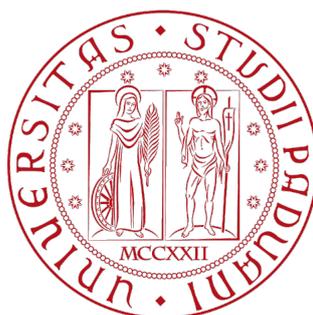


Università degli Studi di Padova
Dipartimento di Scienze Statistiche
Corso di Laurea Magistrale in
Scienze Statistiche



**Identificazione dei soggetti fragili nell'ULSS 15:
una nuova proposta basata sugli ordinamenti
parziali**

Relatore: Prof. Giovanna Boccuzzo
Dipartimento di Scienze Statistiche

Correlatore: Dr. Armando Olivieri
Dipartimento di Prevenzione, ULSS 15 "Alta Padovana"

Laureanda: Margherita Silan
Matricola N. 1076806

Anno accademico 2014/2015

Indice

Introduzione	7
1 La fragilità: definizioni e caratteristiche	9
1.1 Definizioni nella letteratura	10
1.2 I paradigmi della fragilità	13
1.2.1 Il Paradigma Biomedico	13
1.2.2 Il Paradigma dell'accumulo dei deficit	14
1.2.3 Il Paradigma Bio-Psico Sociale	16
1.3 Una definizione di <i>fragilità</i> vincente	18
1.3.1 Validità nei contenuti	20
1.3.2 Validità nell'interpretazione	20
1.3.3 Validità nei criteri	20
1.4 Dalla definizione concettuale a quella operativa	21
1.5 Raccogliere dati sulla fragilità	23
1.6 Screening della fragilità in Italia	24
1.6.1 Uno studio pilota in Toscana	25
1.6.2 I soggetti fragili alle ondate di calore in Lombardia	26
1.6.3 Adjusted Clinical Groups (ACG): dalla Johns Hopkins al Veneto	26
1.6.4 Modelli e fattori di rischio in Emilia-Romagna	27
2 Descrizione dei soggetti fragili secondo il sistema Adjusted Clinical Groups	29
2.1 Introduzione	29
2.2 Cos'è il sistema Adjusted Clinical Groups	30
2.3 La classificazione internazionale delle diagnosi	31
2.4 Dati di input per il sistema ACG	32
2.5 Dati utilizzati per la descrizione dei soggetti fragili nell'Ulss 15	33

2.5.1	La popolazione assistita dall'Ulss 15 nel 2013	33
2.6	Medically Frail Condition Marker	35
2.6.1	Gli utenti fragili dell'Ulss 15 nel 2013	36
2.7	Strumenti forniti come output dal sistema ACG	38
2.7.1	Aggregated Diagnostic Groups	38
2.7.2	Adjusted Clinical Groups	40
2.7.3	Resource Utilization Bands	42
2.7.4	Expanded Diagnosis Clusters	44
2.7.5	Rx-Defined Morbidity Groups	46
2.7.6	Condizioni croniche	48
2.8	Follow-up degli utenti fragili nel 2013	49
2.9	Validazione del Marker per la fragilità di ACG	51
2.10	Conclusioni	54
3	Un modello predittivo per le prognosi negative legate alla fragilità	55
3.1	Introduzione	55
3.2	Il modello MoSaiCo	56
3.3	Il record linkage delle fonti	61
3.3.1	L'indice di deprivazione	63
3.4	La riorganizzazione delle variabili	64
3.4.1	Nuove variabili con riferimento annuale	64
3.4.2	Le classi di età	67
3.5	Analisi per classi di età	68
3.6	Analisi di regressione logistica per prevedere le prognosi negative della fragilità	72
3.7	Un modello loglineare per identificare le interazioni tra le variabili	75
3.8	Stratificazione in base al rischio di incorrere nelle prognosi negative della fragilità	78
4	Un nuovo approccio per la misura della fragilità mediante gli ordinamenti parziali	81
4.1	Introduzione	81
4.2	Ordinamenti parziali e diagrammi di Hasse	82
4.2.1	Alcune basi teoriche per gli ordinamenti parziali	83
4.2.2	Definizioni di base	85
4.3	Estensioni lineari	89
4.4	Approssimazioni per il calcolo degli averaged rank	91

4.4.1	Local Partial Order Model, LPOM0	91
4.4.2	Extended Local Partial Order Model, LPOMext	93
4.4.3	Classi di equivalenza e una modifica all' LPOMext	95
4.5	LPOMext per la misura della fragilità nell'Ulss 15	98
4.5.1	Segmentazione degli utenti dell'Ulss 15 in base all'averaged rank	108
Conclusioni		115
Appendice		119
Ringraziamenti		127
Bibliografia e Sitografia		129

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". Lo scopo principale di questo lavoro è quello di trovare un algoritmo in grado di identificare i soggetti fragili, attraverso un'analisi critica degli strumenti utilizzati più comunemente.

Nel capitolo 1, ci si è interrogati sul significato del termine *fragile*, ricercando in letteratura definizioni e analisi riguardo i soggetti fragili. Definire questo fenomeno è una questione controversa che ha interessato un numero crescente di ricercatori negli ultimi trent'anni, anche a causa del progressivo invecchiamento della popolazione. Di fatto non vi è ancora un punto d'accordo su una definizione né teorica, né operativa. Tuttavia vi sono alcuni punti condivisi, ovvero che si tratti di una condizione legata alla perdita di autonomia in diversi ambiti funzionali e che esponga a maggiori rischi, quali la morte o il ricovero d'urgenza.

Un primo strumento analizzato è quello già in uso dall'Ulss 15, il sistema *Adjusted Clinical Groups (ACG)*. Questo sistema, elaborato dalla *Johns Hopkins University*, fornisce un marcatore che identifica i soggetti fragili attraverso la perdita di alcuni domini funzionali. Per descrivere i soggetti fragili e validare tale definizione di fragilità verranno utilizzati, nel capitolo 2, le classificazioni che il sistema ACG fornisce come output.

Nel terzo capitolo viene replicato un algoritmo utilizzato dall'Ulss di Ravenna per identificare i soggetti fragili. In base ad un paniere di variabili, che rappresentano fattori di rischio per la fragilità, viene stimato un modello logistico per valutare quali siano gli individui con un rischio più alto di morte o ricovero d'urgenza nell'anno successivo. In base ai valori predetti di tale modello si stratificano gli utenti rispetto al livello di fragilità.

I due approcci precedenti risultano molto legati al quadro teorico ipotizzato in partenza, rischiando di imporre vincoli soggettivi ad un fenomeno dinamico e multifattoriale come la fragilità. Per evitare di imporre deficit funzionali che potrebbero non essere discriminanti e variabili dipendenti approssimative rispetto

ad un fenomeno multidimensionale come quello in questione, nel capitolo 4 viene utilizzato un approccio nuovo per identificare i soggetti fragili, la teoria degli ordinamenti parziali (poset). L'intuizione rispetto a tale approccio è quella di mettere in luce l'ordinamento naturale della matrice di dati, tramite un sottoinsieme di variabili validate precedentemente.

Capitolo 1

La fragilità: definizioni e caratteristiche

Il concetto di “*fragilità*” è stato lungamente discusso negli ultimi trent’anni, anche a causa del progressivo invecchiamento della popolazione che ha portato ad un numero sempre maggiore di utenti anziani in carico ai servizi sanitari. Nonostante l’ampia letteratura scientifica sul tema, non è ancora stato possibile raggiungere un accordo in favore di una definizione univoca di “*soggetto fragile*”, questo perchè ogni professionista si occupa di un aspetto diverso della fragilità, quello di sua competenza. Nella loro professione, infatti, psicologi, medici di famiglia, geriatri, infermieri, assistenti sociali vengono a contatto con sfaccettature diverse del medesimo concetto, la *fragilità*, che tutti riconoscono come complesso e multidimensionale (Cavazza & Malvi, 2014).

La “*fragilità*” assume quindi un ruolo chiave in diversi contesti. Per i geriatri, individuare i soggetti fragili potrebbe facilitare il processo decisionale clinico, migliorando le stime di benefici e rischi attraverso una valutazione quantitativa della fragilità. Per i ricercatori in ambito geriatrico, identificare i fattori di rischio biologici che causano la fragilità e il loro rapporto con i normali cambiamenti legati all’età potrebbe consentire analisi e ricerche mirate sulle cause della fragilità, accompagnate dall’identificazione di una popolazione di *pre-fragili* e la possibilità di attuare interventi di prevenzione. Per la programmazione nella sanità pubblica è utile analizzare gli utenti, soprattutto quelli fragili, nella misura in cui si possono prevedere i bisogni di una fascia di popolazione sempre più numerosa, consentendo un uso più efficace delle risorse (Martin & Brighton, 2008).

La mancanza di una definizione univoca comporta, sul versante pratico e analitico, notevoli difficoltà e imprecisioni nello stimare la prevalenza della fragilità e nel programmare interventi di prevenzione specifici.

1.1 Definizioni nella letteratura

In mancanza di una definizione univoca per il termine “*fragilità*”, si presentano diverse analisi sistematiche della letteratura per analizzare le sfaccettature messe in luce nei diversi ambiti di interesse.

Particolarmente interessante, a questo proposito, è uno studio di Marcon del 2010 volto proprio ad identificare le correlazioni del termine “*frailty*” negli abstract degli articoli che in letteratura si sono occupati di fragilità nelle età anziane dal 1983 al 2009. L’analisi della letteratura è stata svolta confrontando gli articoli provenienti da *PubMed*, per rappresentare un approccio medico, e da *PsycInfo*, per avere anche esempi provenienti da un ambito psico-sociale (Marcon *et al.*, 2010). Tra i vocaboli più comunemente associati alla fragilità, ne sono elencati alcuni che mettono in evidenza la ricerca di criteri e definizioni per il concetto di “*fragilità*” (come ad esempio “*definition*”, “*define*” e “*criterion*”); altri invece che riguardano la condizione di rischio e di perdita di salute in cui un soggetto fragile si trova (come ad esempio “*risk*”, “*loss*” e “*health*”); inoltre un gruppo di vocaboli fanno riferimento ai sintomi e ai segni clinici che descrivono un progressivo peggioramento della condizione clinica, fisica e cognitiva.

Confrontando le due fonti prese in considerazione nell’analisi, è possibile osservare la mancanza di un accordo sul significato da attribuire al vocabolo “*fragilità*”, che rimane molto legato all’ambito di ricerca. Mentre in ambito medico le parole che vengono più comunemente associate alla fragilità rientrano nella sfera fisica utilizzando termini specifici come ad esempio “*disability*”, “*mortality*”, “*clinical*” e “*weight*”; negli abstract pubblicati su *PsycInfo* compaiono spesso termini collegati alla persona e alle sue caratteristiche come “*ability*”, “*benefit*”, “*performance*”, “*activity*” e “*relationship*”. Infatti, in campo psico-sociale, la “*fragilità*” non viene letta soltanto come “*deficit*”, ma più che altro come fase naturale della vita con specifiche abilità e risorse.

Le analisi effettuate per cercare una definizione condivisibile e completa di “*fragilità*”, o anche solo per creare un quadro generale, sono numerose e sfruttano metodologie elaborate che coinvolgono l’analisi della letteratura e i pareri degli esperti (Metodo Delphi e Focus Group). Nell’analisi canadese “*Canadian Initiative on Frailty and Aging*”, ad esempio, è stato necessario raggruppare i quesiti di ricerca in 10 ambiti diversi per effettuare l’analisi della letteratura in modo sistematico (Bergman *et al.*, 2004). (Tabella 1.1)

Il Canadian Initiative on Frailty and Aging (CIFA) è una società internazionale di ricercatori e medici formata nel 2002, con l’obiettivo di migliorare la comprensione delle cause, dell’andamento e delle implicazioni della fragilità per

migliorare le condizioni di vita degli anziani che presentano alto rischio di fragilità attraverso la prevenzione, le cure e i servizi (Karunananthan *et al.*, 2009).

Esperti in diverse discipline hanno collaborato a questa revisione sistematica in modo da garantire maggiore consistenza allo studio. Per tale analisi della letteratura non sono stati esaminati soltanto gli articoli e i libri che contenessero il termine “*frailty*” o “*frail*”, ma anche tutte quelle con i termini “*aged*”, “*vulnerable*”, “*vulnerability*”, “*healthy aging*”, “*disability*”, “*disability evaluation*” e “*disabled persons*”, combinati con termini quali “*biomarkers*”, “*prevalence*” o “*diagnosis*”. Non è banale, tantomeno “*a priori*”, definire quali sono le caratteristiche attese di un soggetto fragile in quanto tale, nè valutare i legami tra la “*fragilità*”, la “*disabilità*” e la presenza di “*comorbidità*”, o intuire quali siano le sovrapposizioni tra queste condizioni e stabilire i rapporti causali che possono esservi tra queste.

Dominio	Quesiti di ricerca
Basi biologiche	Quali sono i fattori biologici e fisiologici della fragilità? Come possono essere utilizzati per definire, prevedere e descrivere la fragilità?
Basi sociali	Com'è stata concettualizzata (definita, modellata) la fragilità da un punto di vista sociale?
Prevalenza	Qual è la prevalenza della fragilità tra gli anziani? Come varia in base alle definizioni utilizzate?
Fattori di rischio	Quali sono i fattori utilizzati per predire la fragilità, il declino funzionale, l'invalidità, la mortalità o un maggiore utilizzo delle risorse? Quali sono invece i fattori utilizzati per predire un invecchiamento attivo e in salute?
Impatto	Che impatto ha la fragilità sugli individui che ne sono affetti? Sui loro parenti? E sul sistema sanitario?
Identificazione	Ci sono marker clinici asintomatici e misurabili che possono prevedere la fragilità? Quali sono i criteri clinico-diagnostici operativi? Quali sono gli strumenti per lo screening, la diagnosi e l'analisi della fragilità? Ci sono misure di che indicano la gravità della fragilità?
Prevenzione e Gestione	Esistono interventi che possono prevenire la fragilità? Esistono interventi dedicati alla popolazione generale per prevenire esiti negativi della fragilità (morte, istituzionalizzazione, ecc)? Esistono interventi dedicati alla popolazione fragile o a rischio di fragilità per prevenire le conseguenze della fragilità?
Territorio e Tecnologia	Quali interventi tecnologici hanno prodotto un effettivo miglioramento della qualità della vita e della sicurezza e quali invece no? Quali sono le caratteristiche comuni a quelle efficaci? Quali sono le esigenze che le tecnologie per assistere adulti e anziani fragili e i loro "caregivers" non hanno assolto adeguatamente?

Tabella 1.1: Quesiti di ricerca nella revisione della letteratura, "*Canadian Initiative on Frailty and Aging*". Traduzione da Bergman *et al.*, 2004

1.2 I paradigmi della fragilità

Non essendo possibile un'unica definizione di *fragilità*, sono state raggruppate le proposte all'interno di diversi paradigmi. Nel volume *“La fragilità degli anziani. Strategie, progetti, strumenti per invecchiare bene.”* di Gabriele Cavazza e Cristina Malvi (Cavazza & Malvi, 2014) e nel report tecnico della Regione Toscana (Consiglio Sanitario Regionale della Toscana, 2013) viene proposta una possibile suddivisione tramite tre paradigmi: il *“Paradigma Biomedico”*, il *“Paradigma dell'accumulo dei deficit”* e quello denominato *“Bio-Psico Sociale”*.

1.2.1 Il Paradigma Biomedico

Secondo il paradigma Biomedico la *fragilità* si definisce 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 avverse”*. Questo approccio è stato elaborato da Fried che ha isolato cinque caratteristiche che identificano un fenotipo di fragilità (Fried *et al.*, 2001); la presenza di almeno tre di questi è sufficiente a definire un soggetto *“fragile”* secondo questo paradigma.

Le cinque condizioni biomediche per definire un soggetto fragile sono quindi:

- **Perdita di peso involontaria:** intesa come una perdita di almeno 4.5 Kg (o di almeno il 5% del peso corporeo) nell'ultimo anno, senza che ve ne fosse l'intenzione (senza seguire particolari diete o esercizi).
- **Affaticamento:** facendo riferimento alla *Center for Epidemiologic Studies - Depression scale (CES-D)*, si considera *“affaticato”* un soggetto che dichiara di sentirsi stanco (*“I felt that everything I did was an effort.”*, *“I could not get going.”*) almeno tre volte a settimana.
- **Riduzione della forza muscolare:** misurata come forza di presa (in Kg) e stratificata in base al sesso e al BMI (Body Mass Index).
- **Ridotta attività fisica:** si basa sul *Minnesota Leisure Time Activity Questionnaire*, viene fatta una stima delle chilocalorie consumate facendo attività fisica ogni settimana; i livelli limite variano a seconda del sesso (<383 Kcal per i maschi e <270 Kcal per le femmine).
- **Riduzione della velocità del cammino:** valutata come tempo impiegato a percorrere 4.5 metri a piedi, con *cut off* specifico per sesso ed altezza.

Inoltre Fried si sofferma sulle reciproche relazioni tra “*fragilità*”, “*disabilità*” e “*comorbidità*” affermando che si tratta di concetti diversi, ma che in alcuni casi possono intersecarsi. Secondo le analisi di Fried e collaboratori su un campione di donne proveniente dai dati del *Cardiovascular Health Study (The Cardiovascular Health Study)*, infatti, fragilità e comorbidità si sono rivelati essere predittori per la disabilità; la disabilità tende ad aggravare condizioni di fragilità e comorbidità; e avere comorbidità può contribuire all’insorgenza di fragilità (Fried *et al.*, 2004). Inoltre, la presenza di disabilità e/o di fragilità potrebbe concorrere allo sviluppo o alla progressione di malattie croniche. Per riassumere quindi i risultati di queste analisi, oltre alla definizione biomedica di “*fragilità*”, Fried delinea quattro principali implicazioni in ambito assistenziale ad essa legate:

- Vulnerabilità agli stressor (ad esempio ospedalizzazioni, trattamenti invasivi);
- Necessità di trattare situazioni patologiche concomitanti, debolezza e denutrizione;
- Esigenza di ridurre i rischi di cadute, disabilità, ospedalizzazione, morte;
- Condizioni degeneranti con obbligo di prevenzione primaria e secondaria.

1.2.2 Il Paradigma dell’accumulo dei deficit

Il paradigma elaborato da Rockwood (1994) è più complesso e prende in considerazione un numero più ampio di variabili che non si limitano alla sfera biomedica, ma fanno riferimento anche a fattori cognitivi, psicosociali e geriatrici cercando di ricostruire un quadro più completo possibile della condizione in cui il soggetto si trova.

La “*fragilità*” è quindi 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*” (Rockwood *et al.*, 2007). Questa definizione deriva dall’approccio di Brocklehurst che ha preso in considerazione componenti biomediche e psicosociali per elaborare un modello dinamico per la “*fragilità*”, una sorta di *modello di guasto*. Questo modello si può rappresentare con una bilancia: su un piatto si possono porre le attività e le caratteristiche che aiutano il soggetto considerato a mantenere la sua indipendenza nella comunità (la salute, le capacità funzionali e un atteggiamento positivo e attivo); sull’altro vi sono i deficit che minacciano lo stato di indipendenza (le malattie croniche, la disabilità, la dipendenza da altre

persone nelle attività quotidiane). Per la maggior parte degli anziani, le attività e le caratteristiche positive superano abbondantemente i deficit; in altri si verifica la situazione opposta che causa la perdita dell'indipendenza. C'è però anche un terzo piccolo gruppo per cui i piatti della bilancia sono in un equilibrio precario: si tratta di soggetti in condizioni particolarmente critiche, ma ancora in grado di vivere nella comunità in modo indipendente, i pre-fragili (Rockwood *et al.*, 1994). In questo senso Rockwood definisce anche i *soggetti anziani fragili* come persone i cui deficit si pongono in equilibrio precario (o causano uno sbilanciamento) con lo stato di salute e che quindi dipendono dalla comunità per lo svolgimento delle attività quotidiane.

Sulla base della definizione di “*fragilità*” come accumulo di deficit, è stato costruito l'indice di fragilità (*Frailty Index*). Questo indice conta il numero di deficit che il soggetto possiede. Per deficit si intendono quelli fisici (ovvero sintomi, segni, disabilità, malattie), la riduzione delle attività quotidiane (ADL *Activities of Daily Living*), il deperimento delle abilità cognitive e della performance fisica e la depressione. Tra le linee guida per la costruzione di questo indice (Searle *et al.*, 2008) sono elencati cinque criteri secondo i quali i deficit esaminati (almeno 30-40) possono essere inseriti nel calcolo dell'indice:

- I deficit devono essere associati allo stato di salute;
- La prevalenza di un deficit deve essere positivamente correlata all'età;
- I deficit selezionati non devono arrivare a saturazione troppo presto;
- Nel gruppo devono essere presenti deficit provenienti da tutti i campi considerati;
- Se l'indice va confrontato nel tempo sulla stessa popolazione, i deficit devono essere gli stessi ad ogni misurazione.

Gli item scelti sono dicotomici, il valore 1 indica la presenza del deficit e 0 l'assenza, con la possibilità di creare gradazioni intermedie. Necessità di aiuto per lavarsi, vestirsi, alzarsi e sedersi, fare un giro intorno alla casa, mangiare, usare il bagno, fare la spesa, nei lavori domestici, rimanere a letto per almeno metà giornata a causa di problemi di salute, sentirsi depressi, soli, affaticati (la maggior parte del tempo=1, qualche volta=0.5, raramente=0), tumori maligni, artrite, diabete e obesità sono alcuni esempi di item utilizzati nell'analisi del 2008 di Searle, Mitnitski, Rockwood e collaboratori. Sommando i valori dei deficit stabiliti, si ottiene un indice per cui non è necessario avere un cut-off per distinguere i “*fragili*”

dai “*non fragili*”. Si ricava invece una misura dinamica e comprensiva di molteplici aspetti, utile per predire la probabilità di sperimentare una delle complicazioni legate alla “*fragilità*” o, più in generale, che si verifichi un peggioramento delle condizioni iniziali, ovvero che il numero di deficit aumenti (Rockwood & Mitnitski, 2007).

1.2.3 Il Paradigma Bio-Psico Sociale

La definizione di “*fragilità*” secondo il paradigma bio-psico sociale è stata proposta da Gobbens e collaboratori (2010), che la descrivono 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*”. In questo caso, più che del frutto di analisi quantitative su campioni di anziani o di soggetti a rischio, la definizione è il risultato di un attento studio della letteratura, seguito da uno schema rigoroso di interviste agli esperti (Gobbens *et al.*, 2010).

Per giungere ad una definizione riconosciuta dal maggior numero di esperti, è stato formulato un questionario, risultato da un’attenta analisi della letteratura, con domande su due principi fondamentali che Gobbens e collaboratori avevano individuato (multidimensionalità e continuità) e 11 definizioni possibili di “*fragilità*” presenti negli studi esaminati. Gli esperti avevano il compito di ordinare le definizioni assegnando undici punti a quella migliore e uno a quella che ritenevano meno adatta. Si riportano di seguito le prime cinque definizioni concettuali sulla base dell’opinione degli esperti (con valutazioni nettamente superiori rispetto alle seguenti):

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 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”.*

Dall'osservazione di queste cinque definizioni condivise dalla maggior parte degli esperti, è possibile ricavare i principali temi ricorrenti per descrivere la “*fragilità*”, oltre ai due principi precedentemente citati. Il primo principio è quello della “*multidimensionalità*” che interpreta la “*fragilità*” come una combinazione di problemi appartenenti a diversi domini funzionali (fisici, sensoriali, psicologici, sociali). Questa caratteristica della “*fragilità*” si basa su una valutazione olistica del soggetto fragile, ritenendo una visione unidimensionale non adeguata. L'approccio in questione non è condiviso da tutti gli esperti interpellati, infatti, alcuni di questi hanno dichiarato esplicitamente di preferire un approccio unidimensionale, ritenendo la “*fragilità*” un concetto di per sé già abbastanza complicato se sviluppato in modo unidimensionale. La prima, la seconda e la quinta definizione fanno riferimento proprio al fatto che i domini in cui compaiono i deficit funzionali sono diversi, quindi più di uno.

Il secondo principio, quello relativo alla “*continuità*”, mette in evidenza il fatto che si tratti di un processo degenerativo, eventualmente invertibile, frutto dell'interazione tra fattori personali e ambientali. L'idea di fondo è che non sia corretto rappresentare la “*fragilità*” come un fenomeno dicotomico, ma che sia più utile analizzarne le possibili sfumature.

Le definizioni 1, 3 e 4 fanno inoltre riferimento al concetto di “*capacità di conservazione*”, ovvero la capacità di un individuo di sopportare gli stress ambientali, mantenendo la propria *omeostasi* (equilibrio associato a una relativa stabilità interna delle proprietà chimico-fisiche che accomuna tutti gli organismi viventi). Nel caso l'omeostasi non fosse mantenuta, l'individuo potrebbe incorrere in eventi avversi con esiti negativi a cui fanno riferimento le definizioni 3 e 4.

La prima definizione è tratta da un articolo di Schuurmans e collaboratori (2004), che fa riferimento alla necessità di un criterio migliore dell'età per valutare i bisogni assistenziali delle persone, soprattutto di quelle anziane; infatti non tutti gli ultra settantacinquenni hanno particolari patologie per cui necessitano di assistenza particolare e, allo stesso tempo, ci sono persone anche con meno di 75 anni che potrebbero avere bisogno di servizi sanitari e/o assistenziali specifici e

onerosi. Nell'analisi presentata, viene confrontata la correlazione che il rischio di morte e il calo delle capacità di autogestione presentano con un indice di fragilità, il *Groningen Frailty Indicator*, e con una variabile dicotomica relativa all'età (maggiore o minore di 75 anni).

Il *Groningen Frailty Indicator* è un indice formato da 15 item (da questionario) che aiutano a determinare il livello di fragilità di una persona. Riassume, in particolare, la perdita di funzioni e risorse in quattro domini funzionali: quello fisico (scarsa mobilità, numero maggiore di problemi di salute, stanchezza, problemi di vista e udito), quello cognitivo (di funzionamento cognitivo), quello sociale (isolamento emotivo), e quello psicologico (depressione e ansia).

1.3 Una definizione di *fragilità* vincente

Come già visto, non esiste una definizione univoca e accreditata di “*fragilità*”. Si tratta di un concetto multidimensionale, e i domini che lo costituiscono interagiscono tra loro. Si può leggere una relazione con i concetti di *disabilità*, *comorbilità* (sovrapposizione e influenza reciproca di patologie fisiche o psichiche in un soggetto), *salute percepita* e *vulnerabilità a esiti avversi*. La fragilità è comunemente trattata dai geriatri, ma non vi è alcun consenso su una comune definizione tranne sul fatto che si tratti di una condizione di vulnerabilità e che dipenda da molti fattori. Gli studi approfonditi degli ultimi anni sulla “*fragilità*” forniscono una base concettuale per allontanarsi dai precedenti approcci basati sul singolo organo o su una specifica malattia e iniziare ad avere uno sguardo integrato sulla salute dell'individuo (Bergman *et al.*, 2007).

Un consensus group dell'*American Geriatrics Society* ha tentato di definire il concetto di *fragilità* come una sindrome fisiologica caratterizzata da una diminuzione della resistenza ai fattori di stress, derivante da un calo complessivo generale in più di un sistema fisiologico, e che causa maggiore vulnerabilità ed esposizione ad esiti negativi. È stato proposto anche un fenotipo¹ di fragilità fisica come combinazione di perdita di peso, stanchezza, ridotta forza di ripresa, ridotta attività fisica o lentezza nei movimenti. Inoltre, sono state studiate e proposte numerose misure cliniche che coinvolgono diverse funzionalità come quelle cognitive, funzionali, sociali, e che vanno ben oltre i meri aspetti fisici (Fried e il paradigma biomedico).

Un altro possibile approccio è quello di accettare la varietà di definizioni di *fragilità* esistenti e ritenere che la ricerca dovrebbe produrre una definizione

¹Per fenotipo si intende l'insieme delle caratteristiche osservabili relativamente ad un soggetto.

abbastanza chiara da prendere forma all'interno dello spettro di definizioni già esistente. Questo è l'approccio adottato dall'iniziativa canadese per la fragilità e l'invecchiamento, che ha ripartito le definizioni di fragilità in quattro classi:

- definizioni fisiologiche;
- definizioni basate sulla fragilità come sindrome complessa;
- definizioni sulla base di un modello bilanciato (che aggiunge al concetto di sindrome complessa elementi sociali);
- definizioni di fragilità come sindrome geriatrica (delirio e cadute).

Nel corso del tempo, è probabile che alcune definizioni di fragilità avranno più successo di altre, ma Rockwood sostiene che una definizione vincente della *fragilità* dovrebbe soddisfare i seguenti criteri (Rockwood, 2005):

1. Validità nei contenuti

- Include più fattori;
- È dinamica;
- Sostituisce validamente precedenti definizioni di successo;
- Ampiamente utile in vari contesti (ad esempio, in studi clinici o di popolazione);
- Computazionalmente trattabile.

2. Validità nell'interpretazione

- È più comune nelle donne che negli uomini;
- È più comune con l'aumentare dell'età;
- Correlata con la disabilità;
- Correlata con co-morbilità e con la salute percepita.

3. Validità nei criteri

- Predice la mortalità;
- Prevede anche altri esiti avversi (delirio, cadute);
- Prevede un'età in cui tutti sono fragili;
- Si basa su modelli cellulari e animali per studi di persone.

1.3.1 Validità nei contenuti

Una definizione di fragilità deve rispettare almeno le linee generali condivise in letteratura, per quanto riguarda sia le principali caratteristiche di questo fenomeno (ad esempio la multifattorialità), sia il quadro teorico di riferimento, sia i riscontri operativi e computazionali. Ad esempio, in linea con un approccio multifattoriale, non sembra accettabile che la compromissione di una singola funzionalità possa determinare, da sola, una condizione di fragilità. Le modifiche dei fattori del quadro teorico di riferimento per definire la fragilità di un individuo, dovrebbero essere supportate da un'accurata verifica empirica. Inoltre, sono da tenere presenti anche gli aspetti operativi e computazionali, in quanto spesso non solo vi sono molti fattori da analizzare, ma anche le relative interazioni. La dipendenza tra le variabili non sempre è facile da valutare in modo soddisfacente soprattutto in contesti così complessi. Tali problematiche risultano ancora più critiche se si tiene presente che il fenomeno della fragilità continua ad evolversi di pari passo con i cambiamenti socio-economici.

1.3.2 Validità nell'interpretazione

La definizione operativa deve essere coerente con le altre misure del fenomeno, relativamente a quanto già noto, infatti va misurata la validità di una nuova definizione basandosi sulla correlazione con quelle precedenti. Si deve inoltre ottenere una significativa correlazione con la disabilità, con l'età avanzata, con le comorbidità e con la salute. Dato che le donne disabili vivono più a lungo rispetto agli uomini, è probabile che la fragilità sia più frequente tra le donne.

1.3.3 Validità nei criteri

Si verifica quando la definizione seleziona i soggetti in merito ad un determinato esito negativo (come la morte, l'istituzionalizzazione, prolungati ricoveri in ospedale, o cadute ripetute), infatti la fragilità è senza dubbio legata in qualche modo anche alla vulnerabilità. La capacità predittiva risulta quindi uno degli aspetti più importanti per la definizione di fragilità, soprattutto ai fini della programmazione sanitaria e degli interventi di prevenzione. Tuttavia è bene tenere presente che ci sono decessi e ospedalizzazioni non preceduti da vulnerabilità e che quindi la correlazione tra questi eventi e la fragilità non deve essere perfetta. Un risultato particolarmente ricercato è un'età alla quale si possa stabilire che tutti gli individui sono da considerarsi fragili; che secondo le analisi portate avanti da Rockwood potrebbe essere intorno ai 95 anni.

1.4 Dalla definizione concettuale a quella operativa

Finora si è fatto un excursus su come definire la *fragilità*, fornendo possibili interpretazioni di questo fenomeno da un punto di vista concettuale. Tuttavia, è necessario distinguere tra una definizione concettuale ed una operativa. Infatti in una definizione concettuale, la nozione è definita in termini di altri concetti. Una definizione operativa, nota anche come definizione empirica, definisce l'elemento di interesse sulla base di criteri che devono essere applicati per determinare la presenza e la portata del fenomeno in esame. Una definizione operativa esplicita un concetto in termini di dati osservabili. Queste due definizioni, quella concettuale e quella operativa, devono corrispondere ed essere collegate: la definizione concettuale deve dare una direzione all'operativizzazione del concetto (Gobbens *et al.*, 2010).

La definizione operativa è necessaria per calcolare incidenza e prevalenza della *fragilità*, ma non essendoci uniformità nelle definizioni concettuali, anche i criteri che definiscono se e quanto un soggetto è *fragile* non sono univoci e condivisi.

Sono state portate avanti numerose indagini e analisi per cercare una definizione operativa condivisibile, come ad esempio l'esperienza fatta da Leocadio Rodríguez-Mañas e collaboratori che hanno coinvolto esperti per cercare di selezionare un ridotto numero di criteri utili a definire operativamente un *soggetto fragile* (Rodríguez-Mañas *et al.*, 2013). Sono state alternate iterazioni di un processo Delphi con dei Focus Group per passare da 107 possibili criteri a 39 accettati (anche integrando il paniere di criteri iniziali grazie al contributo degli esperti interpellati). Nel corso dell'indagine, è stato trovato un accordo unanime sulla necessità di combinare più di un fattore, ma, essendo la percezione di *fragilità* diversa a seconda della formazione e della nazionalità di ogni esperto, non è stato identificato un paniere di criteri che incontrasse un accordo totale (Rodríguez-Mañas *et al.*, 2013).

Un altro punto di discussione è quello riguardante l'età dei soggetti fragili: c'è una correlazione innegabile tra anzianità e fragilità, ma è corretto individuare un cut off sull'età come condizione necessaria? Il problema non è soltanto riferito al ruolo dell'età, cioè se sia da interpretare come condizione necessaria, soltanto come criterio aggiuntivo o, nel caso sia molto avanzata, come requisito sufficiente. Non c'è univocità nemmeno nello stabilire quale sia l'età a cui fare riferimento o da usare come cut off.

Come per la definizione concettuale, sono state elaborate alcune revisioni sistematiche relative all'operazionalizzazione del concetto di *fragilità* tra cui quella

Fattori relativi alla fragilità	Operazionalizzazione
Stato nutrizionale	Peso corporeo Appetito Body Mass Index (BMI)
Attività fisica	Livello dell'attività fisica Tempo dedicato ad attività ludiche e fisiche (anche di gruppo)
Mobilità	Con difficoltà o con necessità di aiuto a camminare/spostarsi dentro e intorno alla propria abitazione Velocità del passo
Energia	Stanchezza Livello di energia (ad esempio fiacchezza/affaticamento)
Forza	Sollevamento un oggetto che pesa oltre 5 kg Debolezza di braccia o gambe Alzarsi dalla sedia Salire le scale Forza nella presa
Cognizione	Problemi di memoria Demenza diagnosticata o indebolimento cognitivo
Umore	Depressione Tristezza Ansia Nervosismo
Relazioni/sostegno sociale	Ricorso alla rete sociale (quando c'è bisogno di aiuto, c'è qualcuno disponibile?) Vuoti/Isolamento

Tabella 1.2: Operazionalizzazione dei fattori relativi alla fragilità. Traduzione da De Vriesa *et al.*, 2011

di De Vries e Collaboratori che prende in considerazione un totale di 2232 potenziali studi rilevanti. Questi, eliminati eventuali duplicati, vengono selezionati se corrispondono a due criteri: avere come scopo principale quello di sviluppare uno strumento per la misurazione della fragilità e presentare accuratamente ed esplicitamente la metodologia utilizzata. A seguito di un'ulteriore analisi, discussione e selezione, gli strumenti per misurare la fragilità su cui vengono effettuate le successive considerazioni sono ventuno (De Vriesa *et al.*, 2011).

