S., Chaudhry, R., & 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., & Smith, F. (2004). *Risk factors for nonelective hospitalization in frail and older adult, inner-city outpatients*. The Gerontologist, 44(1), 68-75.
- De Loof, K., De Baets, B., & De Meyer, H. (2011). *Approximation of average ranks in posets*. Match Commun Math Comput Chem, 66, 219-229.
- De Vries, N. M., Staal, J. B., Van Ravensberg, C. D., Hobbelen, J. S. M., Rikkert, M. O., & Nijhuis-Van der Sanden, M. W. G. (2011). *Outcome instruments to measure frailty: a systematic review*. Ageing research reviews, 10(1), 104-114.

- De Vries, O. J., Peeters, G. M. E. E., Lips, P., & Deeg, D. J. H. (2013). *Does frailty predict increased risk of falls and fractures? A prospective population-based study*. *Osteoporosis international*, 24(9), 2397-2403.
- Deyo, R. A., Cherkin, D. C., & Ciol, M. A. (1992). *Adapting a clinical comorbidity index for use with ICD-9-CM administrative databases*. *Journal of clinical epidemiology*, 45(6), 613-619.
- Fabbris, L. (1997). *Statistica multivariata: analisi esplorativa dei dati*. McGraw-Hill Libri Italia.
- Falasca, P., Berardo, A. & 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?*. *Annali dell'Istituto Superiore di Sanità* 47.2, 220-228.
- Fattore, M. (2013). *La valutazione statistica del benessere: ordinamenti parziali e nuovi paradigmi metodologici*. *LAVORO, ECONOMIA, FINANZA*, 10.
- Fried, L. P., Tangen, C. M., Walston, J., Newman, A. B., Hirsch, C., Gottdiener, J., ... & McBurnie, M. A. (2001). *Frailty in older adults evidence for a phenotype*. *The Journals of Gerontology Series A: Biological Sciences and Medical Sciences*, 56(3), M146-M157.
- Fried, L. P., Ferrucci, L., Darer, J., Williamson, J. D. & Anderson, G. (2004). *Untangling the Concepts of Disability, Frailty, and Comorbidity: Implications for Improved Targeting and Care*. *Journal of Gerontology* 59.3, 255–263.
- Gobbens, R.J.J., Luijckx, K.G. & Wijnen-Sponselee, M.T. (2010). *In Search of an Integral Conceptual Definition of Frailty: Opinions of Experts*. *Journal of the American Medical Directors Association* 11.5, 338–343.
- Gray, S. L., Anderson, M. L., Hubbard, R. A., LaCroix, A., Crane, P. K., McCormick, W., ... & Larson, E. B. (2013). *Frailty and incident dementia*. *The Journals of Gerontology Series A: Biological Sciences and Medical Sciences*, glt013.
- Hippisley-Cox, J., & Coupland, C. (2013). *Predicting risk of emergency admission to hospital using primary care data: derivation and validation of QAdmissions score*. *BMJ open*, 3(8), e003482.

- Inouye, S. K., Zhang, Y., Jones, R. N., Shi, P., Cupples, L. A., Calderon, H. N., & 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.
- James, G., Witten, D., Hastie, T., & Tibshirani, R. (2013). *An introduction to statistical learning (Vol. 6)*. New York: springer.
- Jones, D. M., Song, X., & 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., & 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., & 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., & Bernabei, R. (2004). *Comorbidity and social factors predicted hospitalization in frail elderly patients*. *Journal of clinical epidemiology*, 57(8), 832-836.
- López-Aguilà, S., Contel, J. C., Farre, J., Campuzano, J. L., & Rajmil, L. (2011). *Predictive model for emergency hospital admission and 6-month readmission*. *The American journal of managed care*, 17(9), e348-57.
- Luppa, M., Luck, T., Weyerer, S., König, H. H., Brähler, E., & Riedel-Heller, S. G. (2009). *Prediction of institutionalization in the elderly. A systematic review*. *Age and ageing*, afp202.
- Makary, M. A., Segev, D. L., Pronovost, P. J., Syin, D., Bandeen-Roche, K., Patel, P., ... & Fried, L. P. (2010). *Frailty as a predictor of surgical outcomes in older patients*. *Journal of the American College of Surgeons*, 210(6), 901-908.
- Mazzaglia, G., Roti, L., Corsini, G., Colombini, A., Maciocco, G., Marchionni, N., ... & Bari, M. D. (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). *Flusso informativo delle prestazioni farmaceutiche effettuate in distribuzione diretta o per conto*. Roma.
- Ministero della Salute (2016). *Pronto Soccorso e sistema 118*. Roma.
- Ministero della Salute (2016). *Sistema informativo per il monitoraggio dell'assistenza domiciliare*. Roma.
- Ministero della Sanità (1998). *Classificazione delle malattie, dei traumatismi, degli interventi chirurgici e delle procedure diagnostiche e terapeutiche*. Istituto Poligrafico e Zecca dello Stato. Roma.
- Ministero della Sanità (2000). *Classificazione statistica internazionale delle malattie e dei problemi sanitari correlati*. Decima revisione, 3.
- Mitnitski, A. B., Graham, J. E., Mogilner, A. J., & Rockwood, K. (2002). *Frailty, fitness and late-life mortality in relation to chronological and biological age*. BMC geriatrics, 2(1), 1.
- Morley, J. E., Vellas, B., van Kan, G. A., Anker, S. D., Bauer, J. M., Bernabei, R., ... & Fried, L. P. (2013). *Frailty consensus: a call to action*. Journal of the American Medical Directors Association, 14(6), 392-397.
- Paw, M. J. C. A., Dekker, J. M., Feskens, E. J., Schouten, E. G., & Kromhout, D. (1999). *How to select a frail elderly population? A comparison of three working definitions*. Journal of clinical epidemiology, 52(11), 1015-1021.
- Puts, M. T. E., Lips, P. T. A. M., & Deeg, D. J. H. (2005). *Static and dynamic measures of frailty predicted decline in performance-based and self-reported physical functioning*. Journal of clinical epidemiology, 58(11), 1188-1198.
- Ravaglia, G., Forti, P., Lucicesare, A., Pisacane, N., Rietti, E., & Patterson, C. (2008). *Development of an easy prognostic score for frailty outcomes in the aged*. Age and Ageing, 37(2), 161-166.
- Rockwood, K., Fox, R.A., Stolee, P., Robertson, D. & Beattie, B.L. (1994). *Frailty in elderly people: an evolving concept*. Canadian Medical Association Journal 150, 489–495.
- Rockwood, K., Howlett, S. E., MacKnight, C., Beattie, B. L., Bergman, H., Hébert, R., ... & McDowell, I. (2004). *Prevalence, attributes, and outcomes of fitness and frailty in community-dwelling older adults: report from the*

