



Università degli Studi di Milano – Bicocca

Dipartimento di Sociologia

Scuola di Dottorato in Scienze Sociali

Dottorato in Sociologia Applicata e Metodologia della Ricerca Sociale

XXVIII Ciclo

**IL RUOLO DELLE DISUGUAGLIANZE SOCIO-ECONOMICHE
NELLA SOPRAVVIVENZA AL CANCRO PER I TUMORI DELLA
MAMMELLA E DEL COLON-RETTO**

Dottorando: Roberto Lillini

Relatore: Prof. Mario Lucchini
Co-relatore: Prof. Emanuela Sala

Anno Accademico 2014/2015

| <i>Indice</i> | <i>Pag.</i> |
|---|-------------|
| 1. <u>Introduzione: motivazioni, ipotesi e rilevanza dello studio</u> | 6 |
| 1.1 Perché considerare il cancro | 6 |
| 1.2 Gli obiettivi e l'ipotesi alla base dello studio e gli elementi innovativi rilevanti | 8 |
| 1.3 Il disegno di ricerca | 13 |
| 1.4 Articolazione della tesi | 13 |
| | |
| 2. <u>Il quadro teorico e la rassegna bibliografica</u> | 16 |
| 2.1 Gli approcci teorici | 16 |
| 2.2 Salute, malattia e fattori sociali ed economici | 22 |
| 2.2.1 Il supporto sociale | 24 |
| 2.2.2 Tipologia di occupazione e organizzazione del lavoro | 27 |
| 2.2.3 Disoccupazione | 29 |
| 2.2.4 Stili di vita legati alla salute | 30 |
| 2.2.5 La zona di residenza | 30 |
| 2.2.6 Coesione di gruppo e capitale sociale | 32 |
| 2.2.7 Meccanismi biologici influenzati dai fattori socio-economici: alcuni esempi | 33 |
| 2.3 Il ruolo dei fattori socio-economici nella malattia cancro | 34 |
| 2.4 Fattori socio-economici e livello geografico | 40 |
| 2.5 Fattori epidemiologici, socio-economici e sopravvivenza dei pazienti oncologici | 46 |
| 2.5.1 Estensione della malattia alla diagnosi | 49 |
| 2.5.2 Biologia del tumore | 51 |
| 2.5.3 Ruolo delle comorbidità | 51 |
| 2.5.4 Trattamento terapeutico | 52 |
| 2.5.5 Fattori demografici: età e sesso | 54 |
| 2.5.6 L'istruzione | 56 |
| 2.5.7 Il supporto sociale nella sopravvivenza al cancro | 57 |
| 2.6 Cancro, identità e effetti sulla qualità della vita | 59 |

| | |
|--|-----|
| 3. <u>Il disegno della ricerca</u> | 62 |
| 3.1 Gli obiettivi della ricerca | 62 |
| 3.2 Le sedi tumorali e i soggetti inclusi nello studio | 64 |
| 3.3 L'area geografica | 70 |
| | |
| 4. <u>La costruzione del dataset</u> | 76 |
| 4.1 I Registri Tumori e il Registro Tumori umbro | 77 |
| 4.1.1 Funzionamento dei Registri Tumori | 80 |
| 4.1.2 I dati e le fonti dei dati dei Registri Tumori | 84 |
| 4.1.3 La qualità del dato | 89 |
| 4.1.4 Il Registro Tumori umbro | 91 |
| 4.1.5 I dati del Registro Tumori umbro | 93 |
| 4.2 Il Censimento Nazionale della Popolazione e delle Abitazioni | 99 |
| 4.2.1 I dati umbri del Censimento della Popolazione e delle Abitazioni | 103 |
| 4.3 Le Anagrafi comunali | 106 |
| 4.3.1 I dati delle Anagrafi comunali umbre | 108 |
| 4.4 L'unione dei dataset (<i>data linkage</i>) | 112 |
| 4.4.1 Definizione di <i>record linkage</i> | 112 |
| 4.4.2 La costruzione dei <i>dataset</i> finali | 118 |
| | |
| 5. <u>I metodi di analisi</u> | 122 |
| 5.1 La sopravvivenza relativa | 123 |
| 5.1.1 Il modello di base della sopravvivenza relativa | 125 |
| 5.2 Il modello di sopravvivenza relativa a parametri flessibili | 128 |
| 5.2.1 Il modello multilivello a parametri flessibili ed effetti misti per il calcolo della sopravvivenza relativa | 131 |
| 5.3 Il modello analizzato | 132 |
| 5.3.1 Interventi tecnici sulle variabili utilizzate nel modello | 137 |
| 5.4 Le analisi post-modello | 138 |
| | |
| 6. <u>I risultati</u> | 140 |