Di questi ventuno metodi per misurare la fragilità, undici utilizzano dati provenienti da questionari auto-somministrati, uno i dati provenienti da un *performance test* (si tratta di vere e proprie prove fisiche, molto semplici, da sottoporre all'individuo da valutare), sette utilizzano i dati provenienti da entrambe le tipologie di fonti e due si basano soltanto sulla valutazione del geriatra. Per valutare anche la multidimensionalità di questi strumenti sono stati messi in evidenza otto fattori relativi alla fragilità e le rispettive declinazioni operative (Tabella 1.2).

Soltanto cinque, tra i ventuno strumenti considerati, utilizzano almeno un criterio appartenente a ciascun fattore (tra cui l'indice di fragilità di Rockwood e Mitnitski e il *Groningen Frailty Indicator*), anche se probabilmente gli altri sedici sono più ricchi sotto altri aspetti.

1.5 Raccogliere dati sulla fragilità

Finora si è parlato di analisi di dati, di modelli e di validazione di definizioni e di strumenti. Per far sì che questi passaggi siano possibili, è necessario che vi siano dati appositi e, nei limiti del possibile, raccolti ad hoc. Sono stati costruiti, negli ultimi anni, diversi metodi di raccolta dati per far sì che questi siano il meno possibile distorti e il più possibile confrontabili. Come precedentemente accennato, si può riempire un database per le analisi facendo ricorso ai pareri di un geriatra, utilizzando un *performance test* (che può valutare la prestanza fisica e l'indipendenza nella vita domestica), sottoponendo un questionario (auto-somministrato o con intervistatore) oppure utilizzando i flussi informativi già presenti.

I questionari a cui fare riferimento sono numerosi, infatti la maggior parte delle analisi sulla fragilità si basano su indagini campionarie svolte tramite questionario (eventualmente accompagnate da *performance test*).

Il gruppo *Survey of Health Ageing and Retirement in Europe (SHARE)* ha proposto uno strumento di semplice utilizzo per rilevare e quantificare la condizione di fragilità secondo il paradigma biomedico e, quindi, i criteri suggeriti da Fried e collaboratori. Si tratta dello strumento SHARE-FI (*Survey of health Ageing and retirement in Europe Frailty Index*). Questo questionario è infatti strutturato nelle cinque sezioni: stanchezza, perdita di appetito, debolezza, difficoltà di movimento e scarsa attività fisica.

Un altro esempio di questionario è lo *Sherbrooke postal questionnaire*², utilizzato

²Il questionario *Sherbrooke postal questionnaire* è riportato di seguito (la risposta tra parentesi indica la modalità associata ad una condizione di maggiore rischio):

per lo screening postale (Hebert *et al.*, 1996).

Per avere un controllo continuo ed universale sugli utenti dei servizi sanitari, invece, è possibile sfruttare, tramite *record linkage*, le fonti informative necessarie ai processi gestionali, cercando di adattare, quando possibile, i dati raccolti per scopi diversi da quelli della ricerca ai bisogni informativi messi in evidenza per l'individuazione dei soggetti fragili. Utilizzare fonti informative già presenti porta ad avere dati su tutta la popolazione in esame, non soltanto su un campione di essa. Purtroppo spesso questi dati possono avere, essendo database molto ampi, alcune colonne e talvolta interi record mancanti o con molti errori o imprecisioni.

1.6 Screening della fragilità in Italia

Negli ultimi anni la fragilità è stata senza dubbio un tema di grande interesse in Italia, per diverse ragioni:

- l'aumento dell'aspettativa di vita ha portato ad un maggiore sviluppo delle disabilità provocando un aumento dei costi dell'assistenza e la necessità di allocare le risorse in modo più efficiente,
- serve identificare sistematicamente gli anziani fragili per attivare interventi di prevenzione mirati,
- diventa quindi sempre più importante costruire e validare sistemi di identificazione multidimensionale del rischio.

Dopo la prima ondata di calore in Italia nel 2003, il concetto di *fragilità* è stato legato alle ondate di caldo, tanto che il Ministero della Salute formula un vero e proprio piano di risposta per contrastarne gli effetti ogni anno (Bargagli *et al.*, 2013) e incoraggia le regioni a portare avanti progetti e indagini sul tema della fragilità.

Di seguito sono riportate alcune esperienze di ricerca sul tema dell'identificazione dei *soggetti fragili* nel territorio italiano.

-
- Vive solo? (No)
 - Utilizza più di tre diversi farmaci ogni giorno? (Si)
 - Usa regolarmente un bastone, un deambulatore o una sedia a rotelle per spostarsi? (Si)
 - Ci vede bene? (No)
 - Ci sente bene? (No)
 - Ha qualche problema di memoria? (Si)

Un soggetto è a rischio di declino funzionale se presenta più di un fattore di rischio o non restituisce il questionario. (Hebert *et al.*, 1996)

1.6.1 Uno studio pilota in Toscana

In Toscana, nel 2013, è stata condotta un'indagine di screening della fragilità sulla popolazione anziana somministrando un questionario postale ed effettuando una valutazione multidimensionale al domicilio dei residenti per la prevenzione secondaria della disabilità (Razzanelli *et al.*, 2013). Un primo passaggio è stato effettuato attraverso lo screening con questionario postale di tutti gli ultra70enni (esclusi dal campione quelli che percepivano l'indennità di accompagnamento, quelli già a carico dei servizi, per migliorare l'efficienza dell'indagine). Si tratta di un questionario molto semplice, formato da 7 item (una versione modificata del *Sherbrooke postal questionnaire*). In un secondo tempo è stata fatta anche una valutazione a domicilio dei rispondenti che hanno dato almeno tre risposte che rappresentano una condizione di rischio nel questionario precedentemente, attraverso una valutazione multidimensionale domiciliare (VMD) per valutare l'autonomia nelle attività quotidiane (BADL, incidenza di disabilità in almeno un'attività di base della vita quotidiana).

Questa indagine fa riferimento ad un'esperienza canadese: 6 item che corrispondono a 6 predittori indipendenti di perdita di autonomia nell'anno successivo (sensibilità 75%, specificità 52%); lo *Sherbrooke postal questionnaire* è stato tradotto in italiano e modificato per ottenere le seguenti 7 domande a risposta dicotomica.

1. Vive con qualcuno?
2. Usa 4 o più diverse medicine al giorno?
3. Ha problemi di vista?
4. È mai caduto nell'ultimo anno?
5. È mai stato ricoverato nell'ultimo anno?
6. Ha problemi di memoria?
7. Ha qualche difficoltà a camminare per 400 metri?

Ad ogni risposta affermativa viene assegnato un valore unitario che viene sommato per ottenere un punteggio di fragilità. Al termine del questionari è stata aggiunta una domanda come proxy di disabilità conclamata: "Riesce ad alzarsi dal letto senza l'aiuto di un'altra persona?"

La valutazione multidimensionale domiciliare viene effettuata secondo strumenti scelti secondo criteri di semplicità di somministrazione e di affidabilità.

Viene sottoposta la valutazione delle ADL di Katz (Katz *et al.*, 1963) come filtro: coloro che avessero perso l'autonomia anche in una sola delle ADL sono stati indirizzati per una valutazione finalizzata alla redazione di un piano assistenziale personalizzato; gli altri sono stati valutati in base alle attività strumentali della vita quotidiana, al tono dell'umore, alla performance motoria degli arti inferiori, al rischio di malnutrizione per difetto, a deficit uditivi, visivi e cognitivi, a fattori di rischio ambientali e alla politerapia farmacologica. A seguito di tali valutazioni multidimensionali sono stati segnalati, previo consenso dell'intervistato, i casi di non autosufficienza ai medici di medicina generale e agli assistenti sociali per interventi di prevenzione.

Il questionario postale proposto si è rivelato utile e valido, in quanto ha rilevato soggetti con problematiche di salute potenzialmente oggetto di interventi di prevenzione della disabilità.

1.6.2 I soggetti fragili alle ondate di calore in Lombardia

L'azienda Sanitaria Locale di Milano (Servizio di Epidemiologia di Milano, 2015), a seguito delle ondate di calore del 2003, ha elaborato diversi progetti di ricerca e programmi di prevenzione per limitare l'impatto del clima sullo stato di salute degli anziani. L'algoritmo elaborato in Lombardia ha molte similitudini con quello di accumulo dei deficit di Rockwood, infatti si basa sulla stratificazione degli anziani con più di 75 anni in classi di rischio a seconda della presenza di alcune caratteristiche. I deficit considerati si riferiscono all'età (con più di 75 anni), all'isolamento (persone che vivono sole), all'utilizzo di farmaci (sono stati individuati alcuni farmaci che causano scompensi se le temperature sono particolarmente alte), alle malattie croniche (che espongono maggiormente a rischi per la salute durante le ondate di calore) e alla situazione economica (individui residenti in aree di censimento caratterizzate da una situazione di deprivazione economica).

1.6.3 Adjusted Clinical Groups (ACG): dalla Johns Hopkins al Veneto

Il sistema ACG (Adjusted Clinical Groups) è stato sviluppato dalla Johns Hopkins Bloomberg School of Public Health con lo scopo di ottimizzare le risorse e l'offerta dei servizi sanitari (*The Johns Hopkins ACG System*). Questo sistema utilizza l'informazione di alcuni flussi informativi, ne rielabora i dati e produce analisi e previsioni per quanto riguarda la misurazione delle morbosità,

l'identificazione dei pazienti ad alto rischio, la previsione delle risorse sanitarie da impiegare e i report su quelle già erogate. Questo strumento, già utilizzato in diverse parti del mondo (Canada, Israele, Spagna, Germania, Svizzera e Regno Unito) è stato adottato, in modo sperimentale, nel 2012 in Veneto per un progetto biennale limitato a due aziende sanitarie e, dal 2014, il suo utilizzo è stato esteso a tutto il Veneto.

Il sistema ACG produce anche alcuni marker per identificare utenti con particolari esigenze e che richiedono cure ad hoc, uno di questi è il *Medically Frail Condition Marker* che consiste in una variabile dicotomica che vale 1 se l'individuo è da considerarsi fragile e 0 altrimenti. Questo marker e il sistema ACG verranno ulteriormente descritti nel Capitolo 2.

1.6.4 Modelli e fattori di rischio in Emilia-Romagna

L'Ulss di Ravenna ha sviluppato invece uno studio di coorte prospettico storico per identificare i pazienti ad alto rischio di ricovero urgente o di morte nell'Ulss di Ravenna. È stata modellata la probabilità di incorrere in questi *outcome* tipici della popolazione fragile con un modello multivariato di regressione logistica (MoSaiCo, che sta per Modello Statistico Combinato). Le variabili inserite nel modello come predittori sono di tipo demografico, sociale, clinico e di uso dei servizi sanitari e sociali. Il modello ottenuto è stato validato e presenta buoni livelli di potere predittivo, inoltre ha prodotto, per ogni soggetto, un coefficiente di rischio utile a stratificare gli utenti e a pianificare interventi di programmazione futuri. Maggiori dettagli su questo modello sono presenti nel Capitolo 3.

Capitolo 2

Descrizione dei soggetti fragili secondo il sistema Adjusted Clinical Groups

2.1 Introduzione

Lo scopo di questo capitolo è quello di analizzare le caratteristiche dei soggetti fragili secondo il sistema *Adjusted Clinical Groups* (ACG) utilizzando le classificazioni e gli strumenti forniti come output da ACG. Questo sistema, utilizzato solo dal 2012 in Veneto, ha identificato nel 2013 una lista di 1531 soggetti fragili nell'Urss 15 utilizzando il *Medically Frail Condition Marker*. Nella prima parte di questo capitolo si fornisce una descrizione del sistema ACG, del quadro teorico del *Medically Frail Condition Marker* e dei dati dell'Urss 15 che verranno utilizzati per analizzare e tracciare una descrizione della definizione di fragilità secondo ACG.

Nella seconda parte, si confrontano i soggetti fragili con il resto della popolazione sulla base dei principali raggruppamenti di ACG, al fine di mettere in luce le caratteristiche che contraddistinguono i soggetti fragili. Ci si attende che, confrontando questo sottogruppo di popolazione con il resto degli utenti dell'Urss 15, vi siano significative differenze che giustifichino tale suddivisione.

Si effettua in seguito un follow-up degli utenti fragili dal 2013 al 2014 per verificare l'andamento delle loro condizioni di salute. Per validare il *Medically Frail Condition Marker* viene stimato anche un modello che utilizza come variabile dipendente la morte o il ricovero d'urgenza nel 2014, inserendo tra le covariate la variabile dicotomica che vale uno se l'individuo risulta fragile secondo ACG e 0 altrimenti.

2.2 Cos'è il sistema Adjusted Clinical Groups

Il sistema Adjusted Clinical Groups (ACG) è stato sviluppato dalla *Johns Hopkins Bloomberg School of Public Health* con lo scopo di ottimizzare le risorse e l'offerta dei servizi sanitari. Questo sistema utilizza l'informazione di alcuni flussi informativi già disponibili a scopi gestionali per prevedere e pianificare i futuri interventi con i relativi costi in base alle caratteristiche degli utenti. ACG è un case-mix grouper, infatti stratifica l'utenza in base alle risorse utilizzate, alle diagnosi effettuate, alle cronicità manifestate e ai farmaci utilizzati, per valutarne l'impatto sulle risorse sanitarie con un'ottica sia previsiva, sia retrospettiva. Questo approccio di analisi è innovativo in quanto si basa sull'analisi dell'utente e dei suoi bisogni e non soltanto sulle prestazioni erogate. Questo strumento è stato adottato, in modo sperimentale, nel 2012 in Veneto con la DGRV n.439 del 20.03.12 per un progetto biennale «*Sistema ACG per la valutazione del case-mix territoriale nella Regione del Veneto: progetto pilota per l'adozione di strumenti innovativi di misurazione, analisi, programmazione e case-management per il governo clinico del territorio*» (Giunta Regionale del Veneto, 2012). Sono state scelte due Ulss pilota, l'Ulss 16 di Padova e l'Ulss 20 di Verona, per testare il funzionamento del sistema e valutarne l'aderenza alla realtà veneta.

Per poter utilizzare il sistema ACG, è stato necessario effettuare un record linkage tra i diversi flussi informativi presenti: archivio dell'Anagrafe Sanitaria Regionale (con l'aggiunta dei deceduti nel corso dell'anno di riferimento), flusso delle Schede di Dimissione Ospedaliera (SDO) nei 5 anni precedenti a quello di riferimento, flusso della mobilità passiva dei ricoveri ospedalieri nei 5 anni precedenti a quello di riferimento, flusso del Pronto Soccorso Aziendale (PS), flusso della Psichiatria Territoriale, flusso della Specialistica ambulatoriale (SPS), flusso della mobilità passiva della specialistica ambulatoriale, flusso dell'Assistenza Domiciliare Integrata (ADI), archivio delle valutazioni multidimensionali (UVM), archivio delle malattie rare, delle cure termali, delle esenzioni Ticket, riabilitazione ex art. 26, flusso della farmaceutica territoriale, flusso della farmaceutica territoriale in mobilità passiva, farmaceutica a distribuzione diretta o per conto, somministrazione diretta di farmaci in mobilità passiva.

Nel 2013, con la DGRV n.498 del 16 aprile, la Regione ha approvato l'estensione del progetto ad altre quattro Aziende Sanitarie: l'Ulss 4 dell'Alto Vicentino, l'Ulss 6 di Vicenza, l'Ulss 9 di Treviso e l'Ulss 2 di Feltre; raggiungendo così i 2 milioni di assistiti coinvolti nell'esperimento. Infine nel 2014 la giunta regionale del Veneto, in base alla documentazione esaminata e all'esperienza positiva fatta nelle Ulss pilota, ha concordato l'adozione del sistema ACG nelle 15 Ulss rimanenti.

2.3 La classificazione internazionale delle diagnosi

I codici relativi alle diagnosi assegnate ai pazienti sono di fondamentale importanza per l'implementazione del sistema ACG, queste utilizzano l'International Classification of Diseases (ICD) nelle versioni 9 (Ministero del Lavoro, della Salute e delle Politiche Sociali, 2007) e 10 (solo relativamente ai dati provenienti dal flusso della Psichiatria Territoriale). Nonostante i limiti di questo sistema di classificazione, il principale vantaggio è quello di essere confrontabile non solo tra tutti gli enti che hanno adottato ACG, ma anche tra i 194 stati membri del *World Health Organization*. La classificazione ICD è infatti la codifica utilizzata universalmente per la codifica delle diagnosi a scopi clinici, epidemiologici e gestionali. Nel 1893 viene approvata a Chicago dall'istituto internazionale di statistica come classificazione di cause di morte; a seguito di revisioni successive, nel 1948, è stata adottata dalla comunità internazionale anche come classificazione di cause di morbosità (sesta revisione). Nel 1979 un Comitato costituito da rappresentanti di Associazioni professionali ed accademiche di medici, Associazioni degli ospedali, dell'Ufficio regionale della Organizzazione Mondiale della Sanità e dell'agenzia *Health Care Financing Administration* (HCFA) ha approvato la nona versione di ICD: la ICD-9-CM International Classification of Diseases, 9th revision, Clinical Modification. La dicitura "Clinical Modification" mette in luce l'orientamento di questa versione alla classificazione delle informazioni riguardanti le morbosità piuttosto che le cause di mortalità.

La classificazione ICD-9-CM descrive le diagnosi (di malattia o traumatismo), i problemi di salute e le procedure diagnostiche e terapeutiche degli utenti, utilizzando codici numerici e alfanumerici. Tali codici sono stati costruiti tenendo conto di alcuni caratteri fondamentali: l'*esaustività* per cui i raggruppamenti finali della classificazione devono essere sufficienti a rappresentare tutte le entità; la mutua *esclusività* per cui non possono esserci ambiguità di classificazione per le entità; il *numero limitato* di raggruppamenti, circa 16000; e infine la *specificità* dei raggruppamenti tenendo presente la rilevanza di ognuno. In base ad eziologia e collocazione anatomica del problema clinico da classificare, le diagnosi vengono assegnate ad un blocco (per esempio "Osteopatie, Condropatie e Deformità Acquisite del Sistema Muscoloscheletrico" da 730 a 739), una categoria (definita univocamente da un codice di 3 caratteri, per esempio 737 "Deviazioni della colonna vertebrale") ed eventualmente anche una sotto-categoria (individuata da 4 caratteri, di cui uno a seguito di un punto, per esempio 737.2 "Lordosi (acquisita)") e una sotto-classificazione (spiegata da 5 caratteri, di cui 2 a seguito di un punto, per esempio 737.20 "Lordosi posturale (acquisita)"). In aggiunta

sono state predisposte due classificazioni supplementari di codici alfanumerici: i codici V che descrivono le circostanze riconducibili alle cause del ricorso alla prestazione sanitaria o legate alle condizioni di salute del paziente (per esempio V68.0 “Rilascio di certificati medici”); e i codici E che descrivono le cause esterne di traumatismo o di avvelenamento (ad esempio E913 “Soffocamento meccanico accidentale”).

La decima revisione della classificazione ICD (Ministero del Lavoro, della Salute e delle Politiche Sociali, 2008) ha incluso i codici E e V nel nucleo principale e, introducendo anche altri due nuovi settori, conta in tutto 21 capitoli. In questa edizione sono stati utilizzati codici alfanumerici in tutte le categorie, ora identificate da tre caratteri in sequenza (una lettera e due numeri).

2.4 Dati di input per il sistema ACG

Il sistema ACG viene alimentato da tre file di input in formato testo e collegati tra loro da un identificativo individuale. Questi tre file di input sono il frutto dell’aggregazione, secondo criteri minuziosamente descritti nella documentazione tecnica, dei ventidue flussi informativi richiesti.

- **Patient Data Format:** presenta un record per ogni utente contenente alcuni dati anagrafici (genere, età, distretto di appartenenza, . . .) e informazioni relative ai costi sostenuti nell’anno di rilevazione. Ogni assistito è associato ad un codice univoco, anonimo e identificativo che costituisce la chiave di linkage tra i vari file di input.
- **Medical Services Data Format:** in questo file vi sono tutti i codici di diagnosi registrati e disponibili per ciascun utente nel periodo considerato (incluso il flusso delle Schede di Dimissione Ospedaliera relativo ai cinque anni precedenti quello di riferimento). Ad ogni soggetto presente nel Patient Data Format potrebbe non essere associato alcun record o più record (anche con diagnosi duplicate) di questo file a seconda del numero di episodi registrati.
- **Pharmacy Data File Format:** Questo file contiene tutti i codici dei farmaci acquistati e prescritti dagli utenti nel periodo di rilevazione. Anche in questo caso può esserci un numero di record per ogni individuo presente nel Patient Data Format maggiore o uguale a zero.

2.5 Dati utilizzati per la descrizione dei soggetti fragili nell'Ulss 15

In questo capitolo si provvederà ad una descrizione il più possibile completa degli *output* forniti dal sistema ACG, mantenendo una prospettiva di analisi esplorativa e descrittiva della popolazione fragile selezionata dal Medically Frail Condition Marker. I dati utilizzati per le analisi sono il frutto di alcune *query* effettuate tramite il software del sistema ACG e successivamente rielaborate con il software SAS (versione 9.3). Si tratta di estrazioni fatte relativamente all'anno 2013 per tutti i 28 comuni dell' Azienda Ulss 15¹. Per la validazione del Medically Frail Condition Marker sono stati utilizzati anche i primi dati disponibili per l'anno 2014, in modo di avere una sorta di follow-up degli individui marcati come *fragili* nel 2013.

2.5.1 La popolazione assistita dall'Ulss 15 nel 2013

Gli assistiti dall'Ulss 15 nel 2013 e presenti nel database del sistema ACG sono 260676. Come si può vedere in figura 2.1, la struttura per età e genere non presenta particolari asimmetrie, tranne una presenza più marcata di donne, che sono più longeve, nelle età avanzate. La forma della piramide delle età è quella tipica di una popolazione invecchiata: una base piuttosto stretta rispetto alle classi di età intermedie che invece sono quelle maggiormente popolate grazie alla generazione dei *baby boomers*; un' estremità superiore non molto appuntita e sbilanciata verso la parte femminile. Quando si parla di *baby boomers* si intendono i bambini nati in Italia durante la crescita economica avvenuta negli anni '60, che hanno provocato un improvviso e marcato aumento delle nascite. Tale boom demografico si può individuare nelle barre centrali della piramide delle età che facilmente si distinguono per la loro lunghezza. Infatti circa il 13% della popolazione ha almeno 70 anni, mentre solo il 22% ne ha meno di 20. Ci si aspetta che la maggior parte degli utenti per cui sono state erogate prestazioni di carattere medico e sanitario sia anziana, sia a causa della presenza numerosa, sia a causa di una maggiore necessità di cure e assistenza, e che, tra questi, le donne siano presenti in percentuale maggiore.

¹Borgoricco, Campo San Martino, Campodarsego, Campodoro, Camposampiero, Carmignano di Brenta, Cittadella, Curtarolo, Fontaniva, Galliera Veneta, Gazzo Padovano, Grantorto, Loreggia, Massanzago, Piazzola sul Brenta, Piombino Dese, San Giorgio delle Pertiche, San Giorgio in Bosco, Santa Giustina in Colle, San Martino di Lupari, San Pietro in Gu, Tombolo, Trebaseleghe, Vigodarzere, Vigonza, Villa del Conte, Villafranca Padovana, Villanova di Camposampiero

Inoltre tra la popolazione di competenza di questa Ulss, il 10% non è di cittadinanza italiana: di questi, la maggior parte appartiene alla fascia d'età intermedia, tra i 20 e i 70 anni, mentre soltanto una percentuale inferiore al 2% ha più di 70 anni.

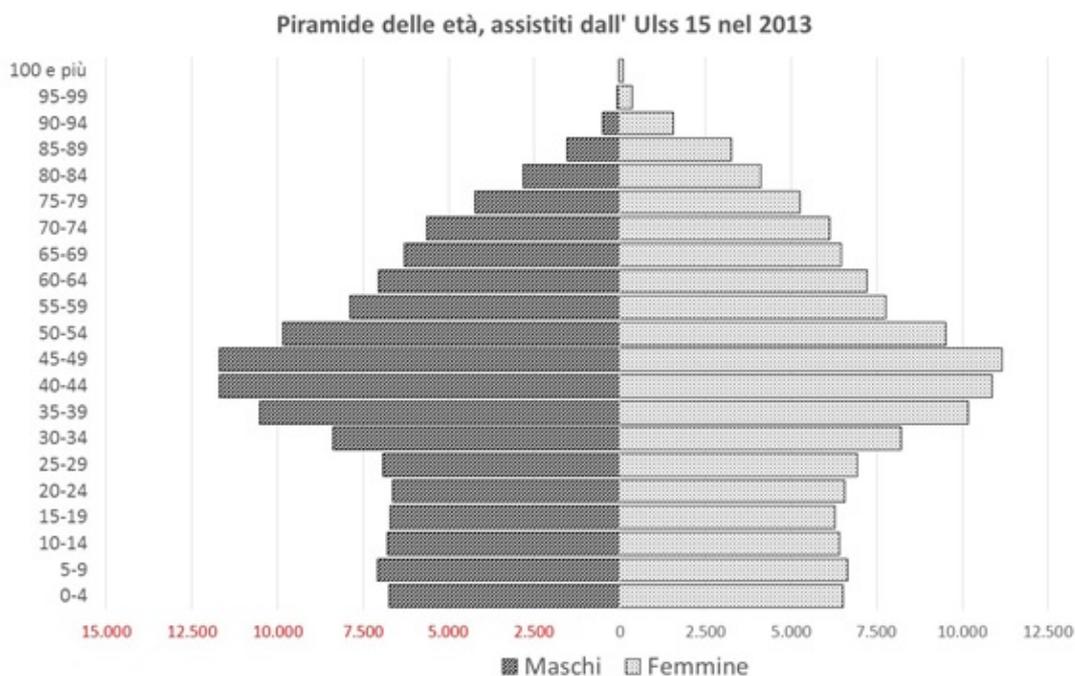


Figura 2.1: Gli assistiti dall'Ulss 15 nel 2013 per età e sesso, valori assoluti

2.6 Medically Frail Condition Marker

La definizione del concetto di fragilità proposta dal sistema ACG è legata a uno dei marker forniti come output. Si tratta di una variabile dicotomica costruita sulla base di alcune caratteristiche dell'utente. Per costruire questa variabile sono stati presi in considerazione dodici cluster che rappresentano problemi medici e sociali associati alla condizione di fragilità (Tabella 2.1). (The Johns Hopkins ACG System, 2011)

Concetti	Diagnosi (Esempi)
Malnutrizione	Marasma nutrizionale. Altre gravi malnutrizioni caloriche e proteiche.
Demenza	Demenza senile con caratteristiche deliranti e depressive. Demenza senile con delirio.
Problemi legati alla vista	Gravi problemi di vista ad entrambi gli occhi. Carenze visive gravi o moderate, con profonde differenze tra gli occhi.
Piaghe da decubito	Piaghe da decubito.
Incontinenza di urina	Incontinenza senza consapevolezza sensoriale. Perdita continua.
Perdita di peso	Perdita anormale di peso e sottopeso. Difficoltà di nutrizione e cattiva gestione.
Incontinenza delle feci	Incontinenza delle feci
Obesità (morbo)	Obesità morbosa
Povertà	Senza dimora. Abitazione inadeguata. Risorse materiali inadeguate.
Problemi di accesso alle cure	Nessuna agevolazione medica per la cura. Nessuna agevolazione medica.
Difficoltà a camminare	Difficoltà a camminare. Anomalie nell'andatura.
Caduta	Caduta dalle scale o camminando. Caduta dalla sedia a rotelle.

Tabella 2.1: Composizione del *Medically Frail Condition Marker*. Traduzione da The Johns Hopkins ACG System, 2011

2.6.1 Gli utenti fragili dell'Ulss 15 nel 2013

Il marker seleziona una particolare porzione della popolazione di assistiti dell'Ulss 15 che conta 1531 fragili nel 2013, pari a meno dell' 1%. La sezione di utenti così individuata presenta una struttura per età significativamente diversa da quella del resto della popolazione (Test di Kolmogorov-Smirnov, livello 0.01 di significatività). Si può infatti notare, osservando la figura 2.2, che la “piramide” ha una base molto larga, così come l'estremità superiore, mentre nelle fasce centrali d'età si stringe particolarmente. Soltanto nella prima classe d'età, relativa ai bambini con meno di 5 anni, si può trovare il 15% della popolazione fragile, mentre quasi la metà degli utenti così evidenziati ha più di 70 anni. Come previsto il grafico è asimmetrico, con un forte sbilanciamento sul lato destro, quello femminile, nelle età avanzate.

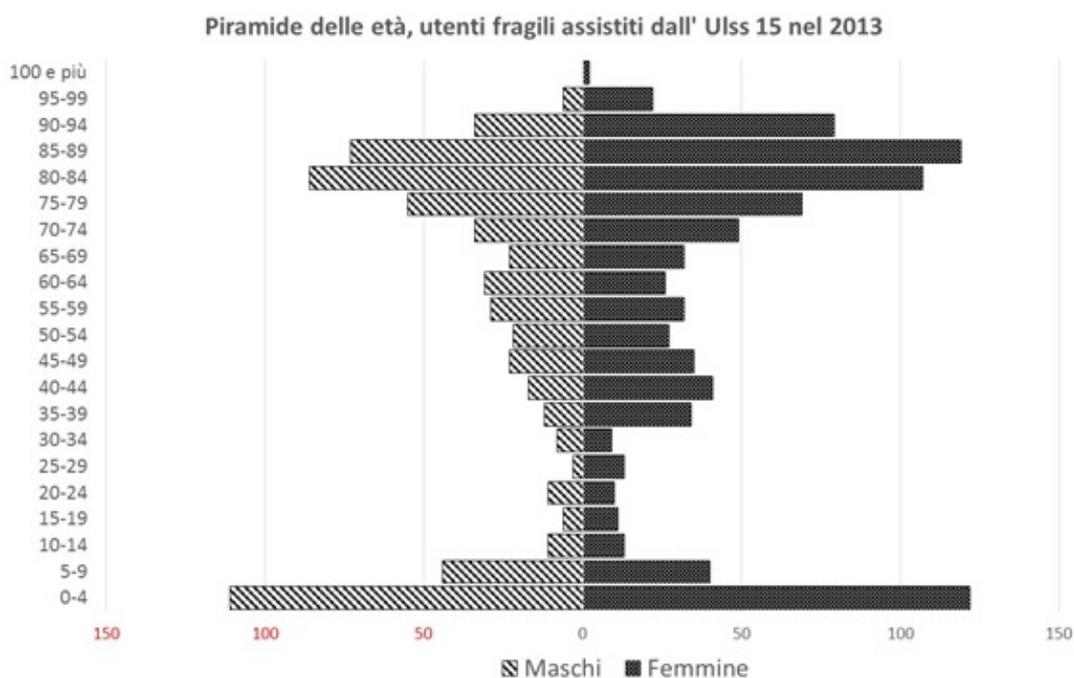


Figura 2.2: Gli utenti fragili assistiti dall'Ulss 15 nel 2013 per età e sesso, valori assoluti

La larghezza della base della piramide delle età è invece piuttosto inattesa: si contano 233 bambini con meno di 5 anni che presentano il marker per la fragilità. Il 97% di questi giovani soggetti presenta la stessa diagnosi, che ha contribuito in larga parte alla loro assegnazione al gruppo dei fragili (il rimanente 3% risulta avere diagnosi correlate a problemi di tipo neurologico). Si tratta della diagnosi ICD9 783.41 (*Failure to thrive, Ritardo dell'accrescimento*) che indica problemi nutrizionali e di alimentazione in età pediatrica tali per cui il bambino

subisce un rallentamento nella crescita o rischia di avere gravi problemi nello sviluppo. Nonostante si tratti di un grave problema, che espone il bambino a numerosi rischi, ci si chiede se sia verosimile e ragionevole tale classificazione: molti dei bambini inseriti nella lista dei fragili non presentano nessun altro segno o sintomo, nè hanno particolari storie o esperienze di ospedalizzazione; ci sono invece pochi altri bambini che, pur accusando diversi disturbi e patologie, non vengono catalogati tra i soggetti fragili. Si possono ipotizzare due possibili spiegazioni per questa apparente incongruenza: una è legata al fatto che una delle dimensioni del *Medically Frail Condition Marker* riguarda proprio problemi legati alla nutrizione e che questa diagnosi abbia un peso molto alto nell'attribuzione della fragilità; l'altra potrebbe essere connessa con l'attribuzione delle diagnosi, ovvero che la specifica diagnosi ICD9 783.41 venga attribuita con più leggerezza e genericità di quanto fosse previsto dal sistema ACG.

Per quanto riguarda la presenza straniera, soltanto il 6% degli utenti fragili non è di cittadinanza italiana (pari a meno dello 0.3% del totale dei residenti stranieri); si potrebbe erroneamente pensare che gli utenti stranieri siano più sani di quelli italiani, ma in realtà si tratta di una distorsione. I residenti stranieri, infatti, sono una popolazione già selezionata dalla migrazione: le motivazioni di tale spostamento sono generalmente di tipo lavorativo, per questo lasciano il loro paese d'origine soltanto le persone in età attiva con una sana costituzione fisica. La relazione causale tra fragilità e cittadinanza è quindi falsata dal processo di autoselezione.

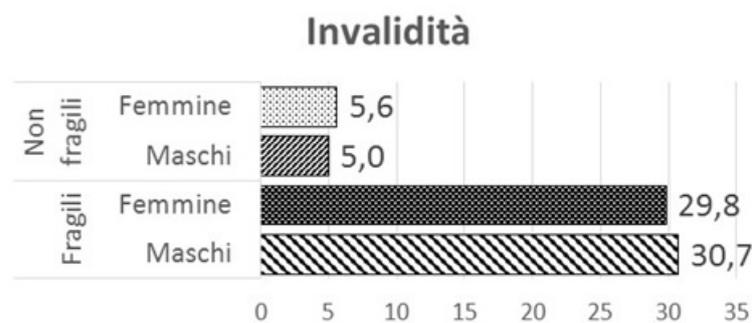


Figura 2.3: Proportione di utenti con invalidità, confronto tra utenti fragili e non fragili per sesso, in percentuali standardizzate per età

Come più volte ribadito nel capitolo 1, i concetti di invalidità e di fragilità non coincidono. Questa affermazione è valida anche per i fragili identificati con il *Medically Frail Condition Marker* come si può vedere dal grafico 2.3. La percentuale di soggetti con invalidità è significativamente diversa se si confrontano gli utenti fragili, 30.7% tra i maschi e 29.8% tra le femmine, e quelli non fragili, 5%

tra i maschi e 5.6% tra le femmine (Test sulla differenza tra due proporzioni, 1% di significatività). Tuttavia l'invalidità non è condizione necessaria, nè sufficiente per la fragilità, infatti non tutti gli invalidi sono fragili e non tutti i fragili sono invalidi.

Le percentuali rappresentate in tabella 2.3 sono state calcolate standardizzando i tassi grezzi per classi quinquennali di età utilizzando il metodo di standardizzazione diretta e la struttura per età della popolazione totale come riferimento.