- Canadian study of health and aging*. The Journals of Gerontology Series A: Biological Sciences and Medical Sciences, 59(12), 1310-1317.
- 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., & 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.
- Rockwood, K. & Mitnitski, A. (2007). *Frailty in Relation to the Accumulation of Deficits*. Journal of Gerontology 62A.7, 722–727.
- Rothman, M. D., Leo-Summers, L., & 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., ... & Wenger, N. S. (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.
- Searle, S.D., Mitnitski, A., Gahbauer, E. A., Gill, T.M. & Rockwood, K. (2008). *A standard procedure for creating a frailty index*. BMC Geriatrics, 8–24.
- Shelton, P., Sager, M. A., & 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. (2015). *Identificazione dei soggetti fragili nell'Ulss 15: una nuova proposta basata sugli ordinamenti parziali*, Tesi di Laurea Magistrale in Scienze Statistiche, Università degli Studi di Padova.
- Solfrizzi, V., Scafato, E., Frisardi, V., Seripa, D., Logroscino, G., Maggi, S., ... & Di Carlo, A. (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., & Rockwood, K. (2014). *Age-related deficit accumulation and the risk of late-life dementia*. Alzheimer's research & therapy, 6(5), 54.

- Sternberg, S. A., Schwartz, A. W., Karunanathan, S., Bergman, H., & Mark Clarfield, A. (2011). *The identification of frailty: a systematic literature review*. *Journal of the American Geriatrics Society*, 59(11), 2129-2138.
- Tammemagi, C. M., Neslund-Dudas, C., Simoff, M., & 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.
- Tom, S. E., Adachi, J. D., Anderson, F. A., Boonen, S., Chapurlat, R. D., Compston, J. E., ... & Nieves, J. W. (2013). *Frailty and fracture, disability, and falls: a multiple country study from the global longitudinal study of osteoporosis in women*. *Journal of the American Geriatrics Society*, 61(3), 327-334.
- Walker, L., Jamrozik, K., & Wingfield, D. (2005). *The Sherbrooke Questionnaire predicts use of emergency services*. *Age and ageing*, 34(3), 233-237.
- Wallace, E., Hinchey, T., Dimitrov, B. D., Bennett, K., Fahey, T., & Smith, S. M. (2013). *A Systematic Review of the Probability of Repeated Admission Score in Community-Dwelling Adults*. *Journal of the American Geriatrics Society*, 61(3), 357-364.
- Walston, J., Hadley, E. C., Ferrucci, L., Guralnik, J. M., Newman, A. B., Studenski, S. A., ... & Fried, L. P. (2006). *Research agenda for frailty in older adults: toward a better understanding of physiology and etiology: summary from the American Geriatrics Society/National Institute on Aging Research Conference on Frailty in Older Adults*. *Journal of the American Geriatrics Society*, 54(6), 991-1001.
- Fugate Woods, N., LaCroix, A. Z., Gray, S. L., Aragaki, A., Cochrane, B. B., Brunner, R. L., ... & 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.
- Zucchelli, G. & Silvi, G. (2001). *La fragilità degli anziani*. Livorno: Debatte.

Sitografia

http://www.salute.gov.it/portale/temi/p2_6.jsp?id=1232&area=ricoveriOspedali&menu=vuoto, consultato in data 16/01/2017.

http://www.salute.gov.it/portale/salute/p1_5.jsp?id=139&area=Servizi_al_cittadino_e_al_paziente, consultato in data 23/01/2017.

http://www.salute.gov.it/portale/temi/p2_6.jsp?lingua=italiano&id=1014&area=esenzioni&menu=vuoto, consultato in data 23/01/2017.

<http://www.epicentro.iss.it/focus/osteoporosi/osteoporosi.asp>, consultato in data 07/02/2017.

http://www.ilsole24ore.com/art/notizie/2016-02-24/perche-2015-italia-sono-morte-54mila-persone-piu-9percento-173100.shtml?uuid=ACB9EFbC&refresh_ce=1, consultato in data 08/01/2016.

[http://www.treccani.it/enciclopedia/istituzionalizzazione_\(Dizionario-di-Medicina\)/](http://www.treccani.it/enciclopedia/istituzionalizzazione_(Dizionario-di-Medicina)/), consultato in data 24/01/2017.