| | |
|--|-----|
| 6.1 Il tumore alla mammella | 140 |
| 6.1.1 Le caratteristiche di base | 142 |
| 6.1.2 I risultati del modello di sopravvivenza a parametri flessibili ed effetti misti | 148 |
| 6.1.3 L'analisi degli effetti dello stato civile sulla diagnosi e sulle cure | 153 |
| 6.2 Il tumore al colon | 157 |
| 6.2.1 Le caratteristiche di base | 157 |
| 6.2.2 I risultati del modello di sopravvivenza a parametri flessibili ed effetti misti | 164 |
| 6.2.3 L'analisi degli effetti dello stato civile e del titolo di studio sulla diagnosi e sulle cure | 168 |
| 6.3 Il tumore al retto | 174 |
| 6.3.1 Le caratteristiche di base | 174 |
| 6.3.2 I risultati del modello di sopravvivenza a parametri flessibili ed effetti misti | 181 |
| 6.3.3 L'analisi degli effetti dello stato civile sulla diagnosi e sulle cure | 185 |
| 6.4 Prime considerazioni sui risultati | 189 |
| | |
| <u>7. Discussione dei risultati e conclusioni</u> | 194 |
| 7.1 Considerazioni sulla scelta dei modelli di sopravvivenza | 194 |
| 7.2 Considerazioni specifiche sui modelli di sopravvivenza relative alle singole sedi | 197 |
| 7.2.1 Il modello di sopravvivenza al tumore della mammella | 204 |
| 7.2.2 Il modello di sopravvivenza al tumore del colon-retto | 209 |
| 7.3 Conclusioni | 213 |
| | |
| Bibliografia | 218 |
| | |
| Appendice | 252 |
| | |
| Ringraziamenti | 268 |

Capitolo 1

Introduzione: motivazioni, ipotesi e rilevanza dello studio

Questa tesi si propone di esaminare alcuni aspetti dell'associazione fra la malattia cancro e le caratteristiche socio-economiche degli individui malati e dell'area in cui risiedono, concentrandosi in particolare sui rapporti esistenti fra tali caratteristiche e la sopravvivenza al tumore.

Lavorando su tale tema, verranno esaminati anche due aspetti collaterali di tipo metodologico e strumentale particolarmente rilevanti, ad esso collegati.

Più precisamente, si cercherà di capire se esista o meno una differenza nel considerare le caratteristiche socio-economiche individuali rispetto a quelle dell'area di residenza del paziente, ossia se esista il cosiddetto *bias* ecologico su cui più volte la letteratura in materia si è interrogata (vedi, ad esempio, Woods et al., 2005).

In più, si valuterà se (e in che condizioni) sia possibile eseguire il *record linkage* fra dati amministrativi provenienti da differenti enti territoriali.

Questo tipo di studio, come vedremo nel capitolo 2, si inserisce in un più ampio quadro teorico e di ricerca empirica, che si è sviluppato in particolare a partire dagli anni '70 in risposta ad un preciso bisogno cognitivo: osservare quali fossero i fattori, non solo di tipo biologico-clinico, in grado di descrivere e spiegare le disuguaglianze di salute fra individui, in modo da fornire elementi utili ai decisori pubblici per intervenire in maniera più efficiente con politiche di sanità pubblica opportunamente mirate.