2.7 Strumenti forniti come output dal sistema ACG

Il Sistema ACG fornisce una serie di strumenti utili per descrivere i bisogni assistenziali di una popolazione su diversi livelli di complessità e dettaglio. In questa sezione, tali strumenti vengono utilizzati per descrivere i particolari bisogni assistenziali degli individui fragili rispetto al resto della popolazione dell'Ulss 15 nel 2013. Nello specifico si descrivono, di seguito:

1. Aggregated Diagnostic Groups (ADG),
2. Adjusted Clinical Groups (ACG),
3. Resource Utilization Bands (RUB),
4. Expanded Diagnosis Clusters (EDC),
5. Rx-Defined Morbidity Groups (Rx-MG);
6. Condizioni croniche.

2.7.1 Aggregated Diagnostic Groups

Gli *Aggregated Diagnostic Groups* (ADG) sono dei raggruppamenti che fungono da base per la classificazione per ACG. Sono costituiti da 32 categorie diagnostiche che raggruppano diagnosi simili per bisogno assistenziale atteso (non per costi effettivi). Vengono definiti sulla base dei codici di diagnosi (ICD9, ICD10) in base a cinque parametri clinici:

- **Durata della condizione:** acuta, ricorrente o cronica.
- **Gravità della condizione.**
- **Certeza diagnostica:** sintomi o malattia documentata.
- **Eziologia della condizione:** malattie infettive, lesioni, o altro.

- **Necessità di cure specialistiche.**

ADG	Descrizione	Prevalenza negli utenti fragili
10	Malattia cronica: stabile	63.4%
11	Malattia cronica: instabile	52.0%
25	Disturbo psicosociale: ricorrente o persistente, instabile	36.5%
4	Temporaneo: grave infezione	22.1%
3	Temporaneo: grave	21.7%
9	Ricorrente degenerativo	20.2%

Tabella 2.2: Esempi di ADG

Nello stesso ADG possono essere raggruppate diverse diagnosi in base all'impatto atteso sul bisogno di risorse sanitarie. Allo stesso tempo, questi raggruppamenti non sono mutuamente esclusivi, quindi ad una stessa persona se ne può associare anche più di uno a seconda delle diagnosi attribuite. Nella tabella 2.2 sono elencati alcuni tra i più comuni raggruppamenti ADG, soprattutto nella popolazione dei fragili, con la relativa percentuale di prevalenza in tale popolazione nel 2013.

Si può notare, osservando la figura 2.4, che il Medically Frail Condition Marker suddivide gli utenti dell'Ulss 15 in due gruppi che si rivelano essere molto diversi anche dal punto di vista dei codici ADG. Le differenze calcolate tra le percentuali relative alla prevalenza delle classificazioni rappresentate, standardizzate per età, nei fragili e nei non fragili sono infatti significative al 1%. Anche patologie comuni, come ad esempio l'asma, sono presenti in percentuale maggiore tra gli utenti fragili rispetto ai non fragili. Più della metà degli utenti fragili è soggetto a patologie croniche con percentuali particolarmente alte soprattutto in età adulta, oltre che anziana (questo spiega la differenza tra i tassi standardizzati per età nel grafico e in tabella).

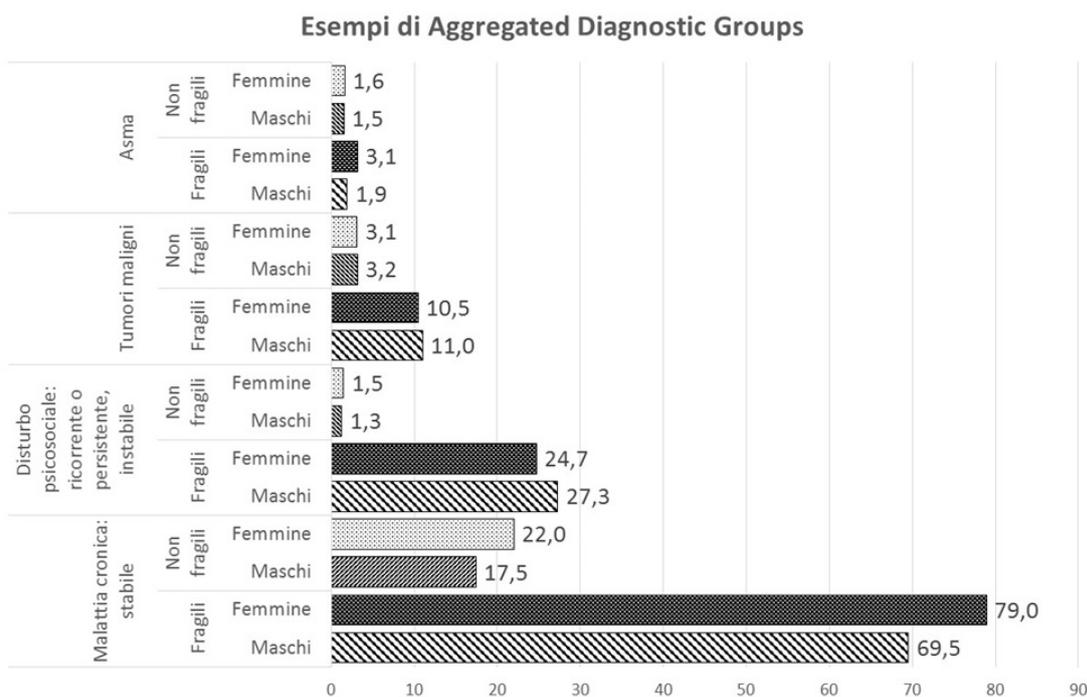


Figura 2.4: Esempi di ADG, confronto tra utenti fragili e non fragili per sesso e ADG, in percentuali standardizzate per età.

2.7.2 Adjusted Clinical Groups

Le categorie *Adjusted Clinical Groups* (ACG) identificano problemi di salute simili per impatto assistenziale. I gruppi sono costituiti da 93 categorie mutualmente esclusive (ad una persona è associato un solo ACG) ed esaustive (ad ogni persona è associata una categoria ACG) basate sulla combinazione individuale di ADG, età e sesso. L'attribuzione dei soggetti alle categorie ACG viene effettuata indipendentemente dai costi osservati.

In tabella 2.3 sono riportati alcuni esempi di raggruppamenti per ACG, con relative percentuali di prevalenza nella popolazione dei fragili identificati dal marker del sistema ACG. Il codice ACG riferito agli utenti senza diagnosi non è assegnato a nessuno dei soggetti fragili. Si può notare che più della metà dei soggetti fragili presenta condizioni multiproblematiche con più combinazioni di ADG, talvolta provenienti da macro-classi diverse, anche più di quattro. Le differenze riguardanti la distribuzione dei codici ACG tra soggetti fragili e non fragili (statisticamente significative allo 0.01) sono molto evidenti se si osserva la figura 2.5 che riporta le percentuali standardizzate per età di quattro esempi di codici ACG. Tuttavia il codice ACG, sebbene possa rappresentare condizioni di salute anche molto gravi con diverse complicanze, non è il criterio secondo cui la condizione di fragilità viene assegnata, infatti, ad esempio, con il codice 4940,

ACG	Descrizione	Prevalenza negli utenti fragili
200	Tra i 2 e i 5 anni, con lieve patologia acuta	6.7%
900	Malattia cronica: stabile	7.5%
4100	Più di 34 anni, con 2-3 diversi codici ADG	17.5%
4430	Più di 44 anni, con 4-5 diversi codici ADG di più di 2 macrocategorie	15.6%
4930	Più di 34 anni, con 6-9 diversi codici ADG di 3 macrocategorie	7.3%
4940	Più di 34 anni, con 6-9 diversi codici ADG di più di 4 macrocategorie	12.4%
5110	Senza Diagnosi	0%

Tabella 2.3: Esempi di ACG

“Più di 34 anni, con 6-9 diversi codici ADG di più di 4 macrocategorie” vi sono molti fragili, ma anche una piccola porzione di non fragili.

Per meglio spiegare la differenza tra ADG e ACG, si utilizza la classificazione corrispondente a “*Malattia cronica: stabile*” presente sia come categoria 10 di ADG, sia come classe 900 di ACG. Le percentuali di prevalenza nella popolazione fragile presenti rispettivamente nella tabella 2.2 nel paragrafo di ADG e in figura 2.5 sono infatti diverse, rispettivamente 63.4% in riferimento al codice ADG, mentre, per il codice ACG, 13.6% tra i maschi fragili e 21.6% tra le femmine fragili. Questo dipende dal fatto che i codici ACG sono univoci e vengono assegnati anche in base al numero di codici ADG attribuiti al singolo utente. Continuando l’esempio dei malati cronici stabili fragili (che sono in totale 971), i codici ACG attribuiti a questi utenti sono ben 36 diversi:

- a circa il 12% di questi viene attribuito il codice ACG 900, che corrisponde a “*Malattia cronica: stabile*”;
- a circa il 20% di questi viene attribuito il codice ACG 4100, che corrisponde a “*Più di 34 anni, con 2-3 diversi codici ADG*”;
- a circa il 18% di questi viene attribuito il codice ACG 4430, che corrisponde a “*Più di 44 anni, con 4-5 diversi codici ADG di più di 2 macrocategorie*”;
- a circa il 10% di questi viene attribuito il codice ACG 4930, che corrisponde a “*Più di 34 anni, con 6-9 diversi codici ADG di 3 macrocategorie*”;

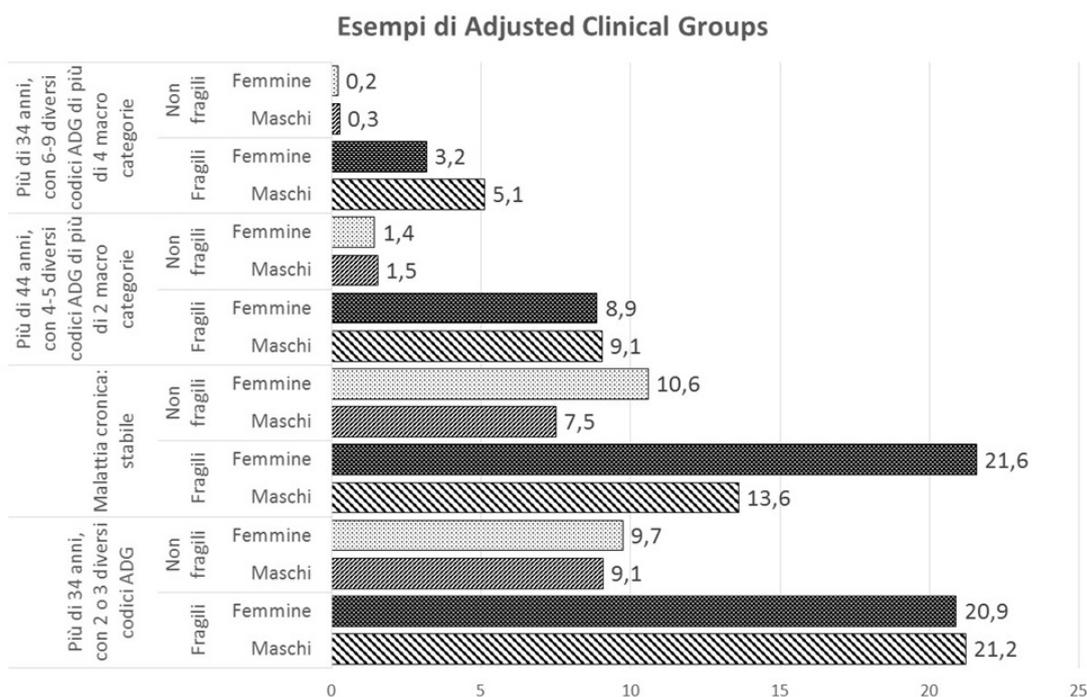


Figura 2.5: Esempi di ACG, confronto tra utenti fragili e non fragili per sesso e ACG, in percentuali standardizzate per età

- a circa il 15% di questi viene attribuito il codice ACG 4940, che corrisponde a *“Più di 34 anni, con 6-9 diversi codici ADG di più di 4 macrocategorie”*;
- il rimanente 25% degli utenti fragili con malattia cronica stabile si distribuisce in altri 31 codici ACG con frequenze molto più basse.

2.7.3 Resource Utilization Bands

Il *Resource Utilization Bands* (RUB) è una suddivisione in cinque categorie in base al consumo atteso di risorse da parte dell'utente, effettuata sui raggruppamenti ACG (a ognuno delle 93 classi cliniche è associata una modalità di RUB):

- 0 = Nessuna diagnosi, nessun consumo
- 1 = Utilizzatori sani
- 2 = Basso consumo di risorse
- 3 = Moderato consumo di risorse
- 4 = Alto consumo di risorse
- 5 = Molto alto consumo di risorse

Questa classificazione viene utilizzata più a scopo gestionale che clinico, infatti, l'obiettivo è quello di descrivere il grado di utilizzo delle risorse assistenziali degli utenti, non il loro pattern di malattie. Ad esempio, si può supporre che una donna incinta con complicazioni importanti, un individuo con una grave condizione psicologica, o, più in generale, un utente con due condizioni mediche croniche, usufruiscano di una quantità di risorse paragonabile in termini quantitativi, anche se ognuno di loro rientra in categorie ACG diverse. Si può eventualmente interpretare come una misura sintetica del grado di complessità assistenziale dell'utente, come proxy di quale sia il livello di autosufficienza rispetto al sistema sanitario.

La distribuzione dell'utilizzo delle risorse degli utenti fragili e non fragili è molto diversa: mentre la maggior parte degli utenti non fragili si distribuisce tra non consumatori e consumatori sani, più del 40% degli utenti fragili ha un consumo moderato di risorse e nessuno ha un consumo atteso pari a zero. La differenza tra le due distribuzioni di utilizzo di risorse per fragili e non fragili è stata verificata con il test di Kolmogorov Smirnov, che non dipende dalle dimensioni delle popolazioni utilizzate, ma soltanto dalla loro funzione di ripartizione empirica. Tale differenza è significativa allo 0.01, anche considerando separatamente maschi e femmine.

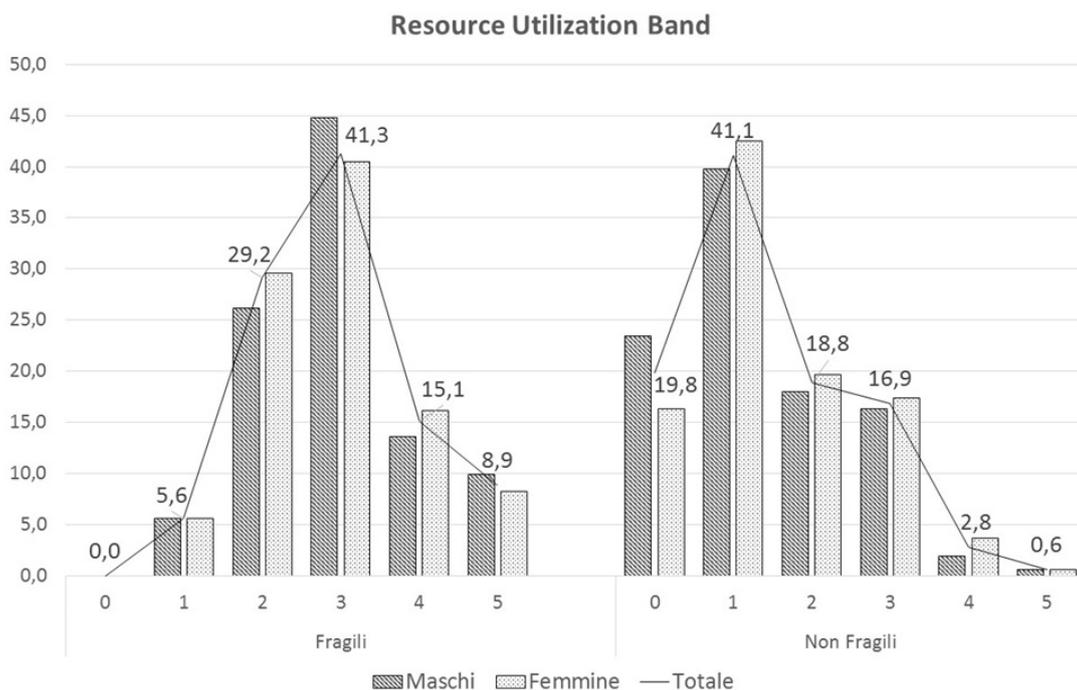


Figura 2.6: Confronto tra utenti fragili e non fragili per sesso e RUB, in percentuali standardizzate per età

C'è una percentuale, seppur minima, di utenti fragili che risultano anche utilizzatori sani: si tratta di utenti con diagnosi legate a problemi di nutrizione.

Questa tipologia di diagnosi associata alla fragilità che compone il Medically Frail Condition Marker crea una sorta di sottogruppo all'interno degli utenti fragili, che si compone di persone con età media più bassa, ma soprattutto caratterizzate da un ridotto consumo di risorse.

2.7.4 Expanded Diagnosis Clusters

I raggruppamenti denominati *Expanded Diagnosis Clusters (EDC)* sono 269 e vengono specificati da categorie di diagnosi (definite dai codici di diagnosi ICD9, ICD10, ICPC) che descrivono una patologia o patologie correlate in base a un organo o a un apparato. Ogni codice diagnostico è legato ad uno e un solo codice EDC. Il criterio alla base di questi raggruppamenti fa riferimento ad un approccio più epidemiologico e non più gestionale legato al consumo di risorse: vengono selezionate sub-popolazioni affette da specifiche patologie. In pratica gli utenti vengono raggruppati per EDC in modo che, all'interno di ogni raggruppamento, vi siano soggetti con malattie e sintomi simili. Dato che si tratta di una classificazione che fa riferimento alle patologie dell'individuo, le categorie in questione non sono mutuamente esclusive (ad una stessa persona possono essere associati più EDC).

Major EDC	Cod. EDC	Descrizione EDC	Prevalenza negli utenti fragili
Disturbi del sistema cardiovascolare	CAR05	Scompenso cardiaco congestizio	14.8%
	CAR09	Aritmia cardiaca	17.4%
	CAR14	Ipertensione senza gravi complicazioni	27.4%
Disturbi del sistema respiratorio	RES10	Insufficienza respiratoria	9.9%
	RES04	Enfisema, bronchite cronica, broncopneumopatia cronica ostruttiva	8.1%
Problemi nutrizionali	NUT01	Ritardo dell'accrescimento	21.2%
	NUT02	Carenze nutrizionali	4.1%
	NUT03	Obesità	24.2%
Disturbi neurologici	NUR06	Morbo di Parkinson	2.9%
	NUR11	Demenza e delirio	33.2%

Tabella 2.4: Esempi di Major EDC con alcuni dei rispettivi EDC

Gli Expanded Diagnosis Clusters possono essere aggregati in 27 macrocategorie chiamate MEDCs (Major EDCs) in base all'apparato a cui fa riferimento la specifica patologia. In tabella 2.4 Sono riportati alcuni esempi di Major EDC particolarmente diffusi tra gli utenti fragili, con alcune declinazioni in termini di EDC più specifiche e con le prevalenze in percentuale rispetto alla popolazione fragile dell'Ulss 15 nel 2013.

Alcuni codici MEDC (Major Expanded Diagnosis Clusters) coinvolgono maggiormente gli utenti fragili dell'Ulss 15 nel 2013, infatti circa la metà di queste persone è soggetta a disturbi del sistema cardiovascolare, il 51% ha disturbi neurologici e il 49% è coinvolto in problemi nutrizionali. Il 98% dei 1531 fragili dell'Ulss 15 è interessato da almeno uno di questi tre codici e il 34% dei 1531 fragili presenta sia disturbi del sistema circolatorio, sia neurologici.

Anche osservando le percentuali standardizzate per età (figura 2.7) si nota come questa suddivisione della popolazione totale sia rilevante anche in riferimento alle patologie organo correlate, infatti, al netto dell'età, le differenze tra le percentuali relative alla prevalenza di queste classificazioni di disturbi nei fragili e nei non fragili sono significative (livello di significatività pari a 0.01, test sulla differenza tra due proporzioni). Mentre la percentuale standardizzata per età di utenti fragili che soffrono di problemi nutrizionali (in figura 2.7) è superiore a quella non standardizzata (in tabella 2.4) perchè si tratta di un disturbo che coinvolge utenti in età particolarmente giovani, quelle relative ai disturbi neurologici e del sistema cardiovascolare risultano ridimensionate perchè presenti in larga parte tra gli anziani. Per problemi nutrizionali si intendono sia le carenze nutrizionali e i ritardi nella crescita, sia l'obesità.

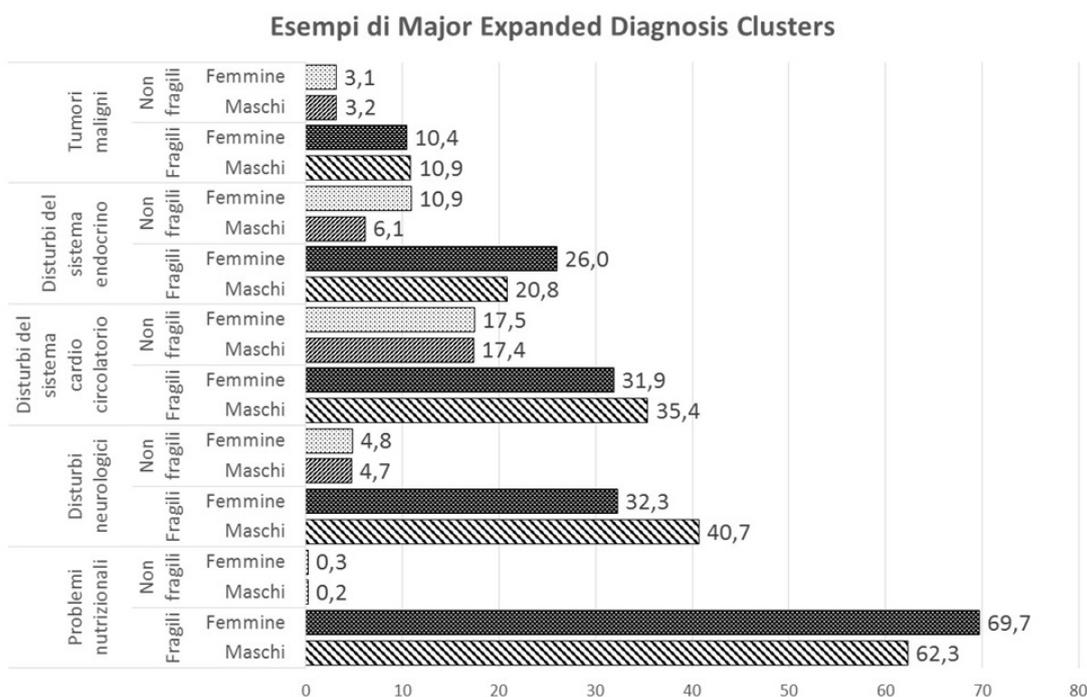


Figura 2.7: Esempi di MEDC, confronto tra utenti fragili e non fragili per sesso e MED, in percentuali standardizzate per età

2.7.5 Rx-Defined Morbidity Groups

Gli *Rx-Defined Morbidity Groups* (Rx-MG) sono gruppi di diagnosi farmaco-correlate, si ricavano dai flussi relativi al consumo di farmaci. Sono costituiti da 64 categorie che raggruppano farmaci proxy di diagnosi cliniche. Dato che uno stesso individuo può essere soggetto a più di una patologia, questi raggruppamenti non sono mutualmente esclusivi. Inoltre queste 64 categorie possono essere ulteriormente raggruppate in 19 MAJOR Rx-MGs in base al settore nosologico di appartenenza. In tabella 2.5 sono elencati alcuni Major Rx-MG con esempi dei relativi Rx-MG più specifici e le percentuali di prevalenza di questi nella popolazione dei fragili dell'Usls 15 nel 2013.

Per quanto riguarda l'utilizzo di farmaci, gli utenti fragili sono caratterizzati da alti consumi di quelli per disturbi cardiovascolari, il 56% ne ha fatto uso almeno una volta nel 2013, di farmaci per le infezioni, utilizzati da più di metà dei fragili, mentre il 46% ha consumato farmaci per disturbi gastrointestinali.

Dal confronto in figura 2.8 si può cogliere un consumo di farmaci più intenso per quanto riguarda i soggetti fragili, con una grande diffusione, come già preannunciato, di terapie per disturbi cardiovascolari. Le percentuali relative all'utilizzo di medicinali in tabella sono molto diverse per fragili e non fragili (test sulla differenza tra due proporzioni significativo allo 0.01); inoltre i soggetti fragili si

Major Rx-MG	Cod. Rx-MG	Descrizione Rx-MG	Prevalenza negli utenti fragili
Farmaci per disturbi cardiovascolari	CARx010	Malattie croniche	9.7%
	CARx020	Scoppenso cardiaco congestizio	26.7%
	CARx030	Pressione alta	40.6%
	CARx050	Patologie vascolari	42.1%
Farmaci per disturbi del sistema endocrino	ENDx030	Diabete con Insulina	7%
	ENDx050	Malattie della tiroide	7.4%
Farmaci per infezioni	INFx020	Acute, lievi	54%
	INFx020	Acute, grave, principale	9.6%

Tabella 2.5: Esempi di Major Rx-MG con alcuni dei rispettivi Rx-MG.

caratterizzano anche per un consumo maggiore di farmaci che, molto spesso, sono i cosiddetti “*salvavita*” o che, essendo legati a malattie croniche, devono essere assunti quotidianamente o regolarmente.

Generalmente i flussi relativi ai farmaci vengono utilizzati come “traccianti” di alcune particolari patologie difficili da rilevare (come ad esempio la broncopneumopatia cronica ostruttiva); infatti alcune patologie conducono raramente al ricovero (e quindi all’assegnazione della diagnosi nel flusso informativo delle Schede di Dimissione Ospedaliera) e spesso vengono sottostimate. Un uso alternativo di queste informazioni riguardanti il consumo di medicinali è quello di creare una variabile proxy della complessità delle condizioni patologiche dell’individuo: si suppone quindi che il consumo di più farmaci appartenenti a diverse categorie nosologiche corrisponda a condizioni di salute multiproblematiche e/o precarie. Il 36% dei soggetti fragili utilizza farmaci provenienti da 5 o più categorie Major Rx-MG.

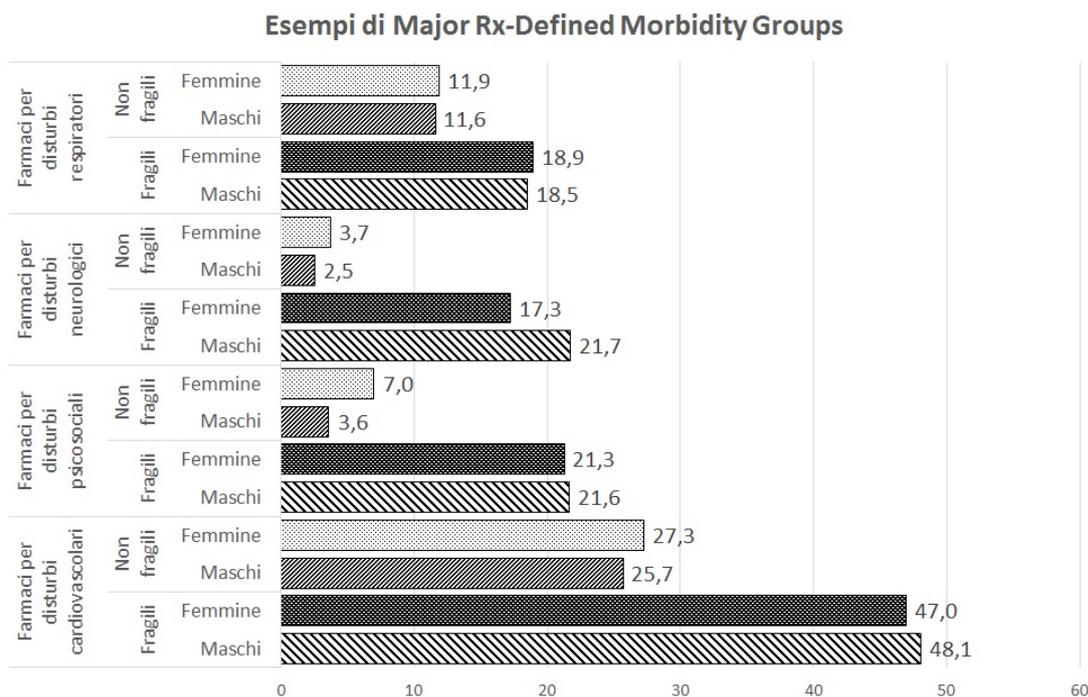


Figura 2.8: Esempi di Major Rx-MG, confronto tra utenti fragili e non fragili per sesso e Major Rx-MG, in percentuali standardizzate per età.

2.7.6 Condizioni croniche

Per ciascun assistito viene segnalata la presenza o meno di alcune condizioni croniche ad alta prevalenza e la fonte in base alla quale la presenza di tale condizione è stata messa in luce. Queste patologie, infatti, spesso non sono facili da identificare utilizzando i flussi informativi di gestione e possono essere individuate in base alle diagnosi, ai farmaci consumati o entrambi. Le diagnosi possono essere raccolte dalle SDO, nel caso si sia verificato un ricovero, dal flusso del pronto soccorso, della psichiatria territoriale, delle esenzioni, dell'assistenza domiciliare, dalle UVMD o dai registri delle malattie rare. Proprio a causa della difficoltà ad identificare i pazienti affetti da queste patologie, per avere abbastanza elementi per determinare con un buon grado di precisione la presenza o l'assenza di una cronicità, il sistema ACG utilizza le Schede di Dimissione Ospedaliera dei 5 anni precedenti a quello di riferimento.

Queste variabili si rivelano utili per identificare sub-popolazioni di interesse, che presentano esigenze particolari, e per analizzare il carico assistenziale relativo alle patologie croniche. Le patologie croniche evidenziate dal sistema di ACG sono: *disordine bipolare, scompenso cardiaco congestizio, depressione, diabete, disordine del metabolismo lipidico, glaucoma, tumore maligno, HIV, ipertensione arteriosa, ipotiroidismo, trapianto/immunodepressione, lombalgia, cardiopatia ischemica,*

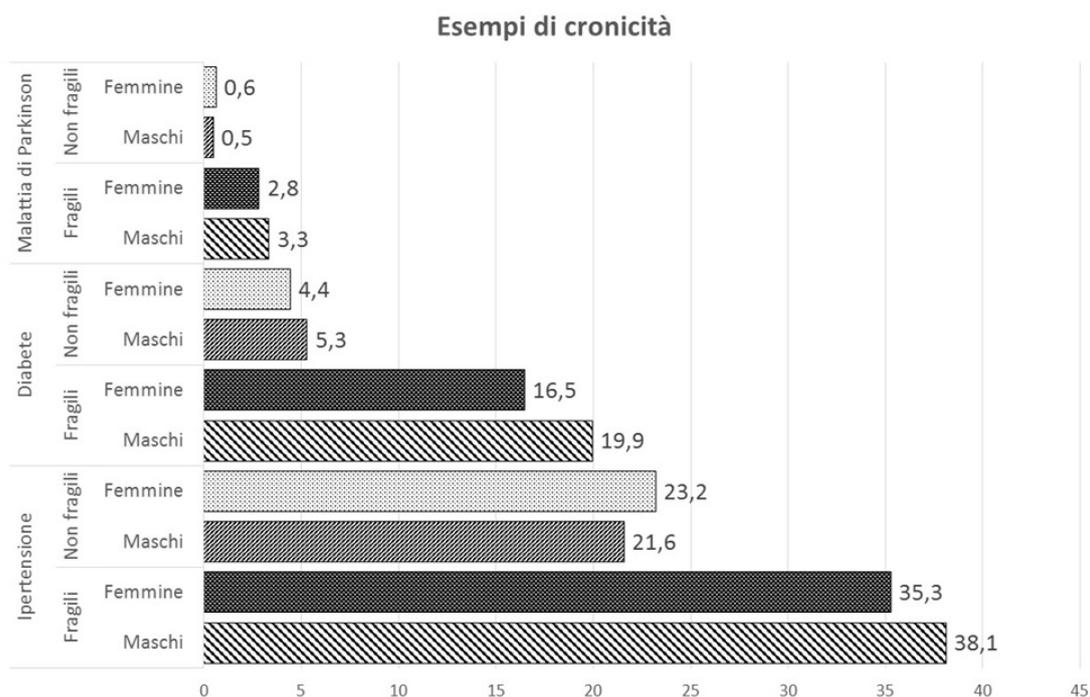


Figura 2.9: Esempi di alcune malattie croniche, confronto tra utenti fragili e non fragili per sesso e malattie croniche, in percentuali standardizzate per età

osteoporosi, morbo di Parkinson, asma persistente, artrite reumatoide, schizofrenia, disturbi epilettici, degenerazione maculare legata all'età, broncopneumopatia cronica ostruttiva (BPCO) e insufficienza renale cronica.

Come più volte si è affermato in questo capitolo, i soggetti fragili spesso hanno almeno una malattia cronica; come si può vedere dalla figura 2.9, questa caratteristica viene confermata dall'osservazione delle percentuali standardizzate per età relative alla prevalenza di certe cronicità in soggetti fragili e non fragili: le percentuali sono significativamente più alte per gli utenti fragili (test sulla differenza tra due proporzioni significativo allo 0.01).

2.8 Follow-up degli utenti fragili nel 2013

Per valutare l'andamento delle condizioni di salute dei soggetti etichettati come fragili dal Medically Frail Condition Marker nel 2013 si è effettuato il record linkage con il database di ACG relativo ai fragili nel 2014, ancora in versione provvisoria, e con l'Anagrafe Sanitaria del 2014 (che fornisce le informazioni riguardanti i decessi nel 2014).

I record linkage sono stati effettuati sfruttando l'univocità del codice fiscale, nelle fonti collegate è presente un record per ogni individuo, quindi è stato possibile

effettuare un record linkage deterministico con rapporto 1 a 1. La percentuale di successo per il collegamento del dataset ACG 2013 con quello relativo al 2014 è pari al 96.9%, mentre quella ottenuta nel linkage con l'Anagrafe Sanitaria del 2014 è pari al 95.1%.

Osservando la tabella 2.6 si può notare che quasi un quarto dei fragili evidenziati nel 2013 è deceduto entro il 31/12/2014, inoltre circa la metà è ancora parte della popolazione fragile nel 2014, ma quasi il 20% non è più fragile. La fragilità secondo ACG è quindi una condizione transitoria che può regredire e non è definitiva.

Esiti possibili per gli utenti fragili Percentuale di utenti fragili dell'Ulss 15 nel 2013

Deceduti nel 2013	15.9%
Deceduti nel 2014	8.8%
Assenti per altri motivi nel 2014	4.4%
Fragili nel 2014	51.3%
Non più fragili nel 2014	19.3%

Tabella 2.6: Follow-up degli utenti fragili nel 2013

2.9 Validazione del Marker per la fragilità di ACG

A questo punto risulta necessario stabilire se il marcatore creato dal sistema ACG sia effettivamente in grado di individuare i soggetti fragili e se le fragilità che rileva siano coerenti con le necessità di previsione e gestione degli utenti e delle prestazioni dell'Ulss 15.

Da quando il sistema ACG è stato sviluppato, sono state svolte diverse analisi per validarne gli output. La maggior parte di queste però sono state condotte con lo scopo di verificare la capacità previsiva in termini di risorse consumate per scopi gestionali (Reid *et al.*, 1999, Zielinski *et al.*, 2009).

Dato che l'interesse principale è rivolto ai soggetti fragili, sembra più appropriata una validazione del Medically Frail Condition Marker come strumento per prevedere ed identificare quali siano i soggetti più bisognosi di cure e più a rischio rispetto a prognosi sfavorevoli (Sternberg *et al.*, 2012). Lo studio israeliano, condotto da Sternberg ed i suoi colleghi, sfrutta il fatto che le persone fragili abbiano un rischio maggiore di ospedalizzazione, di morte e di declino funzionale. Inserendo tra le covariate di un modello di regressione logistica che ha come variabile dipendente la realizzazione degli eventi negativi sopracitati, sono state valutate le capacità previsive dei due indicatori di fragilità.

Per valutare la bontà del Medically Frail Condition Marker del sistema ACG verrà di seguito utilizzato un metodo molto simile con i dati relativi all'Ulss 15 nell'anno 2013 estratti dal sistema di ACG con alcune informazioni ottenute dai flussi informativi: Schede di Dimissione Ospedaliera (SDO) del 2012, 2013 e 2014, Anagrafe Sanitaria del 2013 e 2014 e psichiatria territoriale, assistenza domiciliare, pronto soccorso e farmaceutica territoriale del 2013.

Si stima quindi una regressione logistica utilizzando come variabile dipendente una dicotomica che rappresenta gli esiti negativi associati alla fragilità: vale 1 se l'utente ha avuto un ricovero d'urgenza o è morto nel 2014 e 0 altrimenti. Il modello è stato stimato su un campione di 26010 persone: 13005 con valore unitario nella variabile dipendente e altrettante selezionate casualmente tra gli utenti in vita e che non hanno avuto ricoveri d'urgenza nel 2014.

Si è stimato un primo modello inserendo soltanto il Medically Frail Condition Marker tra le covariate. Il coefficiente relativo a quest'unica covariata risulta significativo con un *odds ratio* pari a 13.9, tuttavia il potere predittivo del modello risulta piuttosto scarso, infatti predice e classifica in modo corretto soltanto il 48% delle unità statistiche del campione. In più la sensibilità è molto bassa (4%) e la percentuale di falsi negativi è pari a 53%. Ciò comporta che questo modello indentifica solo il 4% degli individui per cui $y=1$.

Per migliorare la bontà del modello si sono inserite alcune covariate dicotomiche (con valore unitario in caso affermativo) provenienti dai flussi informativi sopracitati:

- aver ricevuto assistenza domiciliare almeno una volta nel 2013
- avere un Charlson index² (indice relativo alle comorbidità) superiore a uno in almeno uno dei ricoveri degli ultimi due anni (2012 e 2013)
- avere una poliprescrizione negli ultimi tre mesi del 2013 (almeno 4 farmaci appartenenti a macrocategorie diverse)
- avere almeno una condizione di invalidità nel 2013
- aver avuto almeno un ricovero nel 2013 (esclusi quelli con traumi in prima diagnosi)
- aver effettuato almeno un accesso al pronto soccorso nel 2013
- aver effettuato almeno un accesso alla psichiatria territoriale nel 2013
- aver avuto almeno tre ricoveri con diversa prima diagnosi negli ultimi due anni (2012 e 2013)
- aver avuto diagnosi di tumori maligni negli ultimi due anni (2012 e 2013)
- aver avuto diagnosi di scompenso cardiaco congestizio negli ultimi due anni (2012 e 2013).

In più sono state inserite due covariate riguardanti sesso (dicotomica, 1 se maschio) ed età (con tre classi: meno di 20 anni, tra 21 e 69 anni e con più di 70). Modelli così costruiti verranno ulteriormente approfonditi nel prossimo capitolo.

Il modello così stimato (tabella 2.7) presenta caratteristiche migliori (area sotto la curva ROC pari a 0.762): classifica correttamente il 69% degli utenti, il punto di equilibrio tra sensibilità e specificità si trova ad un livello di probabilità pari a 0.4, presso il quale la sensibilità è pari a 65.7%, la specificità è pari a 72.3% e la percentuale di falsi negativi, che in queste circostanze serve sia sufficientemente ridotta in modo da identificare il più possibile degli utenti che subiscono outcome legati alla fragilità, è pari a 32.1%. Dato che entrambi gli eventi utilizzati per determinare la variabile dipendente possono essere ricondotti anche

²Il Charlson Comorbidity Index è un indice che assegna un peso alla comorbidità (considerando le diagnosi assegnate nei ricoveri, dalla seconda alla sesta) sulla base del rischio relativo di morte delle specifiche patologia incluse.

a cause accidentali, oltre che a delle condizioni di fragilità, tale valore sembra accettabile.

Variabile	Odds Ratio	IC 95%
Fragilità per ACG	1,531**	1,035-2,264
Sesso (maschio vs femmina)	0,687***	0,647-0,729
Assistenza Domiciliare Integrata	2,039***	1,622-2,562
Charlson Index	1,863***	1,501-2,313
Poliprescrizione	1,63***	1,489-1,785
Invalidità	2,102***	1,865-2,370
Ricovero	1,717***	1,543-1,910
Pronto Soccorso	1,617***	1,494-1,749
Psichiatria Territoriale	2,09***	1,703-2,566
Giovane (rispetto ad Adulto)	0,47***	0,430-0,515
Anziano (rispetto ad Adulto)	2,563***	2,351-2,794
3 + Diagnosi	1,442**	1,091-1,906
Tumore Maligno	1,516***	1,182-1,944
Livello di probabilità	0,4	
Corretti	69	
Sensibilità	65,7	
Specificità	72,3	
Falsi Positivi	29,8	
Falsi Negativi	32,1	
Area ROC	0,762	
R_2	0.2093	
Pseudo R_2	0.2791	

Tabella 2.7: Regressione logistica, con campione equilibrato rispetto alla variabile risposta (morte o ricovero d'urgenza nell'anno successivo rispetto a quello di riferimento). (Legenda per i livelli di significatività: ***=0.01, **=0.05, *=0.1)

Il coefficiente relativo al Medically Frail Condition Marker risulta comunque significativo allo 0.01, ma con un effetto in termini di *odds ratio* più contenuto rispetto al modello precedentemente stimato, pari a 1.53 (con intervallo di confidenza: 1,035-2,264). Quindi una persona classificata come fragile dal sistema di ACG ha il 53% di probabilità in più in media di incorrere in ricoveri d'urgenza o decessi, al netto delle altre condizioni considerate nel modello.

2.10 Conclusioni

Per concludere, si può affermare che il sistema di gestione di dati ACG si può senza dubbio rivelare utile per la gestione e la programmazione di spese e di interventi nell'Ulss 15. Tuttavia il Medically Frail Condition Marker, pur riconoscendo alcuni soggetti fragili, non è da solo sufficiente a predire le prognosi sfavorevoli legate alla condizione di fragilità.

Si può anche supporre che, con i dati inseriti come input dalle realtà italiane, il Medically Frail Condition Marker sia al momento incompleto dato che, secondo la tabella relativa ai concetti legati alla fragilità utilizzati per la costruzione dell'indice, alcuni campi non sono disponibili o lo sono, ma soltanto con informazioni parziali (per esempio la povertà). Non è invece possibile stabilire se e quanto questa potenziale distorsione sia rilevante dato che le indicazioni specifiche su come la condizione di fragilità venga attribuita non sono state messe a disposizione alle Ulss e non sono state rese pubbliche.

Risulta quindi necessario trovare uno strumento che, con i dati a disposizione, consenta di identificare un gruppo di persone fragili e che risulti utile ad effettuare interventi preventivi sulla popolazione.

Capitolo 3

Un modello predittivo per le prognosi negative legate alla fragilità

3.1 Introduzione

In questo capitolo ci si propone di esplorare la possibilità di identificare i soggetti fragili in base alla loro esposizione al rischio di incorrere nelle prognosi negative legate alla fragilità, che in letteratura sono la morte e il ricovero d'urgenza nell'anno successivo a quello di osservazione. L'esperienza di riferimento per analisi di questo tipo in Italia è quella del *Servizio Valutazione Strategica dell'Azienda USL di Ravenna* (Falasca *et al.*, 2011). Il Modello Statistico Combinato (MoSaiCo), elaborato da Falasca e collaboratori, è un modello multivariato di regressione logistica che prende spunto e cita il *Combined Predictive Model* sviluppato dal gruppo di ricerca di King's Fund (Health Dialog, 2006), adattandolo alla realtà italiana.

Partendo proprio dal modello MoSaiCo dell'Ulss di Ravenna, lo scopo di questo capitolo è quello di elaborare una proposta efficace, ma più parsimoniosa e che sia in grado di identificare un numero ridotto di variabili fondamentali per riconoscere gli individui fragili nell'Ulss 15.

Le analisi relative a questo capitolo sono state eseguite utilizzando il software SAS 9.3 per creare ed analizzare un dataset frutto di un record linkage tra diverse fonti informative a disposizione dell'Ulss 15.

3.2 Il modello MoSaiCo

Il modello MoSaiCo dell'Ulss di Ravenna utilizza i dati di tre anni consecutivi, di cui i primi due per le variabili esplicative, mentre il terzo anno fornisce la variabile risposta relativa alle eventuali prognosi negative legate alla fragilità che coinvolgono gli utenti in osservazione. I flussi informativi utilizzati sono: le schede di dimissione ospedaliera (SDO), le esenzioni dai pagamenti delle prestazioni sanitarie, gli accessi al Pronto Soccorso (PS), l'Anagrafe Sanitaria, i dati relativi alla Farmaceutica Territoriale, l'elenco delle prestazioni erogate di Assistenza Domiciliare Integrata (ADI), il flusso della Psichiatria Territoriale e altre due fonti provenienti dai servizi sociali.

Le informazioni anagrafiche raccolte per il modello MoSaiCo sono l'età, il sesso e la cittadinanza; le esenzioni per invalidità (dal flusso informativo delle esenzioni) e i servizi sociali primari erogati per gli over settantacinquenni, dei quali però non si dispone nell'Ulss 15. Per coloro che hanno almeno un ricovero, vengono raccolti i codici ICD9 relativi alle diagnosi (dalla prima alla sesta), inoltre per coloro che hanno usufruito di ospedale e pronto soccorso, vengono registrate le variabili riguardanti il numero delle prestazioni erogate rispetto ai tempi e alla frequenza di erogazione; in particolare sono di interesse i ricoveri d'urgenza e gli accessi al pronto soccorso tramite ambulanza. I flussi relativi al consumo di farmaci vengono sfruttati per identificare gli utenti che fanno uso di almeno quattro farmaci provenienti da macrocategorie ¹ diverse negli ultimi tre mesi (poliprescrizione). Sono presenti inoltre alcune variabili provenienti dai servizi sociali, quali "Assistito dai servizi sociali", "Anziano solo" e "Coppia di anziani soli".

La variabile risposta (y) è una variabile dicotomica che vale uno se l'individuo è morto nel 2014 o è stato ricoverato d'urgenza nel 2014, 0 altrimenti. Per ricovero d'urgenza si intende un ricovero che abbia nella tipologia di ammissione o "Ricovero d'urgenza" o "Trattamento sanitario obbligatorio".

Prima di eseguire l'analisi logistica del modello Mosaico, il Servizio di Valutazione Strategica dell'Ulss di Ravenna ha scremato le variabili per utilizzare soltanto quelle con frequenza $> 5\%$, passando quindi da 57 a 37; in seguito il

¹Le macrocategorie sono distinte da diversi codici di *Anatomical Therapeutic Chemical classification system* (ATC) al livello di specificazione 3. Il codice ATC è una codifica internazionale e sistematica dei farmaci in commercio che crea una classificazione gerarchica a seconda dei principi attivi. Il primo livello, contrassegnato da una lettera dell'alfabeto, indica l'apparato di riferimento; il secondo, contrassegnato da due cifre, indica il gruppo terapeutico principale; e il terzo, declinato da una lettera, contraddistingue il sottogruppo terapeutico farmacologico. Vi sono ulteriori sottogruppi, il quarto e il quinto livello, in base ai sottogruppi chimici del farmaco.

criterio di selezione stepwise utilizzato nell'analisi ne ha inserite 31. Infatti il modello MoSaiCo è stato stimato come un modello logistico tramite una selezione stepwise delle variabili con criterio di entrata e di uscita pari a 0.15. La misura dell'area sotto la curva ROC per il modello MoSaiCo ottenuto dall'Ulss di Ravenna è 0.774. In tabella 3.1 sono riportate le stime degli Odds Ratio relative a questo modello per l'Ulss di Ravenna, con i relativi intervalli di confidenza e livelli di significatività.

Variabili	OR	IC 95%	p
Età tra i 18 e i 64 anni (Riferimento)	1	-	-
Età tra i 65 e gli 84 anni	2,42	2,28-2,57	<0,0001
Età ultra85	5,69	5,19-6,24	<0,0001
Maschio (vs Femmina)	1,07	1,02-1,12	0,0081
Straniero (vs Italiano)	1,26	1,10-1,45	0,0011
Charlson Index²	1,46	1,33-1,6	<0,0001
1 o più ricoveri d'emergenza negli ultimi 30 giorni	2,2	1,85-2,63	<0,0001
1 o più ricoveri d'emergenza negli ultimi 30-90 giorni	1,25	1,09-1,44	0,0011
1 o più ricoveri d'emergenza negli ultimi 90-180 giorni	1,46	1,29-1,65	<0,0001
1 o più ricoveri d'emergenza negli ultimi 180-365 giorni	1,21	1,07-1,36	0,0245
1 o più ricoveri d'emergenza negli ultimi 365-730 giorni	1,4	1,29-1,53	<0,0001
Invalidità	1,71	1,58-1,85	<0,0001
Over-75 che vive solo	1,44	1,33-1,56	<0,0001
Coppia di over-75	1,36	1,27-1,46	<0,0001
Assistiti dai servizi sociali nel 2007	1,58	1,42-1,76	<0,0001
Accesso alla psichiatria territoriale negli ultimi due anni	1,67	1,45-1,92	<0,0001
Accesso all'Assistenza domiciliare integrata negli ultimi due anni	1,47	1,29-1,67	<0,0001

Continua nella prossima pagina

²Il Charlson Comorbidity Index è un indice che assegna un peso alla comorbidità (considerando le diagnosi assegnate nei ricoveri, dalla seconda alla sesta) sulla base del rischio relativo di morte delle specifiche patologia incluse.

Continua dalla pagina precedente

Variabili	OR	IC 95%	p
Poliprescrizione farmaceutica	1,59	1,45-1,74	<0,0001
Visite strumentali presso il pronto soccorso negli ultimi 90-180 giorni	1,41	1,28-1,55	<0,0001
1 o più accessi al pronto soccorso tramite ambulanza del 118 negli ultimi 30-90 giorni	1,55	1,31-1,82	<0,0001
Ricovero (esclusi quelli con trauma in prima diagnosi) negli ultimi 30-90 giorni	1,53	1,36-1,71	<0,0001
Ricovero (esclusi quelli con trauma in prima diagnosi) negli ultimi 365-730 giorni	1,21	1,13-1,29	<0,0001
Un accesso al Pronto Soccorso negli ultimi 365-730 giorni	1,27	1,18-1,37	<0,0001
Due accessi al Pronto Soccorso negli ultimi 365-730 giorni	1,48	1,31-1,67	<0,0001
Tre o più accessi al Pronto Soccorso negli ultimi 365-730 giorni	1,77	1,5-2,09	<0,0001
Ricoveri con diagnosi di malattia mentale (negli ultimi due anni)	1,51	1,27-1,79	<0,0001
3 o più ricoveri con diverse diagnosi principali (negli ultimi due anni)	1,13	1,01-1,28	0,0036
Ricoveri con coronaropatie (negli ultimi due anni)	1,16	1,03-1,30	<0,0001
Ricoveri con scompenso cardiaco congestizio (negli ultimi due anni)	1,15	1,02-1,30	<0,0001
Ricoveri con tumore maligno (negli ultimi due anni)	1,63	1,47-1,81	0,0036
Ricoveri con broncopneumopatia cronica ostruttiva (BPCO) (negli ultimi due anni)	1,24	1,08-1,43	<0,0001

Tabella 3.1: Modello MoSaiCo, regressione logistica con morte o ricovero d'urgenza nell'anno successivo a quello di riferimento come variabile risposta. Stime degli Odds Ratio (OR) ottenute, con rispettivi intervalli di confidenza al 95% (IC 95%) e livello di significatività (p). Coefficienti calcolati per l'Ulss di Ravenna. Traduzione da Falasca *et al.*, 2011.

Come si può osservare dalla tabella, vengono esclusi dall'analisi i soggetti con

meno di 18 anni, mentre le classi di età utilizzate nell'analisi sono tre: tra i 18 e i 64 anni, utilizzata come riferimento, tra i 65 e gli 84 anni e con più di 85 anni. Relativamente al modello MoSaiCo, si possono fare alcune riflessioni, alcune critiche e proporre alcuni suggerimenti:

- Nell'Ulss di Ravenna viene utilizzata tutta la popolazione per l'analisi, indipendentemente dal numero di individui per cui $y=1$ (casi) e $y=0$ (controlli). Infatti la base di dati utilizzata dal modello MoSaiCo non è bilanciata, questo comporta che il modello, in particolar modo i livelli di probabilità predetti risultino poco equilibrati. Tuttavia, nel caso in cui sia necessario scegliere un livello di probabilità per individuare i soggetti fragili, è preferibile lavorare con matrici di dati in cui vi siano quantità di casi e di controlli quantomeno confrontabili. Il suggerimento in questo caso è quello di costituire un campione contenente tutti i casi e un campione casuale di pari numerosità dei controlli.
- Molte variabili inserite nel modello vengono suddivise in 5 periodi temporali: gli ultimi trenta giorni, gli ultimi 30-90 giorni, gli ultimi 90-180 giorni, gli ultimi 180-365 giorni e gli ultimi 365-730 giorni. Questa suddivisione è presente per tutte le variabili utilizzate dal modello nell'Ulss di Ravenna prima di selezionarle e scremarle, infatti molte di queste non sono riportate nella tabella 3.1 perchè non significative o con frequenze troppo basse. Vengono utilizzate con questa suddivisione anche nel modello sviluppato dal gruppo di ricerca di King's Fund (Health Dialog, 2006). La specificazione temporale delle variabili porta ad avere un numero di covariate molto ampio, senza una reale necessità pratica. Infatti a livello sostantivo, non esiste una particolare differenza rispetto a subire un evento 30-90 giorni prima o 90-180 giorni prima, tanto più che in questo caso gli intervalli temporali sono fissi nell'anno di riferimento e non sono legati alla data in cui accade l'evento corrispondente alla variabile dipendente. Inoltre, osservando gli odds ratio relativi alle 5 variabili corrispondenti ad aver subito uno o più ricoveri d'emergenza stimati nel modello MoSaiCo dell'Ulss di Ravenna, figura 3.1, si può notare che, eccezion fatta per il primo intervallo di tempo (ultimi trenta giorni), i rimanenti quattro odds ratio non sono significativamente diversi tra loro (i valori degli intervalli di confidenza si intersecano e includono anche gli odds ratio attribuiti ad altre variabili). Sarebbe opportuna una riorganizzazione delle covariate da includere nel modello secondo un criterio sostantivo e di significato pratico, prima che basato sulla significatività statistica o su una selezione stepwise.

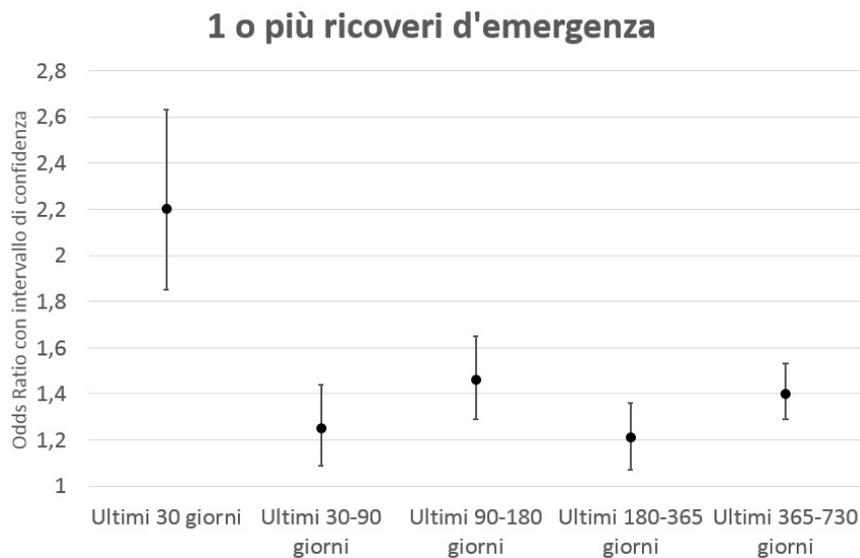


Figura 3.1: Odds Ratio delle cinque variabili associate ad aver avuto almeno un ricovero d'urgenza stimati nel modello MoSaiCo dell'Ulss di Ravenna, con relativi intervalli di confidenza.

- Come si è visto nel capitolo precedente, la fragilità può esistere anche nei minori, quindi escluderli totalmente dal modello sembra una semplificazione pericolosa.
- Le classi di età utilizzate sembrano essere piuttosto arbitrarie e vengono introdotte nel modello senza una spiegazione adeguata. Tuttavia età e sesso sono fondamentali nell'individuare i soggetti fragili, infatti da soli sono in grado di prevedere la morte o il ricovero d'urgenza nell'anno successivo con un buon grado di precisione. Per dimostrare quanto detto si è stimata una regressione logistica su un campione formato da 11130 casi ($y=1$) e altrettanti controlli ($y=0$) di utenti dell'Ulss 15. La variabile risposta utilizzata è costituita dalle prognosi negative legate alla fragilità (morte o il ricovero d'urgenza nell'anno successivo a quelli di riferimento) e come covariate sono state inserite nel modello soltanto età (in anni) e sesso. Queste sole due variabili danno origine ad un modello che presenta un'area sotto la curva ROC pari a 0.7. Dato che, come si è visto, età e sesso hanno questo ruolo determinante è bene che vi sia maggiore precisione e sistematicità anche nella scelta delle classi di età da utilizzare nell'analisi.
- Essendo l'età così legata alla fragilità, ci si chiede se sia ragionevole svolgere le analisi separatamente per classi di età. In questo modo si metterebbero in evidenza insiemi di fattori di rischio eventualmente diversi a seconda dell'età e degli individui.

Date queste osservazioni si procede con le analisi prendendo spunto dal modello MoSaiCo appena descritto, ma con alcuni accorgimenti:

- si stima il modello su un campione equilibrato che comprenda tutte le fasce d'età, anche i minori;
- è necessario riorganizzare le variabili per essere più parsimoniosi e perchè quelle incluse siano più rilevanti;
- si ricercano classi di età più efficienti, identificate mediante un'analisi sistematica;
- si prosegue con analisi distinte per classi di età, per valutare se tale approccio sia valido e realmente necessario.

3.3 Il record linkage delle fonti

Per poter replicare il modello Mosaico è stato necessario creare un dataset che contenesse informazioni provenienti da fonti diverse, sia da un punto di vista cronologico (gli anni di riferimento sono in tutto 3), sia da un punto di vista qualitativo (le tipologie di fonti diverse sono in tutto otto: Anagrafe Sanitaria, Schede di Dimissione Ospedaliera, Pronto Soccorso, Esenzioni, Psichiatria Territoriale, Assistenza Domiciliare Integrata, flussi dei farmaci e Anagrafica Sanitaria georeferenziata con rispettivi indici di deprivazione).

Per effettuare il collegamento tra le tabelle si è utilizzato un criterio deterministico sfruttando il codice fiscale degli utenti, in quanto univoco e presente in tutte le fonti utilizzate. In più, in caso di errori o di incongruenze, il codice fiscale non è altro che una sintesi delle caratteristiche anagrafiche dell'individuo e in quanto tale si può verificare la corretta attribuzione con dei controlli a campione. Inoltre le percentuali di successo e le numerosità molto alte non hanno reso necessario il ricorso a ulteriori criteri per affinare il linkage. Si sono svolti infatti un totale di 14 record linkage con percentuali di successo tra 85.2% e 92.6% (99.2% togliendo i mancati collegamenti a causa di morte). Tenendo presente che si tratta di anni diversi e che potrebbero esserci dei mancati collegamenti a causa di morte o di trasferimento in entrata o in uscita si tratta di percentuali molto soddisfacenti.

Un primo passaggio si è reso necessario per unire i diversi anni appartenenti allo stesso flusso informativo e creare le variabili necessarie per il modello MoSaiCo, come ad esempio le malattie croniche per cui è necessario un follow-up delle Schede di Dimissione Ospedaliera (SDO) di due anni. In questo caso prima di unire le fonti

di dati aggiungendo nuove colonne ad un dataset di partenza, si è preferito agire sulle righe ordinate in base al codice fiscale, creando un dataset contenente un numero variabile di righe per ogni utente (0 se assente nel dataset, 1 se presente con soltanto una prestazione in uno dei due anni, più di uno nel caso vi fossero più prestazioni nei due anni considerati). Una volta calcolate le variabili di interesse, si sono condensate le informazioni rilevanti per ogni utente in una sola riga che è stata selezionata per formare un nuovo dataset contenente un record per ogni utente presente in almeno una delle tabelle originali. I dataset così formati, con codice fiscale univoco, sono stati utilizzati per il record linkage con i dati anagrafici, figura 3.2.

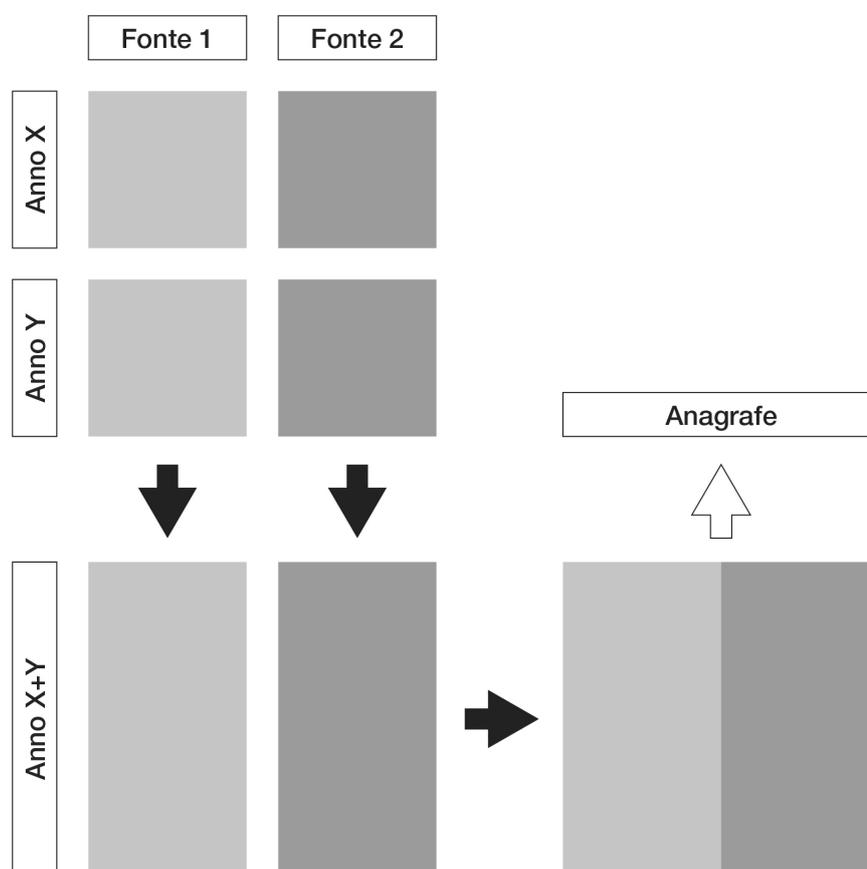


Figura 3.2: Passaggi necessari per effettuare il record linkage delle fonti con i dati anagrafici.

Nella fase di collegamento con i dati anagrafici, si è utilizzato un dataset come base (l'Anagrafe Sanitaria) e si è proceduto ai linkage in una logica di *left join*: il dataset risultante ha tante righe quante quelle del dataset di base, aggiungendo le colonne della tabella da collegare se in corrispondenza di identico codice fiscale. La tabella utilizzata come base è data da l'intersezione (*inner join*) tra l'Anagrafe Sanitaria relativa al 2014, infatti solo per questi utenti sono disponibili i dati

relativi all'eventuale decesso o ricovero d'urgenza nel 2014 necessari per la variabile risposta, e l'Anagrafe Sanitaria 2012. In questo modo si ottengono gli individui presenti per tutta la durata del periodo di osservazione, senza *attrition*³.

3.3.1 L'indice di deprivazione

L'indice di deprivazione è una misura sintetica della condizione economica e del disagio sociale degli abitanti di una determinata zona. Viene utilizzato di seguito l'indice calcolato sulla base dei dati del censimento del 2001 per le sezioni di censimento (Caranci *et al.*, 2010).

Questo indice si ottiene utilizzando alcuni indicatori calcolati per sezione di censimento:

x_1 : % di popolazione con istruzione pari o inferiore alla licenza elementare (mancato raggiungimento obbligo scolastico);

x_2 : % di popolazione attiva disoccupata o in cerca di prima occupazione;

x_3 : % di abitazioni occupate in affitto;

x_4 : % di famiglie monogenitoriali con figli dipendenti conviventi;

x_5 : densità abitativa (numero di occupanti per 100 m^2 nelle abitazioni).

L'indice di deprivazione è dato dalla somma degli scostamenti di questi indicatori di disagio socio-economico dalla media nazionale, secondo la formula

$$z_i = \frac{x_i - \mu_{x_i}}{\sigma_{x_i}}$$

dove μ_{x_i} e σ_{x_i} sono media e varianza rispetto al territorio italiano. Questi indicatori normalizzati vengono poi aggregati come $ID = \sum_{i=1}^5 z_i$.

L'indice viene poi classificato in quintili di popolazione per tutta l'Italia. Il livello "molto deprivato" è contrassegnato dal numero 5 ed è il livello di interesse per la codifica di questa variabile: la variabile relativa all'indice di deprivazione si pone pari a 1 se l'indice si colloca nel 20% di popolazione italiana più deprivata e 0 altrimenti.

In figura 3.3 sono riportate le sezioni di censimento dell'Ulss 15 con diversa colorazione a seconda dell'indice di deprivazione.

Per associare l'indice della sezione di censimento adatta con ogni persona, è stato necessario georeferenziare gli utenti presenti nell'Anagrafe Sanitaria ed

³L'attrition è la perdita nel tempo delle unità statistiche di un'analisi, dovuta a diverse cause.



Figura 3.3: Sezioni di censimento dell’Ulss 15 con diversa colorazione a seconda dell’indice di deprivazione.

effettuare una query geografica. In questo modo si è ottenuta una tabella con i codici fiscali, le coordinate della residenza, la sezione di censimento, l’indice di deprivazione e il quintile corrispondente. La georeferenziazione ha portato a risultati più che soddisfacenti, infatti i dati che non sono stati associati a delle coordinate geografiche rappresentano solo l’1.5% del totale della popolazione. A questo punto è stato possibile effettuare il record linkage dei dati come precedentemente descritto.

3.4 La riorganizzazione delle variabili

3.4.1 Nuove variabili con riferimento annuale

Partendo da una vasta gamma di variabili dicotomiche, costruite sulla base delle analisi relative al modello Mosaico, un primo importante passaggio è quello di ridurre il numero, secondo un criterio sostantivo. Infatti, data la grande numerosità del dataset, il rischio è quello di avere un numero eccessivo di variabili significative anche se non effettivamente rilevanti ai fini di prevedere e spiegare la fragilità. Questa prima scrematura delle variabili risulta quindi di fondamentale importanza in quanto non si basa sulla significatività statistica, ma su ragionamenti relativi all’interpretazione delle variabili. Il primo passaggio di riduzione delle variabili è stato infatti svolto, dopo ampia riflessione, in collaborazione con l’unità di Epidemiologia dell’Ulss 15 di Camposampiero.

In primo luogo si sono riportate le variabili ad un comune riferimento temporale, l'anno 2013, riducendone così il numero. Si sono quindi ottenute le seguenti variabili dicotomiche che tengono conto di quanto accaduto ai soggetti durante l'anno di riferimento (2013):

- Almeno un accesso al pronto soccorso;
- Almeno un ricovero, escludendo i ricoveri con un trauma in prima diagnosi;
- Beneficiario di assistenza domiciliare;
- Accesso ai servizi di salute mentale;
- Presenza di esenzioni per invalidità.

Si sono mantenute invece con un intervallo di due anni (2012 e 2013) le variabili relative alle patologie diagnosticate, soprattutto quelle che sono soggette a cronicità. Questo perché spesso sono difficili da individuare con periodi di osservazione brevi, dato che possono essere state diagnosticate antecedentemente rispetto all'intervallo temporale considerato. Ci sono due modi per stimare la prevalenza di queste malattie croniche senza averne una sistematica sottostima: si può analizzare il flusso dei farmaci, ma con errori e imprecisioni dato che non vi è corrispondenza perfetta e univoca tra farmaco e patologia; oppure si possono ricercare le diagnosi in questione nelle schede di dimissione ospedaliera, ma utilizzando intervalli temporali più ampi. Le variabili con riferimento biennale sono quindi:

- Ricovero con diagnosi di cardiopatia coronarica;
- Ricovero con diagnosi di scompenso cardiaco congestizio;
- Ricovero con diagnosi di tumore maligno;
- Ricovero con diagnosi di broncopneumopatia cronica ostruttiva;

La variabile relativa alla cittadinanza potrebbe essere fuorviante, infatti, chi non ha cittadinanza italiana appartiene a una sotto popolazione molto diversa per età, selezionata a causa dei motivi della migrazione. Generalmente si parla di effetto dello "straniero sano" in quanto essendo gli stranieri in media più giovani, perché in età lavorativa, ed in buona salute, sembrano necessitare in maniera minore di prestazioni sanitarie. Dato che il rapporto causale tra fragilità e cittadinanza, a seguito del processo di autoselezione, non è chiaro, si preferisce escludere tale variabile a priori dal modello.

Le covariate relative ai ricoveri d'urgenza nel 2013 sono state escluse a priori perchè qualitativamente troppo simili alla variabile risposta. Si è preferito non considerare la variabile relativa ai ricoveri con diagnosi di malattia mentale, in quanto sovrapponibile alla variabile relativa agli accessi ai servizi della Psichiatria Territoriale (l'83% delle persone con diagnosi di malattia mentale ha avuto accesso ai servizi della Psichiatria Territoriale).

Vi sono inoltre altre variabili utilizzate per il modello MoSaiCo da mantenere nell'analisi per l'Ulss 15:

- Tre o più ricoveri in presenza di diverse diagnosi principali in due anni di osservazione (2012, 2013), per abbreviare "3 + Diagnosi";
- Indice di Charlson maggiore di uno almeno in occasione di un ricovero avvenuto nel 2013 (una breve descrizione è riportata nel prossimo paragrafo);
- Poliprescrizione: almeno quattro farmaci prescritti negli ultimi tre mesi dell'anno di osservazione;

Le variabili legate ai servizi sociali non sono disponibili per l'Ulss 15, per sostituirle in modo almeno approssimativo viene inserita una variabile basata sull'indice di deprivazione (Caranci *et al.*, 2010) che indica il livello di deprivazione nel 2001 (anno per cui i dati sono ad oggi disponibili) delle sezioni di censimento in cui ogni individuo risulta residente. L'intento è quello di utilizzare questa variabile come proxy di ridotta capacità economica. La variabile corrispondente a quest'indice è dicotomica e vale 1 se l'individuo appartiene al 20% della popolazione più deprivata (corrispondente al valore 5, "molto deprivato") e 0 altrimenti.

Charlson Index

Il *Charlson Comorbidity Index* è un indice che serve a misurare le comorbidità ed è molto correlato con la mortalità. Per la costruzione di questo indice si considerano le diagnosi assegnate nei ricoveri, dalla seconda alla sesta, e si pesano secondo un criterio legato alla mortalità di tale patologia. In particolare vengono assegnati i seguenti pesi:

- 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.

Questi pesi, nel caso un paziente manifesti più di una di queste comorbidità, devono essere sommati in modo da ottenere un indice unico. Sono state elaborate diverse versioni di Charlson index, nel modello MoSaiCo e in quello proposto in questo capitolo si fa riferimento a quella proposta da Deyo (Deyo *et al.*, 1992), che è quella più comunemente utilizzata. Nel modello MoSaiCo viene utilizzato come variabile dicotomica che vale 1 se il valore dell'indice è ≥ 1 e 0 altrimenti.

3.4.2 Le classi di età

Si pone a questo punto il problema dell'arbitrarietà delle classi di età: anche in letteratura, non c'è accordo su quante o quali classi di età siano più opportune per studiare il fenomeno della fragilità. Per ovviare a questo problema si è optato per un'analisi di segmentazione. La variabile risposta (Y) utilizzata è dicotomica, pari ad uno se l'individuo è stato ricoverato d'urgenza o è morto nell'anno successivo a quello di rilevazione delle variabili, uguale a zero altrimenti; come esplicativa è stata utilizzata soltanto l'età per mettere in evidenza le classi che raggruppano meglio gli utenti dell'Ulss. Per far sì che gli individui per cui $Y=1$ (casi) e quelli per cui $Y=0$ (controlli) siano equi-rappresentati nel campione, i controlli sono stati campionati casualmente utilizzando un criterio probabilistico. Inoltre, per non rischiare che i risultati ottenuti tramite l'analisi di segmentazione dipendano dal campione casuale ottenuto e per non perdere informazione scartando un numero così elevato di record relativo agli individui per cui vale $Y=0$, l'analisi è stata ripetuta ventidue volte (pari al risultato arrotondato per eccesso del rapporto tra il numero di controlli e di casi). I controlli sono stati assegnati (con probabilità uniforme da 1 a 22) ad uno dei ventidue campioni che, uniti al gruppo di casi, sono stati utilizzati per l'analisi. I primi tre nodi di ognuna di queste segmentazioni sono molto simili, per ognuno di questi è stata calcolata la media nelle 22 repliche e la varianza dei valori raccolti (tabella 3.2).

	Primo nodo	Secondo nodo
Media	68.5	22.6
Varianza	2.4	3.8

Tabella 3.2: Analisi di segmentazione delle classi di età, media e varianza dei nodi ottenuti

In base a queste analisi sono state costruite tre classi di età, aggiustando i valori limite in modo che siano più *user friendly*:

- 0-20, Giovani (età=1)
- 21-69, Adulti (età=2)
- 70 e più, Anziani (età=3).

3.5 Analisi per classi di età

Nell'ipotesi che i fattori di rischio per la fragilità si differenzino a seconda della classe di età a cui l'individuo appartiene, si effettuano qui di seguito le analisi di regressione logistica per l'Ulss 15 come nel modello MoSaiCo, ma apportando alcune modifiche. La variabile dipendente utilizzata è pari a 1 se l'individuo è morto o è stato ricoverato nel 2014, 0 altrimenti; le covariate inserite nel modello sono quelle riorganizzate e calcolate con riferimento al 2013; i campioni su cui viene effettuata l'analisi sono composti da casi ($Y=1$) e controlli ($Y=0$) in proporzioni uguali.

Si eseguono quindi tre analisi di regressione logistica su campioni di individui composti come descritto in tabella 3.3

Campione	Numerosità del campione	Numero di casi
Giovani	1534	767
Adulti	11154	5577
Anziani	9572	4786

Tabella 3.3: Composizione e numerosità dei campioni per classi di età

In tabella 3.4 sono riportati gli odds ratio relativi ai coefficienti stimati dalle regressioni logistiche per classi di età separate. Si è utilizzato un criterio di selezione stepwise con p-value di entrata pari a 0.15 e p-value di uscita pari a 0.10.

Analisi della classe di età tra 0 e 20 anni, giovani

Nella regressione logistica effettuata per questa classe di età, quella che in assoluto ha una percentuale di fragili inferiore (solo 1.41%), vengono inserite soltanto quattro covariate, più l'intercetta. Seguendo l'ordine di scelta, viene inserita la variabile relativa agli accessi al pronto soccorso, quella relativa ai

ricoveri senza traumi in prima diagnosi, le poliprescrizioni e infine quella relativa ai servizi della psichiatria territoriale. In base agli odds ratio riportati, chi, tra i giovani, ha effettuato almeno un accesso al pronto soccorso nel 2013, ha quasi il doppio della probabilità, in media, di incorrere in uno degli outcome relativi alla fragilità (morte o ricovero d'urgenza) nel 2014. Coloro invece che sono stati ricoverati almeno una volta nel 2013 hanno quasi due volte e mezza la probabilità di morire o essere ricoverati d'urgenza nel 2014, rispetto agli altri. Le variabili poliprescrizione e psichiatria territoriale hanno effetti molto più marcati. Tuttavia, l'area sotto la curva ROC è pari a 0.592, il che indica che il modello è scarsamente accurato nell'individuare i soggetti che nell'anno successivo incorreranno nelle prognosi negative legate alla fragilità. Tale valore potrebbe risultare comunque accettabile se si trattasse di un modello che distingue molto bene i soggetti per cui $Y=1$, a scapito di un ampio numero di falsi positivi, ma non è questo il caso infatti la sensibilità è particolarmente bassa, pari a 34.4%.

Analisi della classe di età tra 21 e 69 anni, adulti

Il numero di adulti fragili (che rappresentano il 3.41% del totale) è superiore rispetto a quello dei giovani ed è quindi più facile rilevare delle variabili significative, l'analisi logistica con criterio di selezione stepwise svolta in questa classe di età, infatti include nel modello undici covariate. La variabile relativa ai ricoveri (non con traumi in prima diagnosi) è la prima ad essere inserita nel modello e indica che coloro che sono stati ricoverati nel 2013 hanno l'80% di rischio in più di incorrere nelle prognosi negative legate alla fragilità, rispetto agli altri adulti. La seconda variabile che viene scelta per entrare nell'insieme delle covariate è il sesso, secondo cui essere maschio invece che femmina risulta un fattore protettivo: i maschi hanno circa metà del rischio di morire o essere ricoverati nell'anno successivo rispetto alle donne. Nessuno degli intervalli di confidenza degli odds ratio calcolati in questo modello contiene il valore 1, questo significa che le variabili hanno tutte una direzione definita che, fatta eccezione per il sesso, fa aumentare il rischio di morte o ricovero d'urgenza nell'anno successivo per coloro che sono stati esposti a questi fattori di rischio.

I parametri inseriti nella parte inferiore della tabella come diagnostiche del modello fanno riferimento al livello di probabilità p . Nello specifico sensibilità, 75%, e specificità, 45.9%, di questo modello fanno riferimento al livello 0.48 di probabilità stimata; generalmente si preferisce un livello che garantisca un buon equilibrio tra queste due, tuttavia anche cercando un altro punto di equilibrio, a costo di penalizzare leggermente la sensibilità a beneficio della specificità, in

corrispondenza del livello di probabilità 0.5, la situazione di sensibilità e specificità appare invertita (sensibilità pari a 35.3%, specificità pari a 83,6%). Quando si verificano situazioni di questo tipo, generalmente, una variabile causa questa repentina inversione: nel caso in esame si tratta della variabile sesso. Stimare un modello senza tenere conto di questa variabile significa ottenere un modello molto meno accurato secondo tutti i parametri diagnostici, infatti il sesso è una tra le variabili più influenti per predire la variabile risposta, dato che viene incluso nel modello per secondo. Si preferisce quindi penalizzare la specificità per ottenere una sensibilità migliore: infatti, avendo lo scopo di identificare gli individui fragili, è molto meno grave avere falsi positivi (e quindi identificare come fragili soggetti che in realtà non sono morti nè hanno subito ricoveri d'urgenza nell'anno successivo) rispetto ad avere falsi negativi.

Analisi della classe di età con più di 70 anni, anziani

Il fenomeno della fragilità è certamente più comune tra gli anziani (in cui i casi rappresentano il 15.26% del totale), per questo motivo il modello ottenuto in questa classe di età si presenta migliore rispetto ai due precedentemente descritti, tuttavia, in base all'area sotto la curva ROC, questo modello risulta moderatamente accurato.

La variabile più importante tra gli anziani per spiegare la morte o il ricovero d'urgenza nel 2014 è l'invalidità, il cui odds ratio mostra che il rischio di incorrere negli esiti negativi legati alla fragilità è circa il doppio per i soggetti invalidi rispetto ai non invalidi. Tra le dieci variabili selezionate in questa classe di età per modellare la variabile dipendente, "3 o più diverse prime diagnosi nei ricoveri degli ultimi due anni" risulta non significativa al 5% e l'intervallo di confidenza ad essa riferito contiene il valore 1; questo vuol dire che la variabile in questione potrebbe avere un effetto anche nullo sul rischio di incorrere negli esiti negativi legati alla fragilità.

La maggior parte delle variabili contenute in questo modello, comunque, nove su dieci, sono presenti anche nella regressione relativa agli adulti, è quindi possibile affermare che non ci sia una differenza abissale tra le classi d'età tale da giustificare tre modelli separati. Infatti le variabili utilizzate nella regressione della classe d'età più giovane sono presenti in quella effettuata per i soggetti adulti. Per evitare di dover stimare senza ragione tre diversi modelli, sembra quindi opportuno analizzare tutte le classi di età attraverso un'analisi unica, ma con la consapevolezza che di fatto non ci sono differenze importanti tra i fattori di rischio che agiscono sulle varie classi d'età.

Variabili	Giovani		Adulti		Anziani	
	Odds Ratio	IC 95%	Odds Ratio	IC 95%	Odds Ratio	IC 95%
Assistenza Domiciliare Integrata			2,631***	1,449-4,775	2,084***	1,792-2,424
Cardiopatia Coronarica			1,326**		1,326**	1,032-1,705
Tumore Maligno			2,200***	1,441-3,359	1,464***	1,172-1,830
Charlson Index			2,760***	1,744-4,367	1,557***	1,318-1,838
Scompenso Cardiaco Congestizio			8,766***	1,163-66,066	1,852***	1,451-2,363
3 + Diagnosi			2,003***	1,230-3,263	1,230*	0,972-1,556
Poliprescrizione	4,69***	1,588-13,849	1,576***	1,375-1,805	1,474***	1,348-1,613
Invaldit�			1,899***	1,550-2,327	2,072***	1,873-2,292
Ricovero	2,387***	1,472-3,872	1,826***	1,583-2,105		
Pronto Soccorso	1,948***	1,502-2,527	1,555***	1,394-1,733	1,684***	1,528-1,857
Psichiatria Territoriale	6,702*	0,823-54,543	2,137***	1,669-2,735		
Sesso (maschio vs femmina)			0,534***	0,493-0,578	1,203***	1,101-1,314
Livello di probabilit�	0,5		0,48		0,44	
Concordi	58,7		60,6		65,7	
Sensibilit�	34,4		75,2		68,3	
Specificit�	82,9		45,9		63,1	
Falsi Positivi	33,2		41,8		35,1	
Falsi Negativi	44,2		35		33,5	
area ROC	0,592		0,657		0,709	
R ²	0,0527		0,0928		0,132	
Pseudo R ²	0,0703		0,1237		0,176	

Tabella 3.4: Analisi di regressione logistica per classi di et  sulla variabile dipendente relativa alle prognosi negative legate alla fragilit . (Legenda per i livelli di significativit : ***=0.01, **=0.05, *=0.1)

3.6 Analisi di regressione logistica per prevedere le prognosi negative della fragilità

Date le ridotte differenze tra i fattori di rischio nelle tre regressioni per classi di età appena descritte, si preferisce stimare un modello unico, con l'aggiunta di una variabile che indica a quale delle tre classi di età appartiene ogni soggetto.

Si costruisce quindi un campione casuale di 22260 individui, i 11130 casi e 11130 controlli scelti in modo casuale. Utilizzando un campione così costruito, si stima un modello logistico con criterio di selezione stepwise con variabile risposta Y pari a 1 se l'individuo è morto o ha subito un ricovero d'emergenza nel 2014 e 0 altrimenti. I p-value critici utilizzati sono 0.15 per l'entrata e 0.05 per l'uscita.

Si ottiene quindi un modello come in tabella 3.5 in cui sono presenti 13 covariate. La variabile in assoluto più utile a prevedere gli outcome legati alla fragilità è l'età riportata in tabella con gli odds ratio relativi al confronto tra giovani e adulti e tra anziani e adulti. È interessante notare che gli anziani hanno quasi tre volte il rischio che hanno gli adulti di incorrere negli outcome relativi alla fragilità, mentre i giovani hanno circa la metà del rischio rispetto agli adulti. I soggetti che usufruiscono di assistenza domiciliare invece hanno il doppio di rischio in più di subire le prognosi negative legate alla fragilità rispetto a coloro che non ne usufruiscono. Gli unici fattori di rischio che sembrano avere, sulla base del coefficiente, un ruolo "protettivo" sono l'essere giovani rispetto all'essere adulti ed essere maschi invece che femmine (circa il 30% di rischio in meno). In questo modello risulta significativo, anche se solo al livello 0.05, l'indice di deprivazione; risulta quindi che chi risiede in un settore di censimento con indice di deprivazione appartenente al quintile più deprivato ha un rischio superiore al resto della popolazione, pari al 10% in più.

Si può notare, osservando la parte inferiore della tabella 3.5, che il modello stimato ha delle caratteristiche diagnostiche soddisfacenti e nettamente migliori rispetto a quelli precedentemente descritti. Infatti, l'area sotto la curva ROC, figura 3.4, è pari a 0.764. Questo modello classifica correttamente il 67.7% dei soggetti che hanno avuto a che fare con le prognosi negative legate alla fragilità nel 2014 (Sensibilità) e classifica correttamente il 70.7% dei controlli (Specificità).

Rispetto alla misura dell'area sotto la curva ROC del modello MoSaiCo calcolato dal Servizio Valutazione Strategica dell'Ulss di Ravenna (0.774), questo modello presenta un valore leggermente peggiore, con una differenza pari ad un centesimo. In compenso si tratta di un modello molto più parsimonioso infatti vengono utilizzate meno della metà delle variabili impiegate nel modello MoSaiCo

Variabile	Odds Ratio	IC 95%	Ingresso
Adulto (rispetto a Giovane)	0.458***	0.419-0.502	1
Anziano (rispetto a Giovane)	2.764***	2.534-3.014	1
Assistenza Domiciliare Integrata	1,990***	1,589-2,491	8
Tumore Maligno	1,648***	1,280-2,121	12
Charlson Index	1,714***	1,370-2,143	7
Scompenso Cardiaco Congestizio	1,614**	1,092-2,388	11
3 + Diagnosi	1,573***	1,191-2,077	10
Poliprescrizione	1,657***	1,514-1,815	4
Invalidità	1,849***	1,644-2,081	3
Ricovero	1,708***	1,535-1,900	2
Pronto Soccorso	1,503***	1,390-1,625	6
Psichiatria Territoriale	1,862***	1,532-2,262	9
Indice di deprivazione	1,087**	1,009-1,172	13
Sesso (maschio vs femmina)	0,673***	0,634-0,715	5
<hr/>			
Livello di probabilità	0,42		
Corretti	62,9		
Sensibilità	67,7		
Specificità	70,7		
Falsi Positivi	30,2		
Falsi Negativi	31,4		
area ROC	0,764		
R_2	0.2104		
Pseudo R_2	0.2806		

Tabella 3.5: Regressione logistica a classi d'età unite, con campione equilibrato rispetto alla variabile risposta relativa alle prognosi negative legate alla fragilità. (Legenda per i livelli di significatività: ***=0.01, **=0.05, *=0.1)

(solo 13 rispetto alle 31 inserite tramite selezione stepwise nel modello MoSaiCo). Inoltre, un pregio non trascurabile è che le variabili riorganizzate per stimare il modello relativo all'Ulss 15 sono più facili da reperire avendo un unico riferimento temporale e non dovendo suddividere gli accessi ai servizi sanitari in base alle cinque suddivisioni precedentemente descritte. Discorso a parte andrebbe fatto per l'indice di deprivazione per cui è necessario georeferenziare l'Anagrafe Sanitaria e che probabilmente implica un impegno maggiore rispetto al record linkage dei dati sociali effettuato per l'Ulss di Ravenna. Tuttavia, stimando nuovamente il modello anche senza l'indice di deprivazione, la misura dell'area sotto la curva ROC rimane praticamente invariata (0.763).

Concludendo, pur perdendo una quantità minima in termini di area sotto la curva ROC, il modello stimato in questo capitolo appare molto più semplice, parsi-

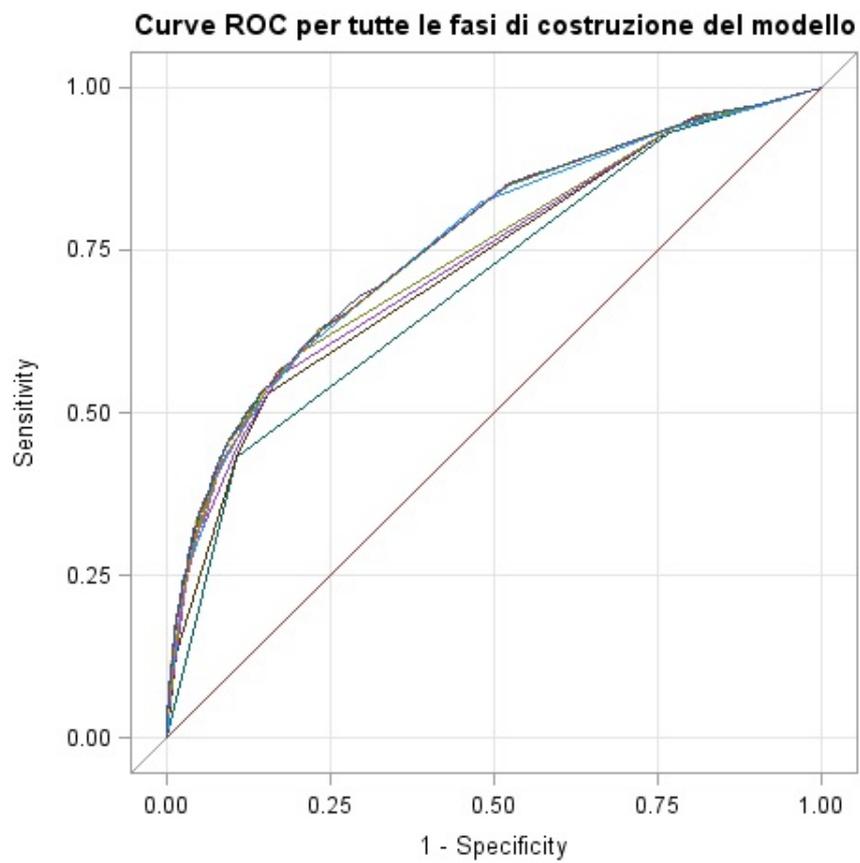


Figura 3.4: Curve ROC relative alla stima del modello in tabella 3.5 tramite selezione stepwise.

monioso e facilmente replicabile rispetto a quello stimato dal Servizio Valutazione Strategica dell'Ulss di Ravenna.

3.7 Un modello loglineare per identificare le interazioni tra le variabili

Dato che in un contesto come quello in esame è comune che vi siano interazioni tra le variabili, in questo paragrafo l'obiettivo è quello di cercare le possibili interazioni significative tra queste e inserirle nel modello logistico per cercare di migliorarne la capacità predittiva. Le variabili inserite nel modello precedentemente descritto sono 13, non è quindi pensabile inserire tutte le interazioni nel modello; per identificare le interazioni presenti tra le variabili esplicative con la variabile risposta (morte o ricovero d'urgenza nell'anno successivo a quello di riferimento), si stima un modello log-lineare. Infatti identificando gli effetti tripli significativi che coinvolgono anche la variabile risposta si ottengono le interazioni tra le coppie di variabili che hanno un effetto sulla variabile dipendente. Si è stimato il modello inserendo, inizialmente, tutte le interazioni doppie e triple; successivamente sono eliminate quelle non significative. Sono state considerate le interazioni triple significative contenenti anche la variabile dipendente: l'interazione fra le due variabili esplicative è stata inserita nel modello logistico.

In tabella 1 dell'appendice sono riportati gli effetti inseriti nel modello con i rispettivi livelli di significatività, mentre nella seconda tabella sono riportate le stime dei parametri legati alle interazioni che coinvolgono la variabile risposta. Le interazioni triple che coinvolgono la variabile risposta che si sono rivelate essere significative sono quindi:

- Invalidità e Psichiatria
- Tumore Maligno e Sesso
- Poliprescrizione e Sesso
- Tumore Maligno ed Età
- Sesso ed Età

Le interazioni significative sono state utilizzate per la stima di un nuovo modello simile a quello in tabella 3.5, ma con l'aggiunta delle interazioni sopra elencate. È stato utilizzato il criterio di selezione stepwise per scremare il paniere di variabili da utilizzare, fino ad ottenere il modello in tabella 3.6.

Come si può notare, sono entrate nel modello solo tre delle 5 interazioni significative. Il modello ottenuto introducendo le interazioni non presenta tuttavia caratteristiche significativamente migliori rispetto al modello precedente senza

interazioni. Dato che l'introduzione degli effetti combinati delle variabili comporta difficoltà maggiori nell'interpretazione del modello e nel calcolo degli odds ratio e minore parsimonia, si decide di mantenere il modello in tabella 3.5 per le analisi successive.

Variabili	DF	Stima	Standard Error	Chi-Square	Pr > ChiQuadr
Intercetta	1	-0,3886	0,0318	1.495.892	<,0001
Adulto (rispetto a Giovane)	1	-10,705	0,0464	5.321.633	<,0001
Anziano (rispetto a Giovane)	1	0,9491	0,043	4.879.834	<,0001
Assistenza Domiciliare Integrata	1	0,7561	0,1154	429.360	<,0001
Tumore Maligno	1	0,0967	0,1797	0,2895	0,5905
Charlson Index	1	0,5828	0,1191	239.444	<,0001
Broncopneumopatia cronica ostruttiva	1	-0,5227	0,2221	55.394	0,0186
Scompenso Cardiaco Congestizio	1	0,5039	0,2008	62.979	0,0121
3 + Diagnosi	1	0,4401	0,1427	95.181	0,002
Poliprescrizione	1	0,2083	0,0623	111.648	0,0008
Invaldit�	1	0,653	0,0605	1.163.189	<,0001
Ricovero	1	0,4938	0,0547	814.190	<,0001
Pronto Soccorso	1	0,4277	0,0401	1.136.146	<,0001
Psichiatria Territoriale	1	0,6455	0,1001	415.769	<,0001
Indice di deprivazione	1	0,0812	0,0383	44.935	0,034
Sesso (maschio vs femmina)	1	-0,2883	0,0431	447.763	<,0001
Tumore Maligno*Sesso	1	0,6876	0,2431	79.966	0,0047
Poliprescrizione*Sesso	1	0,6363	0,0899	501.443	<,0001
Sesso*Et� (Giovane rispetto ad Adulto)	1	0,426	0,0642	439.686	<,0001
Sesso*Et� (Anziano rispetto ad Adulto)	1	0,0128	0,0062	0,0429	0,836
Livello di probabilit�					
Corretti	0,44				
Sensibilit�	69,4				
Specificit�	66,1				
Falsi Positivi	72,7				
Falsi Negativi	29,2				
area ROC	31,8				
R ₂	0,767				
Pseudo R ₂	0,2173				
	0,2897				

Tabella 3.6: Regressione logistica a classi d'et  unite e interazioni, con campione equilibrato rispetto alla variabile risposta relativa alle prognosi negative legate alla fragilit .

3.8 Stratificazione in base al rischio di incorrere nelle prognosi negative della fragilità

Dato l'obbiettivo di questo capitolo, ovvero stratificare gli individui rispetto al rischio stimato di incorrere in un outcome negativo legato alla fragilità, si procede alle stime dei valori predetti per ogni individuo utilizzando il modello senza interazioni in tabella 3.5. Si tratta di valori tra 0 e 1 che dipendono dalle caratteristiche del singolo individuo in base ai coefficienti del modello stimato. Suddividendo i valori in decimi si ottiene una distribuzione come quella in tabella 3.7. Nella prima colonna vi sono i livelli di rischio tali che 9 è alto rischio e 1 corrisponde al rischio basale, nella seconda, invece, sono riportate le numerosità in termini assoluti per ogni classe di rischio. Si può notare che con il diminuire del rischio, i soggetti che popolano le classi aumentano, ad eccezione per le tre con rischio più basso per cui la percentuale di soggetti è minore che nelle classi centrali. Nell'ultima colonna a destra è rappresentata la percentuale di individui con $y=1$ presenti in ogni classe, si può notare che nelle classi con rischio più alto, la percentuale di casi cresce notevolmente, infatti il coefficiente di correlazione di Pearson tra livello di rischio e percentuale di casi è significativo e pari a 0.86 (con p-value pari a 0.003).

Rischio	Soggetti	Percentuale	Percentuale cumulata	Percentuale di casi
9	4282	1,72	100,00	39.58
8	6524	2,61	98,28	20.75
7	9743	3,90	95,67	12.62
6	14382	5,76	91,77	8.52
5	22699	9,09	86,01	5.48
4	76351	30,58	76,92	3.52
3	65983	26,43	46,33	1.65
2	27453	11,00	19,90	1.21
1	22229	8,90	8,90	1.25

Tabella 3.7: Stratificazione in base al rischio di incorrere nelle prognosi negative della fragilità, su 9 livelli di rischio (dove 9 è il più alto e 1 è il più basso)

É possibile inoltre suddividere gli utenti dell'Ulss 15 in base ai valori predetti utilizzando una scala simile a quella del Combined Model (Health Dialog, 2006). Tuttavia è necessaria una leggera modifica alle classi con alto rischio poichè sono state predisposte utilizzando i valori predetti provenienti da un modello stimato su un campione sbilanciato, infatti tutti i valori sono concentrati su livelli di probabilità alti molto alti.

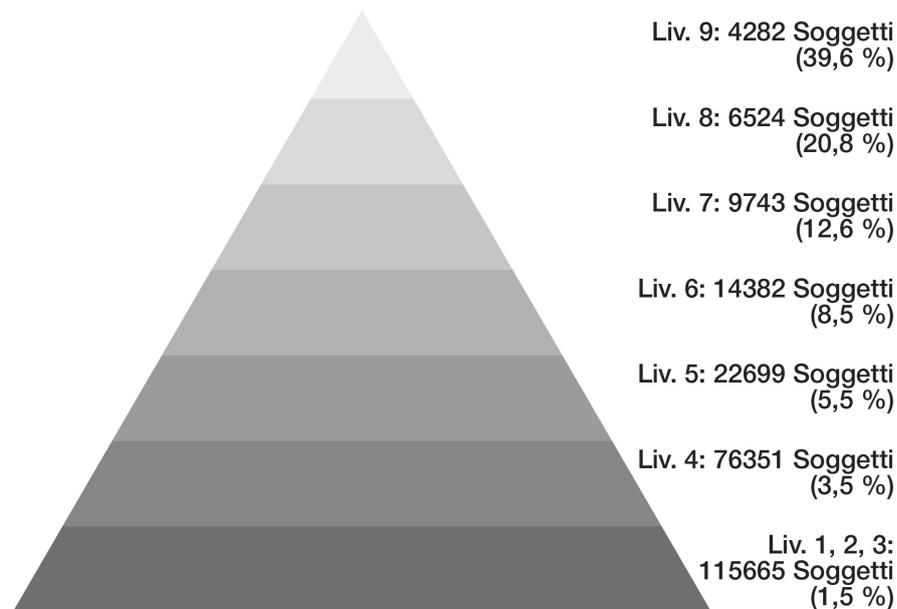


Figura 3.5: Stratificazione in base al rischio di incorrere nelle prognosi negative della fragilità, su 9 livelli di rischio.

- Rischio molto alto, con valori predetti ≥ 0.99 ; nel Combined Model il valore soglia è 0.995, ma in base alla distribuzione dei valori predetti del modello stimato su un campione bilanciato per l'Ulss 15, cadrebbero in questa classe meno di 10 individui.
- Alto rischio, con valori predetti tra 0.95 compreso e 0.99.
- Rischio moderato, con valori predetti tra 0.80 compreso e 0.95.
- Basso rischio, con valori predetti < 0.80 .

Una possibile critica ai modelli fin qui stimati è il fatto che questi, più che essere legati alla fragilità, fanno riferimento ai fattori di rischio relativi ad una variabile dipendente che non è detto sia un buon criterio per riconoscere i fragili. Modelli di questo tipo vengono utilizzati spesso per validare altri metodi di misura della fragilità.

Sarebbe necessario, per avere un metodo migliore per riconoscere i soggetti fragili, rivedere la variabile dipendente, magari aggiungendo altri eventi che possano rappresentare la fragilità o il peggioramento delle condizioni di salute di un individuo (come ad esempio diventare invalidi nell'anno successivo a quello di riferimento, o iniziare a beneficiare di assistenza domiciliare nell'anno successivi); oppure utilizzare metodi che non necessitino di una variabile dipendente.

Capitolo 4

Un nuovo approccio per la misura della fragilità mediante gli ordinamenti parziali

4.1 Introduzione

Dovendo analizzare un fenomeno complesso e multidimensionale, generalmente si cerca di sintetizzare le informazioni a disposizione formando un indicatore composito utilizzando uno o più sistemi di pesi per aggregare le informazioni appartenenti a diverse dimensioni. In questo modo risulta quindi possibile quantificare il fenomeno e ordinare le unità statistiche analizzate secondo tale misura. Tuttavia, l'ordinamento ottenuto in questo caso dipende molto dal sistema di ponderazione utilizzato e questo può risultare soggettivo. È una scelta metodologica soggettiva anche assegnare pesi uguali agli indici che servono a creare l'indicatore composito, come accade per il *Frailty Index* elaborato da Rockwood e per l'indicatore dell'azienda Sanitaria Locale di Milano che rileva i soggetti fragili alle ondate di calore.

Per non dover aggregare indici appartenenti a dimensioni molto diverse tra loro introducendo criteri di ponderazione soggettivi, e tenere maggiormente conto della struttura ordinale dei dati si è pensato di introdurre una metodologia di analisi inedita in ambito epidemiologico: gli ordinamenti parziali.

Gli ordinamenti parziali (poset) sono, in matematica, una generalizzazione del concetto di sequenza o disposizione di oggetti. Si tratta di un campo teorico tra la matematica discreta e la teoria delle relazioni che ha subito un rapido sviluppo dagli anni '70 in poi grazie a nuovi strumenti informatici che hanno consentito l'utilizzo di strumenti altrimenti computazionalmente troppo impegnativi. Un

importante contributo teorico e applicativo a questo tipo di ordinamenti è dato dal lavoro di Rainer Brüggemann che li ha utilizzati soprattutto per analisi in ambito chimico (Brüggemann & Patil, 2010).

L'utilizzo di questo metodo per ordinare gli individui dell'Ulss 15 in base alla fragilità permette di ovviare ad alcuni problemi che presentano altri metodi parametrici già utilizzati in quest'ambito, ordinando soggetti con diverse caratteristiche senza introdurre soggettività, ma basandosi soltanto sulla matrice di dati a disposizione e sui confronti attuabili tra le unità presenti.

In questo capitolo, dopo una prima parte teorica necessaria a chiarire le basi degli ordinamenti parziali, vengono illustrati e descritti i metodi di stima introdotti da Brüggemann per l'ordinamento di alcuni composti chimici in base all'impatto che questi hanno sulla salute umana. Successivamente, per adattare tali metodi ad un contesto caratterizzato da numerosità molto più alte di unità statistiche, come nel caso di studio in questione, viene utilizzata una modifica a questi metodi (Caperna *et al.*, 2015). Nella parte finale di questo capitolo vengono ordinati gli individui dell'Ulss 15 in base alla loro fragilità e vengono evidenziati e descritti i soggetti più fragili.

4.2 Ordinamenti parziali e diagrammi di Hasse

Per intuire come questi metodi possano bene adattarsi ad individuare gli utenti fragili secondo le loro caratteristiche cliniche, senza utilizzare alcun criterio soggettivo, si può utilizzare un esempio semplificato con tre variabili e sei individui che presentano le caratteristiche specificate in tabella 4.1. Si tratta di tre variabili

Soggetto	Variabile1: Poliprescrizione	Variabile 2: Assistenza domiciliare	Variabile 3: Pronto soccorso
A	0	0	0
B	1	0	0
C	0	1	0
E	1	1	0
G	0	1	1
H	1	1	1

Tabella 4.1: Una matrice di dati di esempio per l'applicazione degli ordinamenti parziali all'identificazione degli individui fragili

utilizzate per le analisi anche nel capitolo precedente: la poliprescrizione, cioè aver fatto uso di quattro o più farmaci appartenenti a quattro o più macrocategorie

diverse negli ultimi tre mesi del 2013, l'assistenza domiciliare, che consiste nell'aver beneficiato di questo servizio almeno una volta nel corso del 2013, e il pronto soccorso, ovvero aver fatto almeno un accesso al pronto soccorso nel corso del 2013. Sono variabili che rappresentano fattori di rischio per la fragilità e che aumentano le probabilità di incorrere in prognosi negative nel 2014, come visto nel capitolo precedente, quindi avere valore 1 (presenza della caratteristica a cui la variabile fa riferimento) rappresenta una situazione peggiore rispetto ad avere valore 0 (assenza della caratteristica a cui la variabile fa riferimento).

Come osservato nel capitolo 3, il contributo di queste variabili al rischio di incorrere negli esiti avversi relativi alla fragilità non è uguale e quindi effettuare una somma con un sistema di pesi uguali non sembra essere la soluzione migliore. Tuttavia anche utilizzare un sistema di pesi che faccia riferimento ai coefficienti calcolati nel precedente capitolo significa vincolare l'indicatore ad un criterio (rischio di decesso o di ricovero con urgenza nell'anno successivo) piuttosto arbitrario e discutibile.

Basando il confronto tra gli individui soltanto sulla matrice di dati, possiamo immaginare che C sia in condizioni peggiori rispetto ad A, infatti, a parità di poliprescrizioni e di accessi al pronto soccorso, C usufruisce dell'assistenza domiciliare. Attraverso un simile ragionamento è possibile anche affermare che E probabilmente avrà uno stato di salute più grave rispetto a C in quanto, a parità di servizio di assistenza domiciliare e di accessi al pronto soccorso, E fa un utilizzo di farmaci più esteso. Si può intuire che l'individuo H sia quello con le condizioni di salute peggiori rispetto agli altri rappresentati in tabella 4.1 dato che è l'unico che presenta tutti e tre i fattori di rischio. Non è possibile invece un confronto tra B e C che presentano caratteristiche diverse, non direttamente confrontabili in modo oggettivo; come decidere quale sia la variabile che indica una condizione di salute più grave tra il consumo di farmaci (poliprescrizione) e l'aver usufruito di assistenza domiciliare? Facendo ricorso ad un'aggregazione tramite la somma delle variabili, B e C sarebbero sullo stesso piano a meno che non vengano dati pesi diversi alle variabili che le caratterizzano e le distinguono. Come invece si vedrà in seguito, attraverso l'utilizzo dei metodi basati sugli ordinamenti parziali è possibile stabilire chi tra B e C sia più fragile.

4.2.1 Alcune basi teoriche per gli ordinamenti parziali

Per formalizzare, supponiamo che una matrice di dati X contenga, sulle righe, gli oggetti di interesse (che ad esempio possono essere gli individui precedentemente descritti). Si può quindi scrivere $X = \{A, B, C, E, G, H\}$. Per stabilire delle

relazioni d'ordine tra gli elementi di X , si utilizza il simbolo \leq che rappresenta una relazione binaria specificata dai seguenti assiomi:

Assioma 1: Riflessività: $x \in X : x \leq x$, un oggetto può essere confrontato con se stesso.

Assioma 2: Antisimmetria: $x \leq y, y \leq x$ implica $y = x$, Se entrambi i confronti sono validi, cioè se x è meglio di y e, allo stesso tempo, y è meglio di x , allora questo assioma implica che x sia identico a y .

Assioma 3: Transitività: $x \leq y$ e $y \leq z$ implica $x \leq z$, la transitività è presente se gli oggetti sono almeno su scala ordinale.

Nelle applicazioni pratiche, come in questo caso, non è utile trattare insiemi di oggetti con così pochi elementi, per questo risulta necessario considerare alcuni oggetti come equivalenti. Vengono considerati equivalenti gli oggetti che hanno righe identiche nella matrice dei dati. In questi casi si considera soltanto un oggetto della classe di equivalenza su cui vengono eseguite tutte le operazioni necessarie, tenendo presente che quest'oggetto rappresenta una classe di oggetti più numerosa.

Supponiamo che x e y siano due diversi oggetti dell'insieme X . Sia $q(x)$ la riga nella matrice di dati relativa a x e $q(y)$ quella per y . Allora si dice che $x \leq y$ se e solo se $q(x) \leq q(y)$, $q(x) \leq q(y)$ se e solo se $q_i(x) \leq q_i(y)$ per ogni i . Dato che le variabili che vengono considerate, nello spazio Q^m , hanno un orientamento definito, tali confronti sono legittimi.

Se x e y sono oggetti diversi 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. Per escludere l'equivalenza, si scrive $x < y$. Di conseguenza: $x < y$ se e solo se $q(x) \leq q(y)$, $q(x) \leq q(y)$ con almeno un q_{i^*} , per cui vale $q_{i^*}(x) < q_{i^*}(y)$. Gli ordinamenti descritti si dicono *product order* o *component-wise order*.

Nell'esempio in tabella 4.1 è impossibile confrontare B e C seguendo i criteri appena enunciati, si dice quindi che B e C sono incomparabili, e si scrive $B \parallel C$. Due oggetti x e y per cui risulta vero che $q(x) \leq q(y)$ oppure che $q(x) \geq q(y)$ si dicono comparabili (si scrive $x \perp y$), si vedrà in seguito quanto sia importante la distinzione tra oggetti comparabili e incomparabili per stabilire gli ordinamenti e calcolare il rango medio di ogni oggetto.

Quando l'insieme di oggetti X ha un ordine parziale, cioè gli oggetti in esso contenuti sono collegati tra loro da una relazione che rispetta gli assiomi sopra menzionati, si scrive (X, \leq) . Un insieme di oggetti che rispetta un ordine parziale

è detto anche poset (insieme parzialmente ordinato). Due oggetti risultano quindi comparabili, $x \perp y$, o incomparabili, $x \parallel y$, a seconda di quali variabili vengono utilizzate per il confronto, si parla allora di base informativa (*information base*), IB, con cui si intende l'insieme di variabili nella matrice di dati utilizzate nell'analisi. La base per l'analisi ordinale dei dati si scrive quindi (X, IB) , oppure $(X, \{q_1, q_2, \dots\})$, per cui deve valere $q_i(x) \leq q_i(y)$ per ogni i , dove $q_i \in IB$ e $x \in X$. Con $(X/\cong, IB)$ si intende un ordinamento in cui per almeno un q_{i^*} vale $q_{i^*}(x) < q_{i^*}(y)$, dove $q_i \in IB$ e $x \in X/\cong$.

4.2.2 Definizioni di base

1. Se nessuna delle coppie di oggetti presenti nella matrice dei dati risulta incomparabile, allora parliamo di un ordinamento completo, totale o lineare. In tal caso, gli oggetti $x \in X$ si possono disporre secondo un ordine noto: $x_1 < x_2 < \dots < x_n$.
2. Catena: Se è possibile individuare un sottoinsieme $X' \subset X$ tale che per ogni $(x, y) \in X' \times X'$ si riesce ad identificare un ordinamento completo, allora questo sottoinsieme è detto catena. In una catena tutti gli oggetti sono comparabili tra loro. Un esempio di catena è dato dagli oggetti A(000), B(100), E(110), H(111), in tabella 4.1: questi quattro oggetti sono confrontabili tra loro e vale $q_i(A) \leq q_i(B) \leq q_i(E) \leq q_i(H)$ per ogni i .
3. Ordine lineare: Tutti gli elementi di X formano una catena, l'insieme di oggetti rappresentati in tabella 4.1 non rappresenta un ordine lineare, infatti vi sono coppie di oggetti non confrontabili tra loro, ad esempio B(100) e C(010), E(110) e G(100), B(100) e G(011).
4. Ordine debole (*weak*): Un ordine si dice debole quando è costituito da una catena contenente più di un elemento appartenente alla stessa classe di equivalenza. Un esempio di ordine debole è l'insieme $Y = \{A(000), B(100), B'(100), E(110), H(111)\}$, infatti questi oggetti formano una catena (sono confrontabili e vale $q_i(A) \leq q_i(B) \leq q_i(B') \leq q_i(E) \leq q_i(H)$ per ogni i), ma B e B' presentano le stesse caratteristiche e, pur essendo due oggetti distinti, sono uguali e appartengono alla stessa classe di equivalenza (vale $B \cong B'$).
5. Anticatena: Se è possibile individuare un sottoinsieme $X' \subset X$ per cui non vale $x \perp y$ per nessuna coppia $(x, y) \in X' \times X'$, allora tale sottoinsieme è detto anticatena (per esempio gli elementi B e C in tabella 4.1, infatti questi oggetti non sono confrontabili).

6. In matrici di dati finite, catene e anticatene contengono un numero finito di oggetti. Pertanto, si può parlare di catene o di anticatene aventi una certa lunghezza, in base al numero di elementi che contengono.
7. Altezza: Il numero di elementi della catena più lunga è detto altezza del poset. L'altezza relativa ai dati in tabella 4.1 è 4, infatti le tre catene più lunghe che coinvolgono gli oggetti considerati sono (A,B,E,H), (A,C,E,H) e (A,C,G,H) e sono tutte costituite da un massimo di 4 elementi.
8. Larghezza: Il numero di elementi della massima anticatena è chiamato larghezza del poset (la larghezza relativa ai dati in tabella 4.1 è 2, infatti non esistono sottoinsiemi di oggetti incomparabili con più di due elementi).
9. Massimale: Un elemento $x \in X$ per cui non vale la relazione \leq con alcun elemento y , è detto elemento massimale.
10. Minimale: Un elemento $x \in X$ per cui non vale la relazione \geq con alcun elemento y , è detto elemento minimale.
11. Massimo/Minimo: si definisce un elemento massimo/minimo quando c'è solo un massimale/minimale. Nell'esempio in tabella 4.1 H(111) è il massimo, infatti è un massimale, non c'è nessun elemento y per cui vale la relazione $q_i(H) \leq q_i(y)$, ed è l'unico e A(000) è il minimo, infatti è un minimale, non c'è nessun elemento y per cui vale la relazione $q_i(A) \geq q_i(y)$, ed è l'unico.
12. Elemento Isolato: Un elemento $x \in X$, che sia allo stesso tempo un massimale e un minimale è detto elemento isolato. Dato un elemento isolato i , allora per ogni $x \in X$: $i \parallel x$.
13. Un elemento massimale/minimale è tale se non è isolato.
14. Relazione di copertura: Si dice che y copre x se non esiste un elemento $z \in X$ per cui $x < z$ e $z < y$. Si scrive $x \leq y$. Data la relazione di copertura, che è riflessiva e antisimmetrica, la relazione d'ordine parziale è ottenuta per transitività.

Diagramma di Hasse

Il diagramma di Hasse rappresenta gli oggetti di una matrice di dati in base alle relazioni di copertura che li legano. In figura 4.1 è rappresentato il diagramma relativo ai dati in tabella 4.1. I criteri che vengono utilizzati per la costruzione di questi diagrammi sono:

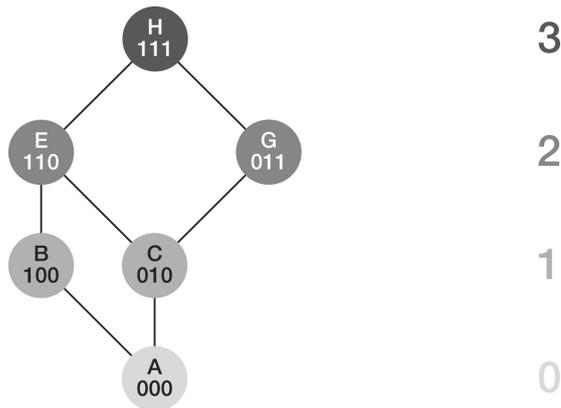


Figura 4.1: Rappresentazione con il Diagramma di Hasse degli oggetti visibili in tabella 4.1, a destra i livelli di una somma equiponderata.

- Se $x \leq y$ il nodo corrispondente all'oggetto x è collocato su un livello inferiore rispetto a quello dell'oggetto y , quindi più in basso.
- Se y copre x , $x \leq y$, viene inserita una linea che lega i nodi di x e di y , detta *edge*. Dato che il grafico è orientato dalla posizione degli oggetti, non è necessario che questi segmenti vengano direzionati.

La matrice ζ

Un altro modo per rappresentare dei dati tenendo conto del loro ordinamento parziale è la matrice ζ : righe e colonne corrispondono ad un oggetto. Se riga < colonna, nella cella corrispondente va inserito un 1, altrimenti vi si trova uno 0. In altre parole, siano $x, y \in X$, allora $\zeta(x, y) = 1 \Leftrightarrow x < y$.

Per esempio, l'ordine parziale dei dati in tabella 4.1 corrisponde alla seguente matrice ζ :

$$\zeta = \begin{matrix} & \begin{matrix} A & B & C & E & G & H \end{matrix} \\ \begin{matrix} A \\ B \\ C \\ E \\ G \\ H \end{matrix} & \begin{pmatrix} 0 & 1 & 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} \end{matrix}$$

Dato che il criterio per determinare ogni elemento della matrice è costituito da un minore stretto ($<$), gli elementi sulla diagonale sono tutti zeri. Inoltre, osservando la prima riga, è facile intuire che A è un minimale e minimo, infatti il primo elemento è uno 0 (perchè non ha senso scrivere $A < A$) e i rimanenti elementi sono pari a 1, dato che tutti gli elementi in tabella 4.1 sono $> A$, tranne A stesso e che A è confrontabile con tutti gli altri. Riflessioni opposte si possono fare per l'ultima riga in cui sono presenti solo elementi nulli: H è massimale,

pertanto per nessun elemento presente in tabella 4.1 si può scrivere la relazione $> H$.

La relazione illustrata attraverso la matrice zeta è asimmetrica, di conseguenza lo è anche la matrice: l'incrocio (E, B) ha valore nullo, infatti $E > B$; invece in corrispondenza dell'incrocio (B, E) si trova un 1, dato che $B < E$. Nel caso i due incroci coinvolgono oggetti incomparabili, il valore è nullo dato che non è possibile fare confronti di alcun tipo tra questi, per esempio (B, C) e (C, B) sono nulli dato che $B \parallel C$. Inoltre è bene specificare che la matrice ζ non è necessariamente triangolare superiore, ma lo è in questo esempio specifico perchè gli oggetti presentano un ordinamento verosimile in base alle caratteristiche della matrice di dati.

Up Sets, Down Sets ed elementi incomparabili

In una matrice di dati non tutte le righe, oggetti, sono confrontabili tra loro. Per ogni oggetto $x \in X$ si distinguono gli altri elementi della matrice in:

- Oggetti incomparabili, $U(x)$, tali che $U(x) := \{y \in X : y \parallel x\}$.
- Down Set $O(x) := \{y \in X : y \leq x\}$, con rispettivo Successor Set $S(x) := O(x) - \{x\}$.
- Up Set $F(x) := \{y \in X : x \leq y\}$, con rispettivo Predecessor Set $P(x) := F(x) - \{x\}$.
- L'insieme di tutti gli oggetti comparabili con x è $C(x) = F(x) \cup O(x)$.

Una caratteristica importante di un poset è il numero totale delle coppie di oggetti incomparabili, U . Si tratta di una misura utile a quantificare le relazioni, generalmente viene normalizzata per il numero di possibili coppie di oggetti contenuti in X ; cioè con $n * (n - 1) / 2$ dove n è $||X||$, ovvero il numero di oggetti contenuti in X . La misura normalizzata delle relazioni all'interno della matrice dei dati X è $P(IB)$, definita come

$$P(IB) = \frac{U}{\binom{n}{2}}$$

In base ai dati in tabella 4.1, le coppie di elementi incomparabili sono (B, C), (B, G) e (E, G), quindi $U = 3$; il numero di possibili coppie di oggetti contenuti nella tabella è 15, dato che $n = 6$; si ottiene quindi che la misura normalizzata delle relazioni all'interno della tabella 4.1 è pari a $P(IB) = 3/15 = 0.2$. Questo valore

può variare da un minimo di 0, situazione in cui tutti gli oggetti sono comparabili tra loro e non vi sono coppie incomparabili, ad un massimo di 1, in cui tutti gli oggetti sono elementi isolati e non risultano quindi comparabili.

Un altro parametro rilevante per misurare e confrontare gli ordinamenti parziali è la lunghezza, che è semplicemente il numero di oggetti nella catena più lunga, tolto uno. Dato che, come precedentemente calcolato, l'altezza relativa alla tabella 4.1 è pari a 4, la lunghezza è 3.

4.3 Estensioni lineari

Un'estensione lineare è un ordinamento lineare che rispecchia tutte le relazioni d'ordine del poset. In generale si possono ottenere diverse estensioni lineari da un poset, dato che la condizione di incomparabilità permette sia $x > y$ sia $x < y$ senza venire meno a nessuna delle imposizioni d'ordine. Il numero di tutte le estensioni lineari di un poset si identifica con LT , dove $LT \leq n!$, con n pari al numero di oggetti nel poset. Utilizzando i dati nella tabella 4.1, le estensioni lineari sono 5, come si può vedere in tabella 4.2.

H	H	H	H	H	H =	6
E	E	E	G	G	E =	4.6
G	G	B	E	E	G =	4.2
C	B	G	B	C	B =	2.8
B	C	C	C	B	C =	2.4
A	A	A	A	A	A =	1

Tabella 4.2: Esempio di estensioni lineari della matrice di dati in tabella 4.1 con rispettivi ranghi medi

Calcolando la media dei ranghi di ogni oggetto in tutte le possibili estensioni lineari si ottiene una misura esatta di *averaged rank*:

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

dove $H(x, le(k))$ è l'altezza dell'oggetto x nella k -esima estensione lineare, $le(k)$. In questo modo si ottengono i ranghi medi riportati in tabella 4.2; ordinando gli oggetti in base al rango medio si osserva quindi una classifica d'ordine degli oggetti rispetto alle caratteristiche di ognuno, senza utilizzare alcun criterio soggettivo. Tale classifica crea inoltre sei livelli diversi, uno per ogni oggetto, suggerendo quindi un ordinamento anche dove una somma equiponderata delle caratteristiche

non sarebbe stata in grado di introdurre alcuna discriminazione. Questa stima dei ranghi viene effettuata senza approssimazioni ed è quindi esatta; l'ordine che ne risulta è detto canonico. Generalmente, se il diagramma di Hasse ha una struttura simmetrica, l'ordine canonico è un ordinamento debole, in quanto a classi di equivalenza diverse, ma simmetriche, potrebbe essere associato lo stesso rango (in questo esempio il diagramma di Hasse è asimmetrico, quindi questo problema non si pone).

Per ottenere questo ordinamento, è necessario però che tutte le estensioni lineari di un poset siano note e che i ranghi di ogni oggetto della matrice siano calcolabili. Questo è possibile soltanto se le variabili e gli oggetti contenuti in una matrice di dati non sono molti e se le coppie di oggetti non comparabili hanno una numerosità contenuta. Con tre variabili dicotomiche le classi di equivalenza teoriche sono otto (2^3); con la sola aggiunta di una variabile dicotomica, le classi di equivalenza possibili sono sedici (2^4), come si può vedere dalla figura 4.2. Nel caso le variabili non fossero dicotomiche, la situazione sarebbe molto più complessa e computazionalmente impegnativa. Con un numero di variabili consistente il procedimento per ottenere i valori esatti dei ranghi medi è estremamente difficoltoso soprattutto da un punto di vista computazionale.

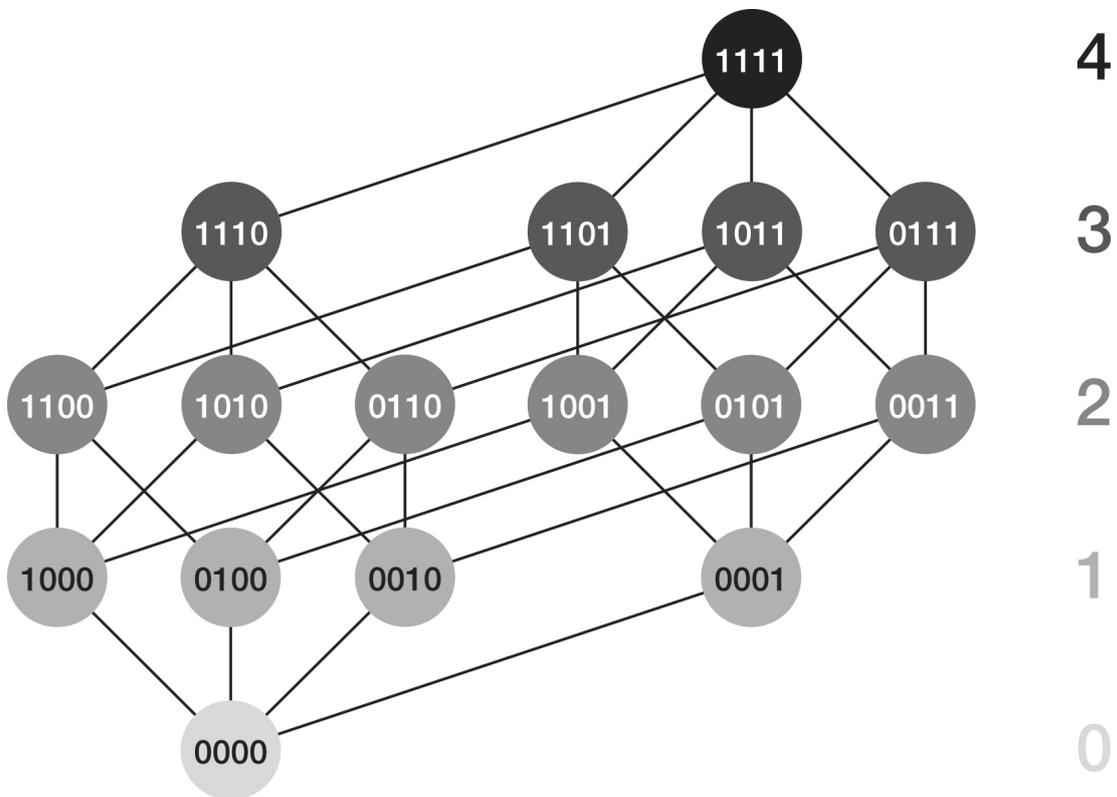


Figura 4.2: Diagramma di Hasse degli elementi rappresentabili con quattro variabili dicotomiche

4.4 Approssimazioni per il calcolo degli averaged rank

Dato che il calcolo esatto del rango medio degli oggetti di un poset nel caso in cui le coppie incomparabili siano molte è poco agevole soprattutto da un punto di vista computazionale, è necessario ricorrere a delle approssimazioni. In questa sezione si riportano due metodi di stima per il rango medio proposti da Rainer Brüggemann: un primo metodo, *Local Partial Order Model* (LPOM0), più semplice ed intuitivo, ma che in presenza di molte coppie incomparabili causa errori non trascurabili; e un secondo, *Local Partial Order Model Extended* (LPOMext), meno intuitivo e lineare anche da un punto di vista computazionale, ma che garantisce un'approssimazione migliore (Brüggemann & Carlsen, 2011).

4.4.1 Local Partial Order Model, LPOM0

Il primo semplice metodo è il Local Partial Order Model, LPOM0; di seguito una breve descrizione del procedimento.

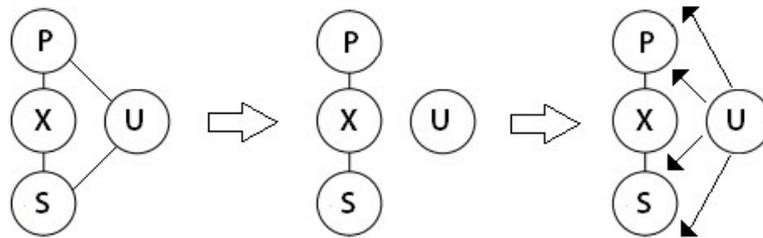


Figura 4.3: Passaggi per la stima dell'averaged rank di X, con il metodo LPOM0

Sia x , l'oggetto per cui è necessario stimare l'averaged rank. Allora:

1. Tutti gli oggetti incomparabili rispetto a x , $y \in X$ con $y \parallel x (= U(x))$, si considerano isolati. Si osservi la figura 4.3, l'oggetto U che pure forma una catena con S e P , nel calcolo dell'averaged rank per X viene considerato come isolato perchè è incomparabile rispetto a X .
2. L'ordine di tutti gli oggetti confrontabili con x viene esteso per ottenere un ordinamento lineare, chiamato catena S - x - P (Successor - oggetto di studio x - Predecessor). In una catena S - x - P gli elementi sopra e sotto a x devono formare delle catene di elementi comparabili con x , mentre non è necessario che lo siano tra loro. Osservando il secondo passaggio in figura 4.3, si vede che, "staccato" l'oggetto U , prende forma una catena formata dagli elementi

S, X e P dove S rappresenta tutti gli elementi di $O(X)$ per cui vale $S < X$ e P rappresenta tutti gli oggetti di $F(X)$ per cui vale $P > X$.

3. Gli oggetti di $U(x)$ potrebbero occupare diverse posizioni nella catena S-x-P; a questo punto si può calcolare l'altezza media tenendo conto dei successori, dei predecessori e di $\|U(x)\|$, secondo la formula:

$$Rkav(x) = (\|S(x)\| + 1) * (n + 1) / (n + 1 - \|U(x)\|)$$

Facendo riferimento alla figura 4.3, una volta isolato, l'oggetto U può ricoprire quattro posizioni come indicato dalle frecce: nelle due posizioni più alte si trova più in alto di x quindi non ne influenza il rango (X ha rango 2), nelle due posizioni inferiori invece finisce tra gli elementi inferiori ad X aumentando di uno il suo rango (X ha rango 3). Effettuando il calcolo in queste condizioni, il rango approssimato di X sarebbe 2.5.

Carenze di LPOM0

Osservando la figura 4.3 si possono cogliere alcuni punti deboli di questa approssimazione, infatti il fatto di trattare tutti gli elementi incomparabili come isolati comporta alcune distorsioni sistematiche:

1. Se gli oggetti incomparabili con l'elemento per cui si calcola l'averaged rank formano a loro volta una catena con alcuni dei predecessori e successori di x, allora gli oggetti incomparabili non avrebbero realmente accesso a tutte le posizioni che gli vengono attribuite nel passaggio tre. Risulta più chiaro osservando la figura 4.3: U, incomparabile con X, forma anche una catena con S e P per cui vale che $S < U$ e che $P > U$; considerare U come oggetto isolato comporta il fatto che U possa ricoprire posizioni anche superiori a P e inferiori a S che, secondo il diagramma originario non sarebbero state lecite.
2. Gli oggetti incomparabili sono diversi tra loro e non sempre hanno lo stesso insieme di posizioni possibili, soprattutto se sono legati da una relazione di copertura tra loro o con altri elementi del poset. Trattandoli tutti come isolati non vengono considerate le relazioni d'ordine non comprese nella catena S-x-P. Per esempio, se si fosse incluso anche un oggetto U_2 in figura 4.3, tale che $U > U_2$, la sua posizione, che dovrebbe dipendere da quella di U dato che vi è una relazione di copertura, sarebbe stata totalmente svincolata.

3. Si creano dei legami artificiali e non verosimili proprio a causa dell'isolamento degli oggetti incomparabili. Infatti i legami esistenti che non sono parte della catena S-x-P vengono totalmente ignorati e il contributo all'averaged rank di X di ogni elemento appartenente a $U(X)$ è basato su un criterio di equidistribuzione anche tra posizioni illecite secondo il diagramma iniziale.

Questo metodo funziona bene e fornisce un meccanismo semplice per capire come si ottiene il rango medio di un oggetto x se $P(IB)$ è piccolo, infatti la gran parte delle distorsioni è provocata da una gestione troppo approssimativa degli oggetti incomparabili. Brüggemann (Brüggemann & Carlsen, 2011) ha introdotto anche alcune modifiche e ha spiegato come migliorare la stima per correggere questi problemi. Il risultato di questi aggiustamenti è l'LPOMext, che però richiede un impegno computazionale maggiore.

4.4.2 Extended Local Partial Order Model, LPOMext

Questo metodo parte dal presupposto che, per ogni riga, x , della matrice di dati c'è un insieme di oggetti comparabili e uno di incomparabili. Ogni oggetto incomparabile, $y \in U(x)$, ha un range di posizioni da poter ricoprire a seconda dell'ordine parziale ereditato dagli oggetti comparabili. Più semplicemente, si tiene conto dei legami e delle relazioni di copertura che gli oggetti incomparabili formano con Up Set e Down Set dell'oggetto x per cui si sta stimando il rango medio, invece di considerarli tutti come se fossero isolati.

Si definiscono quindi due insiemi:

$$P_y^< := O(x) \cap U(y) \text{ e } P_y^> := F(x) \cap U(y) \text{ con } y \in U(x)$$

Questi insiemi rappresentano rispettivamente l'intersezione tra l'insieme degli oggetti incomparabili di y con il Down Set e con l'Up Set di x , dove y sono tutti gli elementi incomparabili rispetto a x . In questo modo si tiene conto per ogni oggetto incomparabile, $y \in U(x)$, della sua posizione relativa rispetto a x .

Il numero di oggetti contenuti in questi due insiemi viene utilizzato nella formula per la stima degli averaged rank, quindi si può scrivere:

$p_y^< = ||P_y^<||$ e $p_y^> = ||P_y^>||$ per contare gli elementi nei due insiemi sopra descritti.

Utilizzando questi dati, la stima approssimata dell'averaged rank secondo LPOMext si può formulare come segue:

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

La prima parte dell'equazione descrive l'altezza minima possibile che x può assumere in base agli elementi del Down Set, mentre la sommatoria quantifica il contributo di ciascun oggetto di $U(x)$. Il quoziente può essere interpretato come la probabilità che un oggetto $y \in U(x)$ si trovi in una posizione inferiore ad x : questo farebbe aumentare l'altezza di x di una unità. Quindi ogni oggetto appartenente a $U(x)$ contribuisce a $H(x)$ in base alla sua probabilità di essere posizionato sotto a x .

Tramite questa formula vengono aggiustati i punti deboli evidenziati da LPOM0:

- Vengono considerate solo le posizioni accessibili agli oggetti incomparabili.
- Ogni oggetto $\in U(x)$ contribuisce a $H(x)$ tramite un quoziente calcolato singolarmente.
- La causa principale per cui si ottenevano i legami artificiali con LPOM0, dove tutti gli elementi di $U(x)$ erano considerati isolati, in questo caso non sussiste.

Soggetto	O(x)	F(x)	U(x)	H(x)
A	A	A B C E G H		1
B	A B	B E H	G C	2.833
C	A C	C E G H	B	2.333
E	A B C E	E H	G	4.333
G	A C G	G H	E B	4.833
H	A B C E G H	H		6

Tabella 4.3: Down Set, Up Set, oggetti incomparabili e valore dell'averaged rank con LPOMext per gli oggetti in tabella 4.1

Basandosi sulla tabella 4.3, che riporta per ogni oggetto dell'esempio di tabella 4.1 Down Set, Up Set, oggetti incomparabili e valore dell'averaged rank, si può calcolare: $O(B)=\{A,B\}$ quindi $H(B)=2+1/3$ (come contributo di G)+ $1/2$ (come contributo di C), quindi $H(B)=2.833$, da confrontare con il valore ottenuto con il metodo LPOM0, che è 2.8. Rispetto al valore calcolato con il metodo esatto (2.8) sembra che LPOM0 goda di un'approssimazione migliore, ma questo dipende dalla struttura semplice della matrice dei dati che presenta poche coppie di oggetti incomparabili. Infatti, all'aumentare del numero di coppie incomparabili, l'approssimazione di LPOM0 tende a peggiorare nettamente mentre LPOMext dovrebbe mantenere una buona performance.

Le approssimazioni calcolate in questo modo sono soggette a tre tipi di scostamenti dalla realtà:

“Effetto combinatorio”: mentre LPOMext tiene conto soltanto della configurazione di oggetti presente, per il calcolo esatto si calcola la media di tutte le possibili combinazioni lineari.

“Effetto uno-dopo l’altro”: LPOMext esprime il collegamento di oggetti di $U(x)$ nella catena S-x-P come un processo eseguito in sequenza e non in modo simultaneo, come invece si fa con le estensioni lineari. Ciò significa che con questo approccio gli elementi di $U(x)$ contribuiscono alla somma indipendentemente dall’ordine, cioè senza considerare che gli elementi già ordinati hanno cambiato la configurazione per l’elemento successivo. Se così non fosse, il risultato sarebbe legato all’ordine con cui vengono effettuate le analisi degli oggetti nella matrice dei dati. Questo scostamento dalla realtà può essere anche interpretato come una sorta di effetto di “interazione” tra gli oggetti di $U(x)$.

Restrizioni: Gli elementi di $U(x)$ potrebbero essere collegati tra loro e in tal caso l’aggiunta di un elemento di $U(x)$ alla catena S-x-P dovrebbe limitare le posizioni disponibili per gli altri elementi di $U(x)$. Tuttavia, nel modello LPOMext, nonostante si tenga conto della relazione d’ordine tra gli elementi di $U(x)$ nel calcolo dell’average rank, la sequenzialità con cui vengono svolti i calcoli potrebbe comportare distorsioni delle relazioni tra gli elementi.

4.4.3 Classi di equivalenza e una modifica all’ LPOMext

Con i metodi descritti finora è possibile ordinare insiemi di oggetti di dimensioni contenute, che sono ben lontani dalla numerosità degli utenti dell’Ulss 15 per i quali è necessario stabilire quali siano i più fragili. Tuttavia, come già sottolineato precedentemente, è possibile trattare righe uguali della matrice di dati come oggetti equivalenti. In particolare, con un insieme di variabili dicotomiche che rappresentano le caratteristiche cliniche degli utenti c’è un numero massimo di oggetti diversi che possono esistere in una matrice di dati; questo numero dipende da quante variabili vengono utilizzate nell’analisi (se le variabili sono n , allora il numero massimo di oggetti è 2^n). Ordinare gli oggetti soltanto in base alla loro classe di appartenenza sembra un approccio piuttosto sterile in un contesto in cui si considerano gruppi di persone: il rango medio del gruppo di individui sani, che si suppone sia oggetto minimo perchè non presenta alcun fattore di

rischio, non dovrebbe essere 1 (come risulterebbe con il metodo LPOMext), ma sembra ragionevole che sia pari a metà della numerosità del gruppo (il rango medio delle persone in quel gruppo corrisponde infatti alla media dei ranghi del primo e dell'ultimo individuo).

Applicare il metodo degli LPOMext ad una matrice di dati costruita utilizzando soltanto le classi di equivalenza presenti porta a calcolare e ad attribuire un averaged rank per ogni classe di equivalenza senza tenere conto della sua numerosità. Per includere anche questa caratteristica nell'analisi, è stata fatta una modifica alla formula di LPOMext in modo da inserire anche il valore della frequenza di ciascuna classe di equivalenza, $f(x)$ (Caperna *et al.*, 2015). Si ottiene quindi:

$$H(x) = \max\{1, f(O(x)) - \frac{1}{2}f(x)\} + \sum_{y \in U(x)} \frac{p_y^<}{p_y^< + p_y^>} \cdot f(y)$$

Applicando questa approssimazione per l'averaged rank si ottengono delle stime dei ranghi maggiormente distanziate dato che, oltre ad ordinare le classi, si tiene conto del rango della prima e dell'ultima unità presente in ogni raggruppamento e che la classe successiva può iniziare soltanto dopo l'ultimo elemento di quella precedente. La prima parte della formula tiene conto della posizione media minima della classe di equivalenza in base al Down Set, con una correzione nel caso nella prima ci sia soltanto un elemento. La seconda parte pesa i contributi degli oggetti incomparabili per la loro frequenza e quindi per l'altezza che ogni classe di equivalenza y farebbe guadagnare al rango medio di x se questa si trovasse sotto all'oggetto x considerato.

Utilizzando un sottoinsieme (come esempio) di variabili per i dati relativi all'Ulss 15 nel 2013, e alle tre variabili dell'esempio in tabella 4.1 (Poliprescrizione, Assistenza domiciliare e Accessi al Pronto Soccorso), si sono calcolati gli averaged rank utilizzando il metodo LPOMext sia tenendo conto della numerosità delle classi di equivalenza sia attraverso il metodo classico; i valori normalizzati utilizzando il valore minimo e il valore massimo sono riassunti in tabella 4.4.

Si può notare che in questo caso l'averaged rank calcolato attraverso il metodo senza frequenze presenta soltanto quattro valori: questo dipende dal fatto che il Diagramma di Hasse, che è un riflesso della struttura della matrice dei dati utilizzata come esempio, è simmetrico e regolare a causa di una numerosità molto elevata, figura 4.4, che fa sì che ogni classe contenga almeno un elemento. Si può notare che i ranghi calcolati tenendo conto delle frequenze delle classi di equivalenza sono più alti rispetto agli altri: questo dipende dal fatto che la prima

Soggetto	Variabile 1: Farmaci	Variabile 2: ADI	Variabile 3: PS	LPOMext	LPOMext modificato
A	0	0	0	0	0
B	1	0	0	0.274	0.776
C	0	1	0	0.274	0.716
D	0	0	1	0.274	0.782
E	1	1	0	0.726	0.961
F	1	0	1	0.726	0.971
G	0	1	1	0.726	0.982
H	1	1	1	1	1

Tabella 4.4: Esempio di calcolo dei valori di LPOMext tenendo conto della numerosità delle classi di equivalenza

classe in basso è la più popolata e comprende più di tre quarti della popolazione totale. Si tratta di una notizia abbastanza confortante dato che è ragionevole etichettare la classe A come quella dei soggetti sani.

In una situazione di analisi reale e con un numero di variabili maggiore, il diagramma di Hasse sarà molto più complesso e non tutte le classi di equivalenza teoricamente possibili (tenendo conto di tutte le combinazioni delle variabili inserite nell'analisi) avranno una numerosità non nulla. In questi casi, calcolare gli averaged rank utilizzando le tecniche appena descritte può risultare più utile per identificare quali classi di equivalenza siano più vicine a quella massima e, quindi, quali siano i pazienti con criticità più gravi e complesse, che necessitano di interventi mirati di cura e prevenzione.

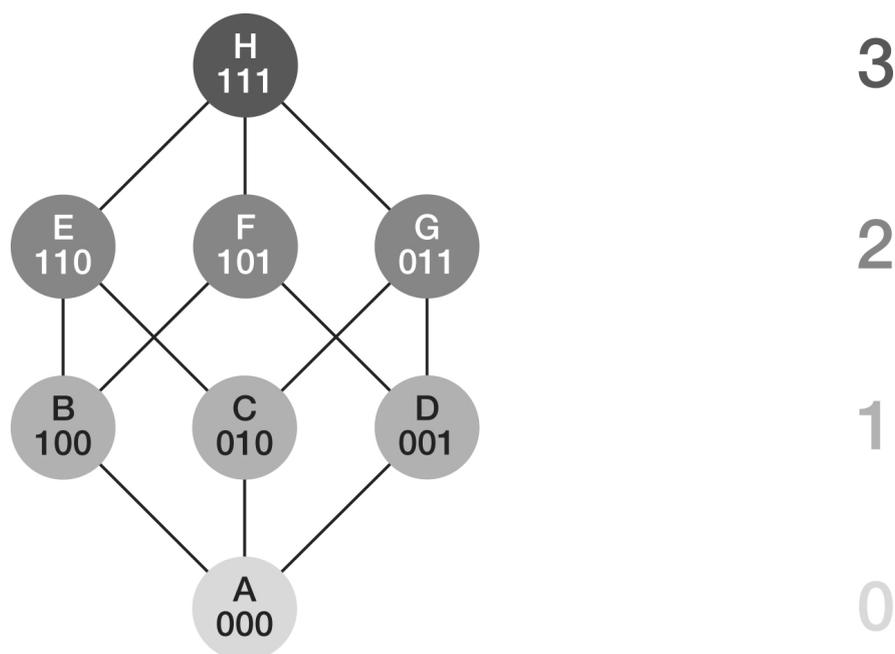


Figura 4.4: Diagramma di Hasse delle classi di equivalenza rappresentate in tabella 4.4

4.5 LPOMext per la misura della fragilità nell’Ulss 15

Le tecniche appena descritte, applicate finora soltanto in ambiti molto diversi, sono state applicate alla popolazione di utenti dell’Ulss 15 per individuare quali siano le classi di equivalenza da considerarsi più fragili. Per l’analisi è stata utilizzata un’apposita funzione in R ((Caperna *et al.*, 2015)). Con l’utilizzo di questo approccio innovativo si ottiene una stima approssimata del rango di ogni individuo basata sulle sole informazioni contenute nella matrice di dati e sulla numerosità degli individui presenti in ogni classe di equivalenza. Infatti, la modifica che tiene conto delle frequenze appena descritta sembra fornire una rappresentazione più fedele della situazione reale considerata. In questo modo, i soggetti fragili si distinguono per avere un rango particolarmente alto, vicino a 1. Il dataset utilizzato nelle analisi è quello ampiamente descritto nel capitolo 3.

Gli averaged rank stimati di seguito si sono ottenuti grazie alla modifica del metodo LPOMext che tiene conto delle numerosità delle classi di equivalenza e sono stati poi normalizzati utilizzando i valori di minimo e di massimo. Come spiegato precedentemente, i tempi di calcolo e la difficoltà computazionale aumentano con il numero delle variabili utilizzate; per questo motivo sono state prese in considerazione le variabili messe in luce nel capitolo 3 come maggiormente

significative attraverso il metodo di selezione delle variabili *stepwise*. Sono stati utilizzati 5 insiemi di variabili diversi per valutare quale fosse il più adatto ad ordinare gli utenti dell'Ulss 15.

- L'insieme di variabili A è formato dalle prime 7 che vengono scelte dal criterio di selezione stepwise nella regressione logistica per modellare la probabilità di morire o di essere ricoverati d'urgenza nel 2014, in riferimento al capitolo 3, tabella 3.5 (le classi di equivalenza teoriche risultano quindi $2^6 \cdot 3 = 192$, di cui 170 hanno un numero di elementi non nullo).

Età, che presenta tre categorie, ovvero giovani (con meno di 21 anni), adulti (con un'età compresa tra i 20 ei 70 anni) e anziani (con più di 69 anni);

Assistenza Domiciliare Integrata (ADI), aver ricevuto assistenza domiciliare almeno una volta nel 2013;

Invalidità, avere almeno una condizione di invalidità nel 2013;

Farmaci, avere una poliprescrizione (almeno 4 farmaci appartenenti a macrocategorie diverse) negli ultimi tre mesi del 2013;

Charlson index (indice relativo alle comorbidità) superiore a uno in almeno uno dei ricoveri degli ultimi due anni (2012 e 2013), si ricorda che tale indice è legato alla presenza di comorbidità;

Pronto Soccorso, aver effettuato almeno un accesso al pronto soccorso nel 2013;

Ricovero, aver avuto almeno un ricovero nel 2013 (esclusi quelli con traumi in prima diagnosi).

- L'insieme di variabili B è formato dalle 7 variabili presenti in A, con l'aggiunta della **Psichiatria**, ovvero aver effettuato almeno un accesso alla psichiatria territoriale nel 2013. Per un totale di otto variabili che danno origine a 384 classi di equivalenza teoriche ($2^7 \cdot 3 = 384$), di cui 278 hanno un numero di elementi non nullo.
- L'insieme di variabili C è formato dalle 8 variabili presenti in B, in cui viene inserita **3 + Diagnosi**, ovvero aver avuto almeno tre ricoveri con diversa prima diagnosi negli ultimi due anni, 2012 e 2013 (le classi di equivalenza teoriche risultano quindi $2^8 \cdot 3 = 768$, di cui 449 hanno un numero di elementi non nullo).
- L'insieme di variabili D è formato dalle 9 variabili presenti in C, in cui viene inserita **Scompenso cardiaco congestizio (Congestive Heart Failure)**,

CHF), ovvero aver avuto diagnosi di scompenso cardiaco congestizio negli ultimi due anni, 2012 e 2013 (le classi di equivalenza teoriche risultano quindi $2^9 \cdot 3 = 1536$, di cui 606 hanno un numero di elementi non nullo).

- L'insieme di variabili **E** è formato dalle 10 variabili presenti in **D**, in cui viene inserita **Tumori maligni**, ovvero aver avuto diagnosi di tumori maligni negli ultimi due anni, 2012 e 2013 (le classi di equivalenza teoriche risultano quindi $2^{10} \cdot 3 = 3072$, di cui 857 hanno un numero di elementi non nullo).

Si è proceduto alle aggiunte di variabili rispetto all'insieme **A** in base alle analisi svolte nel capitolo 3 secondo l'ordine di ingresso nel modello in tabella 3.5 con criterio di selezione *stepwise*.

In figura 4.5, sono riportati i risultati in sintesi delle cinque stime effettuate degli averaged rank, si può notare che l'aggiunta di una variabile non stravolge l'ordinamento delle classi in modo importante, nei valori di rango medio stimato.

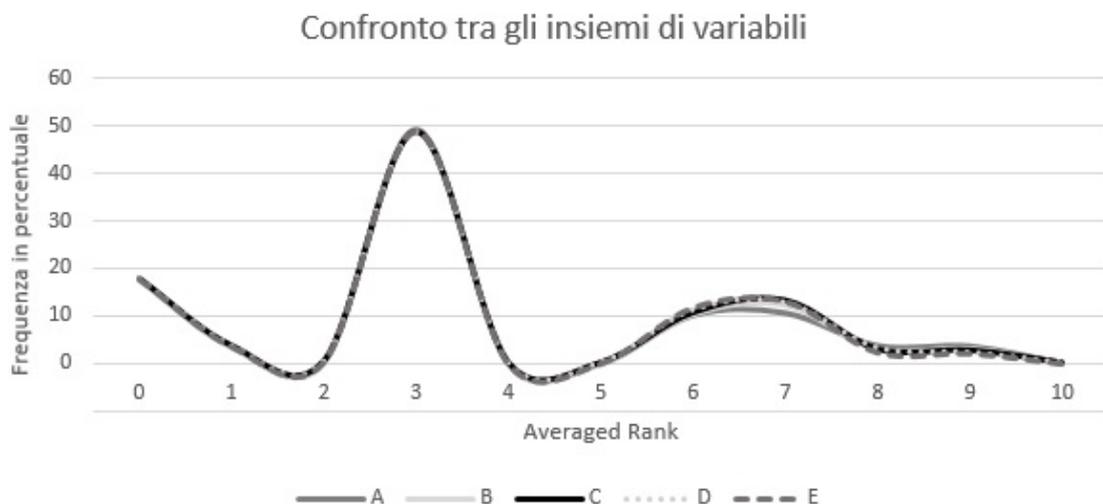


Figura 4.5: Ripartizione degli utenti dell'Ulss 15 nei decimi dell'averaged rank, confronto tra le distribuzioni dei diversi insiemi di variabili (A, B, C, D, E), valori in percentuale.

In tabella 4.5 sono riportati i coefficienti e alcune misure diagnostiche delle regressioni logistiche effettuate utilizzando la stessa variabile risposta dei modelli del capitolo 3 (aver avuto almeno un ricovero d'urgenza nel 2014 o decesso nel 2014, Y) e inserendo di volta in volta il rango medio stimato con i cinque diversi insiemi di variabili. Le analisi sono state fatte su un campione di 22260 individui composti da tutti quelli per cui $Y = 1$ e da un numero uguale di soggetti selezionati con un criterio casuale dall'insieme di individui per cui $Y = 0$. È interessante notare che con una sola variabile esplicativa e l'intercetta si riesce ad ottenere un

modello che riconosce i soggetti fragili con quasi la stessa precisione dei modelli rappresentati nel capitolo 3, che invece presentano molte più variabili. Non essendoci una significativa differenza (test di Kolmogorov-Smirnov) tra le curve prodotte utilizzando i diversi insiemi di variabili, si è optato per utilizzare le stime dell'averaged rank sull'insieme di variabili C, dotato di capacità predittive uguali o leggermente migliori rispetto agli altri due set di variabili considerati (Tabella 4.5).

	A	B	C	D	E
Intercetta	-1,702 ***	-1,720 ***	-1,750 ***	-1,750 ***	-1,751 ***
LPOMext	3,245 ***	3,320 ***	3,413 ***	3,426 ***	3,440 ***
% concordi	68,711	68,738	68,841	68,787	68,769
% discordi	31,289	31,262	31,159	31,213	31,231
Falsi positivi	29,6	29,8	29,7	29,8	29,8
Falsi negativi	32,7	32,5	32,4	32,5	32,5
Specificità	72,8	72,3	72,5	72,4	72,3
Sensibilità	64,6	65,1	65,2	65,2	65,3
R^2	0,174	0,175	0,178	0,176	0,175
R^2 corretto	0,232	0,233	0,237	0,235	0,234
Area sotto la curva ROC	0,738	0,740	0,743	0,742	0,741

Tabella 4.5: Modelli logistici per la probabilità di morte o ricovero d'urgenza nell'anno successivo a quello di riferimento stimati con solo intercetta e averaged rank. Si confrontano gli ordinamenti in base agli insiemi di variabili A, B, C, D ed E. Il livello di probabilità a cui fanno riferimento la percentuale di concordi, di discordi, di falsi positivi, di falsi negativi, specificità e sensibilità è 0.5.

Gli averaged rank calcolati in questo modo risentono molto della numerosità di individui sani giovani, adulti e anziani; si distinguono tre scalini principali nella distribuzione delle frequenze cumulate (figura 4.6): uno iniziale causato dalla numerosità dei giovani sani; uno intorno al rango medio normalizzato pari a 0.4, dovuto al fatto che agli adulti sani è stato attribuito questo valore; il terzo dislivello si colloca tra i valori 0.6 e 0.7 ed è provocato dagli anziani sani.

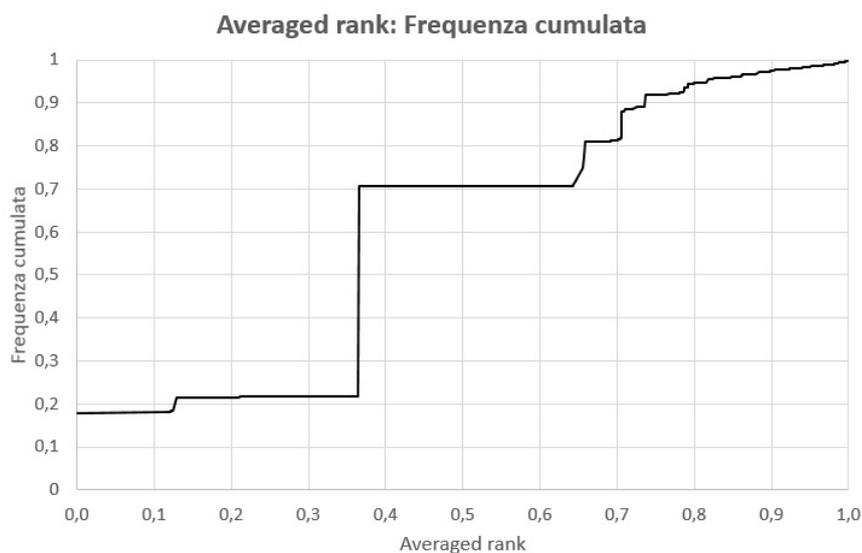


Figura 4.6: Frequenze cumulate degli utenti dell'Ulss 15 in percentuale, in base all'averaged rank.

Averaged Rank	Età		
	Giovani	Adulti	Anziani
0	81,54	0,00	0,00
1	16,56	0,00	0,00
2	1,52	0,00	0,00
3	0,08	74,55	0,00
4	0,15	0,00	0,00
5	0,03	0,00	0,00
6	0,06	10,24	30,72
7	0,03	13,54	34,44
8	0,03	1,04	17,15
9	0,01	0,63	17,65
10	0,00	0,00	0,04
Totale	54537	163750	31359

Tabella 4.6: Ripartizione degli utenti dell'Ulss 15 per fasce d'età in percentuale, in base all'averaged rank in decimi.

Gli aspetti legati all'età si possono approfondire osservando la tabella 4.6. Giovani e adulti sani hanno rispettivamente rango intorno allo 0 e pari a circa 0.35. Gli anziani hanno ranghi più variabili, ma solo a partire da un valore

minimo di 0.6: infatti nessuna persona anziana ha rango inferiore a questo valore. Osservando le classi più finali (con valori di rango superiore a 0.9) si può notare che i giovani sono praticamente assenti, gli adulti sono presenti soltanto in percentuale molto bassa, mentre quasi un quinto degli anziani cade proprio in questa fascia di rango medio stimato. In figura 4.7 sono rappresentate le frequenze cumulate degli utenti divisi per classi d'età: si può osservare che le forme delle distribuzioni risultano molto diverse tra loro (test di Kolmogorov-Smirnov significativo allo 0.01). Si può notare che gli adulti, oltre a non avere averaged rank inferiori a 0.3, non hanno alcun averaged rank tra 0.4 e 0.6. Questo dipende dal fatto che la classe di adulti sani è molto numerosa e fa sì che il primo individuo appartenente alla classe immediatamente successiva agli adulti sani abbia rango più alto (da 0.6 in poi).

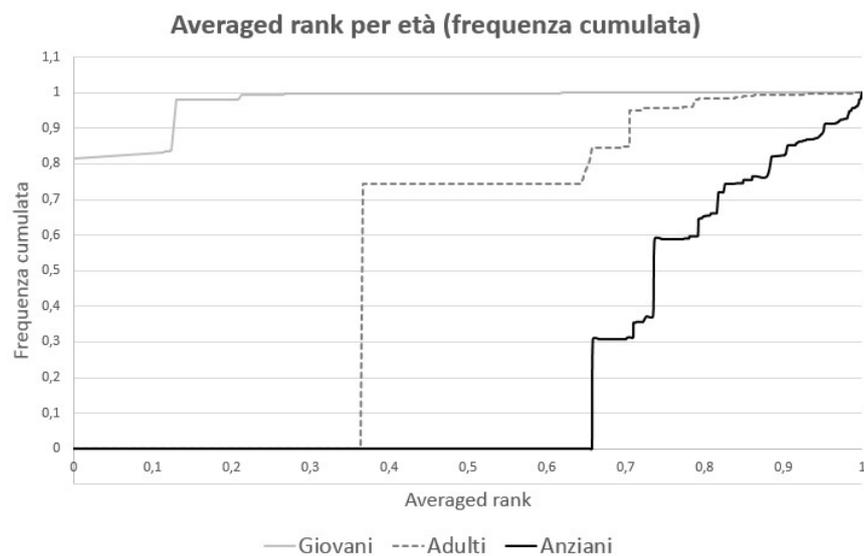


Figura 4.7: Frequenze cumulate degli utenti dell'Ulss 15 per fasce d'età in percentuale, in base all'averaged rank.

Le frequenze cumulate dei giovani partono da una baseline di giovani sani che rappresentano circa l'80% del totale, il primo scalino evidente nel grafico relativo ai giovani è originato da un gruppo numeroso di individui con meno di vent'anni che hanno avuto almeno un accesso al pronto soccorso nel 2013, che comunque presentano un rango medio molto basso.

Osservando le frequenze cumulate relative agli adulti in figura si osserva che la maggior parte di questi individui è sana (circa il 75%) e ha quindi un basso rango medio. Osservando il grafico, si notano altri due dislivelli importanti: il primo è dato dalla somma di quattro classi di equivalenza caratterizzate da adulti con un solo fattore di rischio, almeno un'invalidità, poliprescrizione, almeno un

accesso alla psichiatria territoriale nel 2013 e almeno un ricovero (esclusi quelli con un trauma in prima diagnosi); il secondo invece è causato da una sola classe di equivalenza molto numerosa, ovvero gli adulti che hanno avuto almeno un accesso al pronto soccorso nel 2013.

Averaged Rank	Assistenza Domiciliare		Invalidità		Farmaci		Charlson Index		Pronto Soccorso		Ricovero		Psichiatria Territoriale		3 o più ricoveri con diagnosi diverse		Totale
	No	Si	No	Si	No	Si	No	Si	No	Si	No	Si	No	Si	No	Si	
0	18,05	0,00	18,79	0,00	20,18	0,00	18,10	0,00	21,01	0,00	19,23	0,00	18,08	0,00	17,98	0,00	17,81
1	3,66	0,12	3,74	1,35	4,02	0,62	3,67	0,30	0,65	20,16	3,49	5,19	3,66	1,09	3,64	1,80	3,62
2	0,33	0,15	0,32	0,56	0,35	0,20	0,33	0,53	0,03	2,03	0,05	3,89	0,34	0,00	0,33	0,51	0,33
3	49,56	0,00	51,60	0,04	55,43	0,01	49,70	0,25	57,68	0,07	52,81	0,08	49,63	0,95	49,37	0,43	48,92
4	0,03	0,28	0,03	0,08	0,02	0,11	0,03	0,28	0,01	0,17	0,00	0,42	0,03	0,03	0,02	1,80	0,03
5	0,01	0,03	0,00	0,10	0,01	0,01	0,01	0,00	0,00	0,04	0,00	0,07	0,01	0,00	0,01	0,00	0,01
6	10,69	2,51	10,08	19,86	8,91	23,13	10,65	6,35	12,47	0,07	8,86	32,17	10,11	42,19	10,64	4,89	10,59
7	13,21	14,02	12,58	24,87	10,15	36,25	13,16	16,70	5,78	54,64	12,51	22,05	13,07	23,11	13,18	16,59	13,22
8	2,74	10,40	1,81	21,74	0,61	19,65	2,74	9,54	1,58	9,90	2,09	12,26	2,66	14,98	2,75	12,56	2,84
9	1,72	72,18	1,06	31,33	0,32	19,98	1,61	65,80	0,79	12,90	0,94	23,82	2,41	17,39	2,08	60,99	2,63
10	0,00	0,31	0,00	0,08	0,00	0,03	0,00	0,25	0,00	0,03	0,00	0,05	0,00	0,27	0,00	0,43	0,00
Totale	246415	3231	236649	12997	220316	29330	245663	3983	211657	37989	231195	18451	245960	3686	247313	2333	249646

Tabella 4.7: Ripartizione degli utenti dell'Ulss 15 in percentuale, in base all'averaged rank (in decimi) e alle caratteristiche anagrafiche e cliniche.

Per quanto riguarda la curva relativa alle frequenze cumulate degli anziani, si osservano due classi di equivalenza in cui è coinvolta circa la metà del totale degli anziani: i sani che rappresentano il 30.7%, e le persone con più di settant'anni che fanno un largo uso di farmaci di diverso tipo (poliprescrizione), che rappresentano il 20.6% circa del totale degli utenti con più di 70 anni. La curva relativa ai soggetti anziani, in generale, è molto più ripida della altre due appena descritte, questo perchè questi utenti sono tutti concentrati in valori di rango medio particolarmente alti.

Nella tabella 4.7 è indicato come si distribuiscono i soggetti esposti ai fattori di rischio nei diversi valori degli averaged rank (in decimi) a seconda dei fattori di rischio clinico. Si può notare che, mentre i non possessori dei fattori di rischio presentano una distribuzione di valori di averaged rank che segue l'andamento generale, coloro che invece vi sono soggetti hanno distribuzioni più varie. Per esempio, chi usufruisce dell'assistenza domiciliare ha rango in media più alto rispetto a chi non ne usufruisce.

Per alcune variabili, la metà degli esposti ha rango tra 0.9 e 1: Assistenza Domiciliare, Charlson Index e "3+ Diagnosi"; le percentuali elencate nella tabella, tuttavia, non sono calcolate al netto dell'età e di altri fattori che potrebbero influire creando distorsioni.

Averaged Rank	Esiti negativi legati alla fragilità		Fragili secondo ACG	
	No	Si	No	Si
0	18.44	4.46	17.86	11.48
1	3.7	1.79	3.61	7.62
2	0.33	0.37	0.31	4.43
3	49.89	28.13	49.02	7.79
4	0.03	0.05	0.03	0.9
5	0.01	0.01	0	0.16
6	10.49	12.72	10.61	9.43
7	12.84	21.37	13.24	14.59
8	2.49	10.47	2.82	9.59
9	1.8	20.58	2.49	33.61
10	0	0.06	0	0.41

Tabella 4.8: Ripartizione degli utenti dell'Ulss 15 per condizioni legate alla definizione di fragilità in percentuale, in base all'averaged rank.

Sfruttando la definizione utilizzata nel capitolo precedente di *outcome* relativi alla fragilità, morte o ricovero d'urgenza nell'anno successivo, si possono confrontare gli averaged rank degli individui che vi sono stati soggetti rispetto al resto della

popolazione, figura 4.8; parallelamente è interessante osservare anche i valori di averaged rank attribuiti ai soggetti fragili secondo la definizione di ACG (capitolo 2) rispetto ai non fragili, figura 4.9.

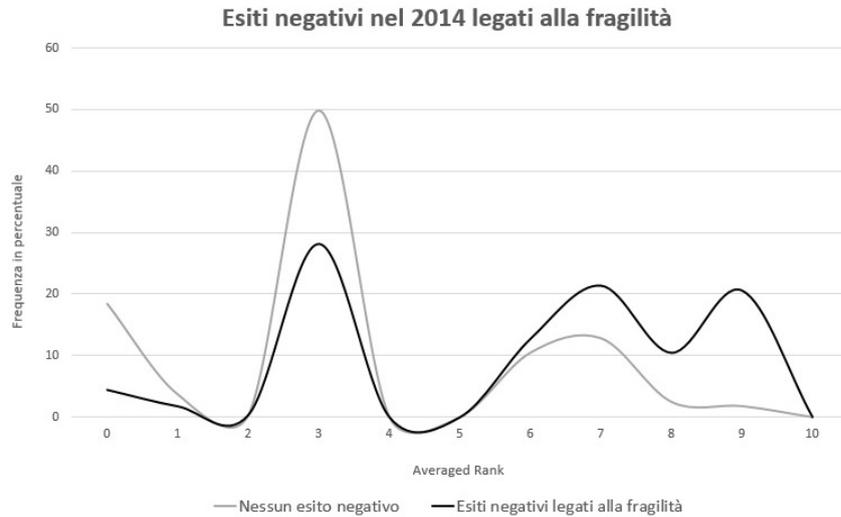


Figura 4.8: Esiti negativi nel 2014 legati alla fragilità. Distribuzione degli utenti dell'Ulss 15 in percentuale, in base all'averaged rank (in decimi).

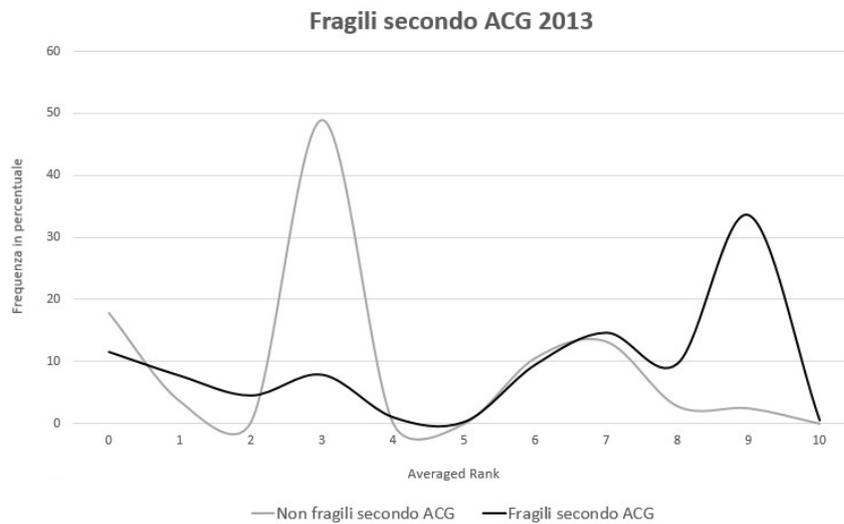


Figura 4.9: Fragili secondo ACG 2013. Distribuzione degli utenti dell'Ulss 15 in percentuale, in base all'averaged rank (in decimi).

Si può notare che tra coloro che hanno subito le prognosi negative relative alla fragilità nell'anno successivo, la maggior parte ha un rango medio stimato maggiore di 0.6 e che più del 30% di coloro che sono stati classificati come fragili dal sistema ACG ha rango medio superiore a 0.9. Anche in questo caso, come

nella tabella relativa ai fattori di rischio, si può notare che la popolazione che ACG non classifica come fragile ha un andamento che segue quello della popolazione totale, mentre la porzione di soggetti fragili ha un andamento significativamente diverso (test di Kolmogorov-Smirnov, livello di significatività 0.01).

4.5.1 Segmentazione degli utenti dell'Ulss 15 in base all'averaged rank

In questo paragrafo si mira a creare dei sottogruppi tra gli utenti totali per ridurre le classi determinate dai profili (449), in base agli averaged rank calcolati. Si effettua un'analisi di segmentazione (utilizzando il software SAS 9.3) per creare gruppi il più possibile omogenei rispetto all'averaged rank, utilizzato come variabile risposta. Vengono inserite come esplicative le nove variabili che hanno contribuito al calcolo del rango medio in analisi. Il criterio scelto per creare i nodi nella segmentazione illustrata in seguito è quello che SAS utilizza di default: Chi Quadrato con soglia pari a 0.2. In questa analisi il Misclassification rate è inferiore a 0.1 già dal primo nodo, infatti le variabili utilizzate per l'analisi sono le stesse che determinano valori diversi di rango medio.

Tramite quest'analisi di segmentazione i 449 profili vengono raggruppati in 23 gruppi, utilizzando soltanto 6 delle 9 variabili utilizzate per l'analisi (le tre variabili che non vengono utilizzate sono gli accessi alla Psichiatria Territoriale, quella relativa al Charlson Index e quella legata al numero di diverse prime diagnosi). La segmentazione, come si può vedere dalla figura 4.10 non è simmetrica, si è scelto di non forzare la simmetria in quanto non sembra realistico che, ad esempio in gruppi di età diversi, debbano esserci suddivisioni analoghe legate alle stesse variabili. Infatti, come si è visto nell'analisi con il modello loglineare, la variabile relativa all'età sviluppa interazioni significative con diverse variabili. Inoltre, come si può osservare nella figura 4.10, vi sono ampi gruppi di persone che rappresentano gli individui sani, più numerosi nel caso di giovani ed adulti, per i quali forzare una suddivisione ulteriore non avrebbe senso.

Il grafico a torta è formato da sette cerchi concentrici che rappresentano i livelli gerarchici su cui vengono effettuate le segmentazioni:

- Il primo cerchio interno rappresenta il totale degli individui che hanno averaged rank medio pari a 0.39.
- Nella prima corona circolare viene a crearsi una parziale suddivisione per età, i giovani, con meno di vent'anni, vengono separati dal resto degli individui in

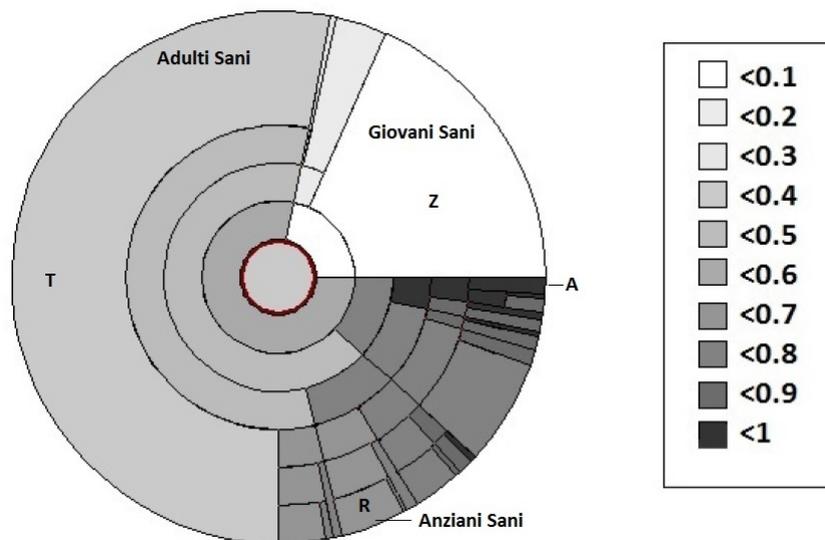


Figura 4.10: Segmentazione rappresentata con un diagramma a torta, colorazione in base all'averaged rank.

quanto sono caratterizzati da un rango medio pari a 0.026 (rispetto a 0.504 del gruppo formato da adulti e anziani).

- La seconda suddivisione è simmetrica, infatti in entrambi i gruppi la variabile utilizzata è quella relativa agli accessi al Pronto Soccorso; si mette a questo punto in evidenza già uno dei 23 gruppi definitivi il cui rango medio è pari a 0.004 dove si collocano i giovani sani.
- Nella terza fascia si distinguono tre diverse variabili che operano la suddivisione: la più influente in termini di varianza spiegata è l'età che suddivide adulti e anziani tra coloro che non hanno effettuato alcun accesso al pronto soccorso nel 2013; la seconda, in termini di influenza, è quella relativa alla Poliprescrizione che suddivide il gruppo di adulti e anziani, ancora unito, che ha effettuato almeno un accesso al pronto soccorso nel 2013; la terza variabile suddivide i giovani che hanno avuto almeno un accesso al pronto soccorso nel 2013 in due gruppi in base al fatto di essere stati ricoverato o no.
- Nelle ultime tre suddivisioni (che avvengono soltanto per adulti e anziani) si utilizzano principalmente quattro variabili: l'età per distinguere adulti e anziani, la poliprescrizione, quella relativa ai ricoveri (esclusi quelli con traumi in prima diagnosi) e l'invalidità. Nell'ultima suddivisione viene aggiunta anche la variabile relativa all'assistenza domiciliare per suddividere il gruppo di persone anziane invalide e con poliprescrizioni.

- L'ultima corona circolare è suddivisa nelle 23 sezioni che rappresentano i gruppi identificati con delle lettere (che ordinano i gruppi da quello con averaged rank più alto, contrassegnato dalla lettera A, a quello con averaged rank più basso, contrassegnato dalla lettera Z) nella tabella in figura 4.11; la sezione più ampia rappresenta gli adulti sani (gruppo T), la seconda per dimensioni rappresenta i giovani sani (gruppo Z), mentre la sezione che raccoglie gli anziani sani non è facilmente riconoscibile per le sue dimensioni ed è contrassegnata dalla lettera R.

Osservando la figura 4.11 si può notare che il gruppo di soggetti più fragili, secondo questo criterio di analisi, è costituito da 2446 soggetti anziani che fanno un largo utilizzo di farmaci (poliprescrizione) e che nel 2013 hanno avuto almeno un ricovero (senza considerare quelli con traumi come prima diagnosi) e un accesso al pronto soccorso; in questo gruppo, l'averaged rank medio è molto alto, cioè pari a 0.981. Seguendo l'ordine, il gruppo B è il secondo gruppo più fragile ed è costituito da anziani invalidi che consumano almeno quattro farmaci appartenenti a macrocategorie diverse e che, sebbene non siano stati ricoverati nel 2013, hanno avuto almeno un accesso al pronto soccorso in quest'anno (l'averaged rank medio in questo gruppo è pari a 0.964). Nei primi quattro gruppi vi sono soltanto individui anziani; il quinto gruppo (E) è caratterizzato da persone adulte con poliprescrizione che nel 2013 hanno effettuato sia un accesso al pronto soccorso, sia sono state ricoverate, la media degli averaged rank in questo gruppo è 0.913. Il gruppo di giovani che presenta rango medio più alto è quello formato da individui che hanno sia un ricovero sia un accesso al pronto soccorso nel 2013; la media degli averaged rank in questo gruppo è comunque molto bassa (0.261), ma con uno standard error più grande rispetto a quello di tutti gli altri gruppi: gli individui in questa sezione (U) non sono molti, ma si tratta di soggetti particolarmente eterogenei, tra cui probabilmente si trova anche qualche situazione critica o problematica di fragilità che viene trascurata a causa della giovane età.

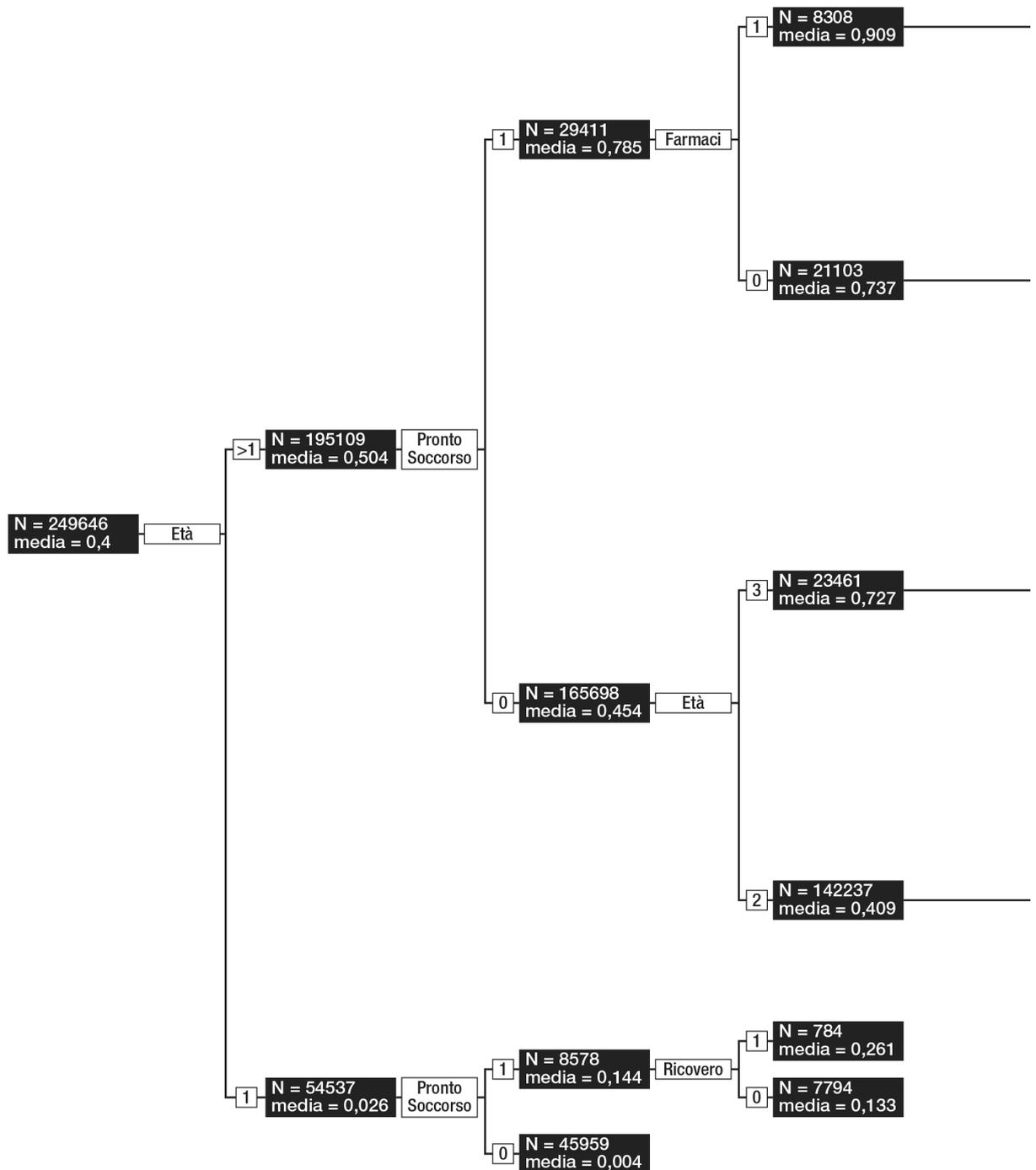


Figura 4.12: Analisi di segmentazione degli utenti dell'Ulss 15 in base all'averaged rank. Parte 1. La variabile età ha tre modalità corrispondenti alle tre macro-classi: Giovani (età=1), Adulti (età=2) e Anziani (età=3).

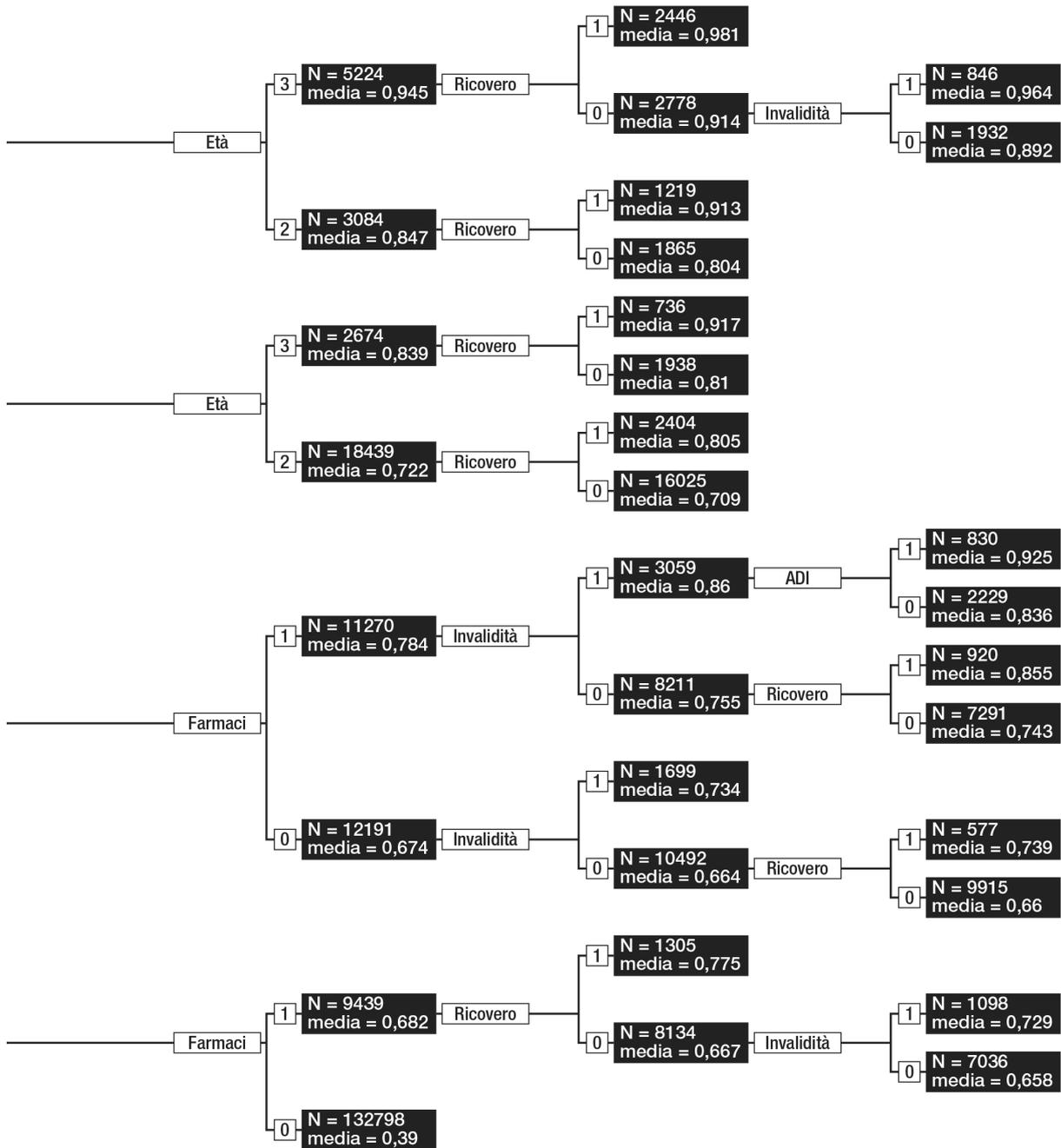


Figura 4.13: Analisi di segmentazione degli utenti dell’Ulss 15 in base all’averaged rank. Parte 2. La variabile età ha tre modalità corrispondenti alle tre macro-classi: Giovani (età=1), Adulti (età=2) e Anziani (età=3).

Conclusioni

La fragilità è un fenomeno complesso e multidimensionale per cui non esiste ancora un accordo univoco su una possibile definizione. Alcuni fattori che la caratterizzano sono tuttavia ampiamente condivisi: il fatto che si tratti di una patologia multifattoriale, che sia legata ad un peggioramento delle condizioni di salute e ad una perdita dell'autonomia e che esponga l'individuo che ne è affetto a maggiori rischi e a minori capacità di ripresa rispetto a stress ambientali anche minimi. In mancanza di una definizione concettuale condivisa, non c'è accordo nemmeno su un possibile criterio operativo per individuare i soggetti fragili.

Nel presente elaborato si sono approfonditi alcuni metodi esistenti mettendone in luce punti di forza e debolezze ed è stato proposto un nuovo strumento per la classificazione dei soggetti fragili, introducendo un percorso metodologico ancora inedito in quest'ambito.

Dall'analisi approfondita delle caratteristiche dei soggetti fragili secondo il sistema ACG è emerso che tale popolazione ha effettivamente un pattern di criticità più problematico rispetto al resto della popolazione dell'Ulss 15. Tuttavia, essendo molto legato al quadro teorico di definizione delle variabili, questo strumento non è un predittore sufficiente ad identificare gli individui che l'anno successivo a quello di riferimento potrebbero incorrere negli esiti avversi legati alla fragilità. La capacità predittiva di uno strumento per l'identificazione dei soggetti fragili è molto importante per mettere in atto interventi di prevenzione.

Nel capitolo 3, viene descritto ed applicato agli utenti dell'Ulss 15 un criterio per l'identificazione dei soggetti fragili basato su un modello predittivo. Viene stimata una regressione logistica utilizzando come variabile risposta la morte o il ricovero d'urgenza nell'anno successivo a quello di riferimento, prendendo spunto dal modello MoSaiCo elaborato per l'Ulss di Ravenna. Per ottenere un modello con variabili interessanti anche da un punto di vista dell'interpretazione sostantiva, e non soltanto statisticamente significative, è stata fatta un'accurata scelta e scrematura rispetto al modello MoSaiCo di riferimento, ottenendo un insieme molto più parsimonioso, ma paragonabile in termini di varianza spiegata.

Gli utenti dell'Ulss 15 sono stati stratificati in base ai valori di rischio predetti dal modello definito nel corso del capitolo 3. Tale criterio è stato costruito in modo da poter prevedere gli esiti avversi legati alla fragilità, ma si rischia a questo punto di essere troppo vincolati alla definizione utilizzata come “esiti avversi” ovvero la morte o il ricovero d'urgenza nell'anno successivo a quello di riferimento. Questi due eventi, utilizzati come variabile dipendente, risultano sicuramente utili per validare una misura di fragilità, ma rappresentano un criterio soggettivo e discutibile se impiegati come parte integrante della definizione.

Nella parte finale della tesi si cerca di trovare un metodo che non sia vincolato da una variabile dipendente, ma che sia comunque in grado di fornire una classificazione dei soggetti fragili. Per ottenere questo risultato si utilizzano le teorie sugli ordinamenti parziali (poset).

Nonostante finora le teorie sui poset non siano ancora mai state applicate a contesti epidemiologici, e soprattutto con numerosità così grandi, l'applicazione appena descritta mette in luce un buon adattamento che apre la strada ad un percorso di analisi nuovo e promettente.

Le potenzialità di questo metodo per identificare gli individui fragili sono innegabili, infatti con un percorso libero da assunzioni soggettive iniziali si ottiene un ordinamento degli individui in base al grado di fragilità che permette alla struttura ordinale dei dati in matrice di “parlare da sola” e di creare ordinamenti anche laddove indicatori equiponderati non creerebbero distinzioni. Inoltre, in questo modo la gerarchia delle variabili non viene imposta, ma prende forma grazie alla struttura della matrice di dati.

Il metodo LPOMext modificato per tenere conto delle frequenze delle classi di equivalenza, in particolare, sintetizza l'informazione della matrice di dati per creare un ordinamento inedito, inserendo anche una dimensione che altrimenti verrebbe ignorata. Ad ogni classe viene associato un rango che non soltanto tiene conto delle caratteristiche qualitative e di numerosità della classe stessa, ma anche di quelle delle altre classi di equivalenza presenti nel dataset di riferimento; garantendo una visione d'insieme delle condizioni di salute degli individui dell'Ulss 15 utile e fondamentale alla programmazione ed alla gestione degli interventi di prevenzione e di screening.

Concludendo, si sono ottenuti 23 gruppi che descrivono le caratteristiche di salute degli utenti dell'Ulss 15 sia secondo un criterio quantitativo, l'averaged rank, di gravità dello stato di salute, sia da un punto di vista qualitativo che ne rende agevole l'interpretazione. A questo punto delle analisi, ragionare su un valore di soglia per dicotomizzare un'informazione rispetto alle condizioni di salute degli

individui dell'Ulss 15 non sembra una priorità. Si ha infatti una classificazione graduata degli individui che fornisce una semplificazione di qualità di un concetto complesso e multidimensionale come la fragilità.

La teoria degli ordinamenti parziali si è molto bene adattata al contesto epidemiologico di identificazione dei soggetti fragili. Tuttavia, c'è ancora ampio spazio di analisi su questa strada, sarebbe interessante approfondire il ruolo che l'età assume nell'ordinamento e se valga la pena effettuare delle stime per gli averaged rank separatamente in base alla classe d'età. Inoltre per le analisi svolte nel capitolo 4 sono state utilizzate soltanto variabili dicotomiche (ad eccezione appunto della variabile relativa all'età); sarebbe interessante introdurre nell'Information Base variabili ordinali con più modalità per distinguere i soggetti fragili anche in base all'intensità dei fattori di rischio. Inoltre un aspetto da approfondire per le future applicazioni è quello relativo alla misura della bontà dell'ordinamento, soprattutto in applicazioni pratiche in cui l'ordinamento deve rispecchiare delle caratteristiche qualitative degli oggetti.

Appendice

Tabella 4.9: Modello loglineare: effetti stimati e rispettivo livello di significatività.

Variabili	DF	Chi-Square	$Pr > ChiQuadr$
ADI	1	52,82	<,0001
Y	1	0,53	0,4648
ADI*Y	1	24,51	<,0001
Tumore Maligno	1	0,52	0,4709
ADI*Tumore Maligno	1	3,29	0,0699
Y*Tumore Maligno	1	18,19	<,0001
Charlson Index	1	82,11	<,0001
ADI*Charlson Index	1	11,45	0,0007
Y*Charlson Index	1	1,96	0,1615
Tumore Maligno*Charlson Index	1	2,16	0,1421
Scompenso Cardiaco	1	4,07	0,0438
ADI*Scompenso Cardiaco	1	1,56	0,2113
Y*Scompenso Cardiaco	1	10,28	0,0013
Tumore Maligno*Scompenso Cardiaco	1	2,37	0,1238
Charlson Index*Scompenso Car- diaco	1	95,86	<,0001
BPCO	1	6,9	0,0086
ADI*BPCO	1	18,05	<,0001
Y*BPCO	1	30,9	<,0001
Tumore Maligno*BPCO	1	0,69	0,4062
Charlson Index*BPCO	1	52,63	<,0001
Scompenso Cardiaco*BPCO	1	0,99	0,3189
3 + Diagnosi	1	0,62	0,4319
ADI*3 + Diagnosi	1	1,27	0,2596

Continua nella prossima pagina

Continua dalla pagina precedente

Variabili	DF	Chi-Square	Pr > ChiQuadr
Y*3 + Diagnosi	1	0,12	0,7248
Tumore Maligno*3 + Diagnosi	1	90,16	<,0001
Charlson Index*3 + Diagnosi	1	0	0,9636
Scompenso Cardiaco*3 + Diagnosi	1	38,41	<,0001
BPCO*3 + Diagnosi	1	11,71	0,0006
Poliprescrizione	1	3,6	0,0577
ADI*Poliprescrizione	1	119,04	<,0001
Y*Poliprescrizione	1	129,01	<,0001
Tumore Maligno*Poliprescrizione	1	14,87	0,0001
Charlson Index*Poliprescrizione	1	0,16	0,6895
Scompenso Cardiaco*Poliprescrizione	1	6,19	0,0129
BPCO*Poliprescrizione	1	0,43	0,5136
3 + Diagnosi*Poliprescrizione	1	11,43	0,0007
Invalidità	1	2,4	0,1216
ADI*Invalidità	1	3,26	0,0708
Y*Invalidità	1	2,15	0,1422
Tumore Maligno*Invalidità	1	5,15	0,0233
Charlson Index*Invalidità	1	0,9	0,3416
Scompenso Cardiaco*Invalidità	1	0,03	0,8525
BPCO*Invalidità	1	1,12	0,2898
3 + Diagnosi*Invalidità	1	15,77	<,0001
Poliprescrizione*Invalidità	1	53,91	<,0001
Ricovero	1	88,68	<,0001
ADI*Ricovero	1	22,94	<,0001
Y*Ricovero	1	70,51	<,0001
Tumore Maligno*Ricovero	1	0,01	0,9352
Charlson Index*Ricovero	1	66,55	<,0001
Scompenso Cardiaco*Ricovero	1	8,74	0,0031
BPCO*Ricovero	1	17,34	<,0001
3 + Diagnosi*Ricovero	1	2,57	0,1087
Poliprescrizione*Ricovero	1	6,04	0,014

Continua nella prossima pagina

Continua dalla pagina precedente

Variabili	DF	Chi-Square	Pr > ChiQuadr
Invalità*Recupero	1	6,52	0,0106
Pronto Soccorso	1	0,38	0,5373
ADI*Pronto Soccorso	1	25,64	<,0001
Y*Pronto Soccorso	1	115,51	<,0001
Tumore Maligno*Pronto Soccorso	1	3,6	0,0577
Charlson Index*Pronto Soccorso	1	9,26	0,0023
Scompenso Cardiaco*Pronto Soccorso	1	4,31	0,0378
BPCO*Pronto Soccorso	1	0,17	0,6774
3 + Diagnosi*Pronto Soccorso	1	3,34	0,0676
Poliprescrizione*Pronto Soccorso	1	147,55	<,0001
Invalità*Pronto Soccorso	1	3,38	0,0658
Recupero*Pronto Soccorso	1	19,21	<,0001
Psichiatria	1	6,87	0,0087
ADI*Psichiatria	1	1,34	0,2471
Y*Psichiatria	1	0,1	0,7518
Tumore Maligno*Psichiatria	1	0,27	0,6007
Charlson Index*Psichiatria	1	9	0,0027
Scompenso Cardiaco*Psichiatria	1	0,2	0,6572
BPCO*Psichiatria	1	0,55	0,459
3 + Diagnosi*Psichiatria	1	17,22	<,0001
Poliprescrizione*Psichiatria	1	1,03	0,3113
Invalità*Psichiatria	1	9,6	0,0019
Recupero*Psichiatria	1	116,32	<,0001
Pronto Soccorso*Psichiatria	1	2,81	0,0936
Deprivazione	1	3,11	0,0776
ADI*Deprivazione	1	4,53	0,0333
Y*Deprivazione	1	2,25	0,1336
Tumore Maligno*Deprivazione	1	29,4	<,0001
Charlson Index*Deprivazione	1	3,28	0,0701
Scompenso Cardiaco*Deprivazione	1	6,1	0,0135
BPCO*Deprivazione	1	1,46	0,2262

Continua nella prossima pagina

Continua dalla pagina precedente

Variabili	DF	Chi-Square	Pr > ChiQuadr
3 + Diagnosi*Deprivazione	1	7,81	0,0052
Poliprescrizione*Deprivazione	1	0,5	0,4784
Invalità*Deprivazione	1	0,27	0,6034
Ricovero*Deprivazione	1	0,03	0,873
Pronto Soccorso*Deprivazione	1	0,16	0,6876
Psichiatria*Deprivazione	1	14,14	0,0002
Sesso	1	0,3	0,5868
ADI*Sesso	1	33,67	<,0001
Y*Sesso	1	6,71	0,0096
Tumore Maligno*Sesso	1	5,6	0,0179
Charlson Index*Sesso	1	0,22	0,6355
Scompenso Cardiaco*Sesso	1	0,69	0,4046
BPCO*Sesso	1	7,56	0,006
3 + Diagnosi*Sesso	1	1,53	0,2164
Poliprescrizione*Sesso	1	1,44	0,2304
Invalità*Sesso	1	0,07	0,7869
Ricovero*Sesso	1	2,25	0,1334
Pronto Soccorso*Sesso	1	6,52	0,0107
Psichiatria*Sesso	1	0	0,9688
Deprivazione*Sesso	1	0,01	0,9229
ADI*Tumore Maligno*Charlson Index	1	4,47	0,0346
Tumore Maligno*Charlson Index*3 + Diagnosi	1	17,1	<,0001
Charlson Index*Scompenso Cardiaco*3 + Diagnosi	1	12,32	0,0004
Charlson Index*BPCO*3 + Diagnosi	1	7,18	0,0074
Tumore Maligno*Charlson Index*Poliprescrizione	1	26,6	<,0001
ADI*Tumore Maligno*Invalità	1	11,57	0,0007
ADI*Charlson Index*Invalità	1	27,13	<,0001
Tumore Maligno*Charlson Index*Invalità	1	9,51	0,002

Continua nella prossima pagina

Continua dalla pagina precedente

Variabili	DF	Chi-Square	Pr > ChiQuadr
Tumore Maligno*Charlson Index*Ricovero	1	61,93	<,0001
Charlson Index*Scompenso Cardiaco*Ricovero	1	48,4	<,0001
Charlson Index*BPCO*Ricovero	1	34,77	<,0001
Tumore Maligno*3 + Diagnosi*Ricovero	1	30,92	<,0001
Charlson Index*3 + Diagnosi*Ricovero	1	61,15	<,0001
Scompenso Cardiaco*3 + Diagnosi*Ricovero	1	8,77	0,0031
Charlson Index*Invalidità*Ricovero	1	17,73	<,0001
Poliprescrizione*Invalidità*Ricovero	1	82,41	<,0001
ADI*Invalidità*Pronto Soccorso	1	33,1	<,0001
Tumore Maligno*Ricovero*Pronto Soccorso	1	25,07	<,0001
ADI*Tumore Maligno*Psichiatria	1	9,06	0,0026
ADI*Invalidità*Psichiatria	1	28,22	<,0001
Y*Invalidità*Psichiatria	1	18,83	<,0001
Poliprescrizione*Invalidità*Psichiatria	1	5,54	0,0185
Charlson Index*Pronto Soccorso*Psichiatria	1	37,79	<,0001
Tumore Maligno*Ricovero*Deprivazione Ricovero*Psichiatria*Deprivazione	1	6,13	0,0133
Y*Tumore Maligno*Sesso	1	6,97	0,0083
Y*Poliprescrizione*Sesso	1	8,15	0,0043
Poliprescrizione*Pronto Soccorso*Sesso	1	57,99	<,0001
Charlson Index*Deprivazione*Sesso	1	13,99	0,0002
	1	6,5	0,0108

Continua nella prossima pagina

Continua dalla pagina precedente

Variabili	DF	Chi-Square	Pr > ChiQuadr
Invalità*Deprivazione*Sesso	1	13,91	0,0002
Y*Tumore Maligno*eta	2	952,76	<,0001
Tumore Maligno*Charlson Index*eta	2	56,4	<,0001
Charlson Index*3 + Diagnosi*eta	2	48,12	<,0001
Tumore Maligno*Poliprescrizione*eta	2	25,15	<,0001
Charlson Index*Poliprescrizione*eta	2	83,99	<,0001
Tumore Maligno*Invalità*eta	2	55,37	<,0001
BPCO*Invalità*eta	2	8,63	0,0133
Poliprescrizione*Invalità*eta	2	237,3	<,0001
Invalità*Ricovero*eta	2	33,81	<,0001
Ricovero*Pronto Soccorso*eta	2	44,22	<,0001
Poliprescrizione*Psichiatria*eta	2	112,64	<,0001
Invalità*Psichiatria*eta	2	24,06	<,0001
Y*Sesso*eta	2	75,26	<,0001
Poliprescrizione*Sesso*eta	2	15,16	0,0005
Invalità*Sesso*eta	2	44,95	<,0001
Ricovero*Sesso*eta	2	110,61	<,0001
Pronto Soccorso*Sesso*eta	2	15,47	0,0004
Rapporto di verosimiglianza	2000,00	2095,57	<,0001

Tabella 4.10: Modello loglineare: Stime delle interazioni triple contenenti la variabile Y.

Variabili	Casi	Stima	Standard Error	ChiQuadr	Pr > ChiQuadr
Y*Invalità*Psichiatria	000	0,1213	0,0279	18,83	<,0001
	001	-0,1213			
	010	-0,1213			
	011	0,1213			

Continua nella prossima pagina

Continua dalla pagina precedente

Variabili	Casi	Stima	SE	ChiQuadr	Pr > ChiQ
	100	-0,1213			
	101	0,1213			
	110	0,1213			
	111	-0,1213			
Y*Tumore Maligno*Sesso	000	-0,0813	0,0285	8,15	0,0043
	001	0,0813			
	010	0,0813			
	011	-0,0813			
	100	0,0813			
	101	-0,0813			
	110	-0,0813			
	111	0,0813			
Y*Poliprescrizione*Sesso	000	-0,0828	0,0109	57,99	<,0001
	001	0,0828			
	010	0,0828			
	011	-0,0828			
	100	0,0828			
	101	-0,0828			
	110	-0,0828			
	111	0,0828			
Y*Tumore Maligno*eta	001	0,4152	0,016	676,97	<,0001
	002	0,0428	0,011	15,14	<,0001
	003	-0,458			
	011	-0,4152			
	012	-0,0428			
	013	0,458			
	101	-0,4152			
	102	-0,0428			
	103	0,458			
	111	0,4152			
	112	0,0428			
	113	-0,458			

Continua nella prossima pagina

Continua dalla pagina precedente

Variabili	Casi	Stima	SE	ChiQuadr	Pr > ChiQ
Y*Sesso*eta	001	0,1056	0,0157	45,34	<,0001
	002	-0,0861	0,0108	63,82	<,0001
	003	-0,0195			
	011	-0,1056			
	012	0,0861			
	013	0,0195			
	101	-0,1056			
	102	0,0861			
	103	0,0195			
	111	0,1056			
	112	-0,0861			
	113	-0,0195			

Ringraziamenti

Giunta al termine di questo percorso, desidero ringraziare di cuore tutte le persone che mi hanno accompagnata in questi cinque anni, incoraggiandomi, supportandomi, sopportandomi e credendo in me.

In particolare ringrazio la Professoressa Giovanna Boccuzzo non solo per quello che mi ha insegnato, ma anche perché senza la sua pazienza e disponibilità infinita questa tesi non sarebbe ancora conclusa. Grazie anche al Dottor Giulio Caperna che mi ha generosamente dedicato il suo tempo dandomi un prezioso aiuto e suggerimenti indispensabili.

Un sentito ringraziamento anche al Dottor Armando Olivieri e al Dottor Luca Benacchio che mi hanno premurosamente accompagnata nella mia esperienza di stage, proponendomi questo argomento per la tesi che si è dimostrato utile e interessante e fornendomi dati e strumenti necessari a svolgerla al meglio.

Grazie a mia mamma che più di tutti ha contribuito a farmi completare questo percorso, aiutandomi in tutti i modi, come solo una mamma può. Ringrazio tutta la mia famiglia, che anche se devono ancora capire bene cos'ho studiato per tutti questi anni mi hanno sempre incoraggiata e appoggiata. In particolare grazie anche a mia cugina Giuditta per le consulenze di latex e a Paolo per il contributo grafico.

Ringrazio anche Federico che è sempre al mio fianco e sa sempre cosa dire per spronarmi e sostenermi e che ormai capisce di cosa ho bisogno, anche con solo uno sguardo. Senza dilungarmi, sono davvero felice che sia nella mia vita e che voglia continuare a farne parte.

Grazie alle mie amiche Francesca, Annapaola e Marcela con cui ho vissuto e sto vivendo un rapporto quasi simbiotico, per il supporto morale nei momenti di pazzia, di gioia e di sconforto. Un caro grazie anche a Riccardo che mi ha offerto asilo nella casa dei festini e mi ha davvero sopportata.

Un caro grazie ai Fioi per tutte le belle serate, le giornate spensierate e i momenti speciali che mi avete fatto vivere lontano dai libri, perché ci vuole anche quello.

Avrò sicuramente dimenticato qualcuno, perché le persone che in questo percorso mi hanno fatto sentire sostenuta e amata sono davvero tante. Non sono certa che quanto ho scritto in queste poche righe renda l'idea di quanto mi sento grata e riconoscente nei confronti di tutti voi. Grazie a tutti!!!

Bibliografia e Sitografia

- Bargagli, A.M., De Sario, M., Michelozzi, P., Vecchi, S., Cialdea, L. *et al.* (2013). *Linee di indirizzo per la prevenzione degli effetti del caldo sulla salute*. Ministero della Salute. Ultimo accesso: 07/08/2015. URL: http://www.salute.gov.it/imgs/C_17_pubblicazioni_1959_allegato.pdf.
- Bergman, H., Béland, F., Karunanathan, S., Hummel, S., Hogan, D. *et al.* (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. *et al.* (2007). *Frailty: An Emerging Research and Clinical Paradigm - Issues and Controversies*. *Journal of Gerontology* 62A.7, 731–737.
- Brüggemann, R. & Carlsen, L. (2011). *An Improved Estimation of Averaged Ranks of Partial Orders*. *Communications in Mathematical and in Computer Chemistry* 65, 383–414.
- Brüggemann, R. & Patil, G.P. (2010). *Ranking and Prioritization with Multiple Indicators Introduction to Partial Order Applications*. Department of Statistics The Pennsylvania State University.
- Caperna, G., Boccuzzo, G. & Brüggemann, R. (2015). *Measuring well-being: poset theory as a tool for synthesis of information from a big set of data*. Dealing with complexity in society: from plurality of data to synthetic indicators.
- Caranci, N., Biggeri, A., Grisotto, L., Pacelli, B., Spadea, T. *et al.* (2010). *L'indice di deprivazione italiano a livello di sezione di censimento: definizione, descrizione e associazione con la mortalità*. *Epidemiologia e Prevenzione* 34.4, 167–176.
- Cavazza, G. & Malvi, C. (2014). *La fragilità degli anziani. Strategie, progetti, strumenti per invecchiare bene*. Santarcangelo di Romagna: Maggioli Editore.
- Consiglio Sanitario Regionale della Toscana (2013). *La fragilità dell'anziano*. A cura di Regione Toscana. Ultimo accesso: 07/08/2015. URL: https://www.ars.toscana.it/files/aree_intervento/salute_di_anziani/lg_toscana_fragilita.pdf.

- De Vriesa, N.M., Staal, J.B, Ravensberg, C.D van, Hobbelenc, J.S.M., Olde Rikkert, M.G.M. *et al.* (2011). *Outcome instruments to measure frailty: A systematic review.* Ageing Research Reviews 10, 104–114.
- 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.
- 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.
- 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.
- Fried, L.P., Tangen, C.M., Walston, J., Newman, A.B., Hirsch, C. *et al.* (2001). *Frailty in Older Adults. Evidence for a Phenotype.* The Journals of Gerontology 56.3, 146–157.
- Giunta Regionale del Veneto (2012). *Il sistema ACG (Adjusted Clinical Group) per la valutazione e gestione del case-mix territoriale nella Regione Veneto: progetto pilota per l’adozione di strumenti innovativi di misurazione del rischio, analisi, programmazione.* Ultimo accesso: 07/08/2015. URL: <http://bur.regione.veneto.it/BurvServices/Pubblica/DettaglioDgr.aspx?id=238986>.
- Gobbens, R.J.J., Luijkx, 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.
- Health Dialog King’s Fund, New York University. (2006). *Combined predictive model. Final Report & technical documentation.* London: King’s Fund. Ultimo accesso: 30/08/2015. URL: www.kingsfund.org.uk/document.rm?id=6745..
- Hebert, R., Bravo, G., Korner-Bitensky, N. & Voyer, L. (1996). *Predictive Validity of a Postal Questionnaire for Screening Community-dwelling Elderly Individuals at Risk of Functional Decline.* Age and Ageing 25, 159–167.
- Karunanathan, S., Wolfson, C., Bergman, H., Béland, F. & Hogan, D.B. (2009). *A multidisciplinary systematic literature review on frailty: Overview of the methodology used by the Canadian Initiative on Frailty and Aging.* BMC Medical Research Methodology 9.68.
- Katz, S., Ford, A.B., Moskowitz, R.W., Jackson, B.A. & Jaffe, M.W. (1963). *Studies of illness in the aged. The index of ADL: a standardized measure of*

- biological and psychosocial function*. The Journal of the American Medical Association 185, 914–919.
- Marcon, A., Accorsi, A., Di Tommaso, F., Falasca, P., Berardo, A. *et al.* (2010). *Frailty in older people: analysis of literature from 1983 to 2009*. Journal of Gerontology 58, 179–183.
- Martin, F.C. & Brighton, P. (2008). *Frailty: different tools for different purposes?*: Age and Ageing 37, 129–131.
- Ministero del Lavoro, della Salute e delle Politiche Sociali (2007). *Classificazione delle malattie, dei traumatismi, degli interventi chirurgici e delle procedure diagnostiche e terapeutiche, Versione italiana della icd-9-cm*. Dipartimento della Qualità. Ultimo accesso: 07/08/2015. URL: http://www.salute.gov.it/imgs/C_17_pubblicazioni_2251_allegato.pdf.
- (2008). *Classificazione Statistica Internazionale delle Malattie e dei Problemi Sanitari Correlati Decima Revisione*. Sistema Epidemiologico Regionale del Veneto (CRRC-SER). Ultimo accesso: 07/08/2015. URL: http://www.ser-veneto.it/public/File/documents/documentazione_tecnica/classificazionevol12008.pdf.
- Razzanelli, M., Profili, F., Mossello, E., Bandinelli, S., Corridori, C. *et al.* (2013). *Screening della fragilità e valutazione multidimensionale degli anziani residenti a domicilio per la prevenzione secondaria della disabilità: risultati di uno studio pilota*. Epidemiologia e Prevenzione 37, 271–278.
- Reid, R., MacWilliam, L., Roos, N.P., Bogdanovic, B. & Black, C. (1999). *Measuring Morbidity in Populations: Performance of the Johns Hopkins Adjusted Clinical Group (ACG) Case-Mix Adjustment System in Manitoba*. Manitoba Centre for Health Policy & Evaluation, Department of Community Health Sciences, Faculty of Medicine, University of Manitoba. Ultimo accesso: 07/08/2015. URL: <http://mchp-appserv.cpe.umanitoba.ca/reference/acg.pdf>.
- Rockwood, K. (2005). *What would make a definition of frailty successful?*: Age and Ageing 34, 432–434.
- Rockwood, K. & Mitnitski, A. (2007). *Frailty in Relation to the Accumulation of Deficits*. Journal of Gerontology 62A.7, 722–727.
- 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, R., Andrew, M. & Mitnitski, A. (2007). *A Comparison of Two Approaches to Measuring Frailty in Elderly People*. Journal of Gerontology 62A.7, 738–743.

- Rodríguez-Mañas, L., Féart, C., Mann, G., Viña, J., Chatterji, S. *et al.* (2013). *Searching for an Operational Definition of Frailty: A Delphi Method Based Consensus Statement. The Frailty Operative Definition-Consensus Conference Project.* Journal of Gerontology 68, 62–67.
- Schuurmans, H., Steverink, N., Lindenberg, S., Frieswijk, N. & Slaets, J. P. J. (2004). *Old or Frail: What Tells Us More?:* Journal of Gerontology 59A.9, 962–965.
- 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.
- Servizio di Epidemiologia di Milano (2015). *Piano di Interventi per la Prevenzione e il Contenimento degli Effetti sulla Salute Umana delle Onde di Calore e degli Eccessi di Caldo Estivo.* Azienda Sanitaria Locale Di Milano. Ultimo accesso: 07/08/2015. URL: http://www.asl.milano.it/user/download.aspx?FILE=OBJ02011.PDF&TIPO=FLE&NOME=piano_anticaldo_2010.
- SHARE. Ultimo accesso: 04/09/2015. URL: <http://www.share-project.org/>.
- Sternberg, S.A., Bentur, N., Abrams, C., Spalter, T., Karpati, T. *et al.* (2012). *Identifying Frail Older People Using Predictive Modeling.* The American Journal of Managed Care 18.10, 392–397.
- The Cardiovascular Health Study.* Ultimo accesso: 04/09/2015. National Heart, Lung & Blood Institute. URL: <https://chs-nhlbi.org/>.
- The Johns Hopkins ACG System.* Ultimo accesso: 07/08/2015. URL: <http://acg.jhsph.org/>.
- The Johns Hopkins ACG System (2011). *Technical Reference Guide, Version 10.0.* Johns Hopkins Bloomberg School of Public Health. Ultimo accesso: 07/08/2015. URL: https://www.healthpartners.com/ucm/groups/public/@hp/@public/documents/documents/cntrb_035024.pdf.
- Zielinski, A., Kronogård, M., Lenhoff, H. & Halling, A. (2009). *Validation of ACG Case-mix for equitable resource allocation in Swedish primary health care.* BMC Public Health 9.347.