1.1 Perché considerare il cancro

Nello studio qui descritto ci si è concentrati sulla patologia cancro per una serie di caratteristiche che la rendono particolarmente rilevante sia per i sistemi sanitari basati sull'intervento pubblico, sia per quelli fondati sull'assicurazione sanitaria privata.

Il cancro è, infatti, una malattia cronica la cui storia clinica ha una forte relazione con i fattori socio-economici, in particolare per quanto riguarda l'accesso alle cure sanitarie e la possibilità di una diagnosi tempestiva, ossia i fattori che influenzano in maniera importante la prognosi e i risultati terapeutici (Hansen et al., 2008; Kogevinas et al.,

1997).

Inoltre, il cancro è una malattia la cui diffusione è tipicamente legata al crescere dell'età. Essa rappresenta il principale fattore di rischio di morte per la popolazione sopra i 65 anni, con incidenza e mortalità che crescono esponenzialmente dopo i cinquant'anni.

In Europa, già nel 2000 più del 65% dei decessi per cancro si registrava nella popolazione anziana sopra i 65 anni di età e, attualmente, tale valore è salito al 68% (dati al 2012; WHO, 2015).

Con il progressivo aumento dell'aspettativa di vita e l'incremento dell'invecchiamento della popolazione (in particolare nei paesi sviluppati) si sta, quindi, facendo sempre più pesante il carico di malattia oncologica che i sistemi sanitari e le politiche sociali devono e dovranno affrontare.

Di conseguenza, capire come ottimizzare i processi di diagnosi e cura e quali fattori individuali di tipo biologico, clinico, demografico e socio-economico condizionano l'efficacia di tali politiche è diventato un obiettivo di grande rilevanza per migliorare le opportunità di sopravvivenza dei malati (cfr., ad esempio, Klassen et al., 2011; Whitaker et al., 2015).

Occorre anche considerare che il già consistente impatto, esercitato dalle patologie neoplastiche su spesa sanitaria e sociale, è destinato ad aumentare oltre che per l'ingente numero di pazienti interessati, anche per i particolari costi della ricerca e delle procedure cliniche tecnologicamente sempre più avanzate, che possono rivelarsi particolarmente pesanti da sostenere in periodi di crisi economiche globali, nonché ad avere un ruolo sempre maggiore nel condizionare la situazione socio-economica delle famiglie in cui un membro incorre in una diagnosi di tumore (Martin-Moreno et al., 2012; Timmons et al., 2013).

Per dare un'idea della dimensione di tali costi, si può ricordare come la stima del costo annuale dei nuovi casi di tumore nel 2009 in Italia fosse pari a oltre 8,3 miliardi di euro, con un costo complessivo per paziente pari a circa 26 mila euro, con un impatto che può essere considerato pari, in valore, allo 0,6% circa del prodotto interno lordo annuale (Osservatorio sulla Condizione Assistenziale dei Malati Oncologici, 2010).

La tabella 1.1 riporta invece il confronto, sempre per il 2009, dei costi complessivi e per paziente ad un anno dalla diagnosi fra diversi paesi europei e l'Italia.

Tab. 1.1 - Costi delle patologie tumorali. Confronto Italia-Paesi europei (valori assoluti in euro).

| Paese | Costo complessivo 2009 | Costo unitario 2009 |
|-------------|------------------------|---------------------|
| Belgio | 1.963.902.754 | 35.388 |
| Francia | 9.929.017.054 | 33.329 |
| Germania | 14.751.886.786 | 32.303 |
| Italia | 8.337.852.450 | 25.815 |
| Regno Unito | 6.348.245.019 | 21.321 |
| Spagna | 2.874.051.975 | 15.453 |
| Eu27 | 52.959.144.243 | 22.696 |

Fonte: Osservatorio sulla Condizione Assistenziale dei Malati Oncologici, 2010.

Se si considera, poi, il miglioramento generale della prognosi dei pazienti oncologici, è necessario mettere in conto che il numero sempre maggiore di persone che vivono dopo una diagnosi di tumore è già alto e continuerà a crescere (Allemani et al., 2015). Di conseguenza, la maggiore prevalenza di malati verrà così a determinare un ulteriore aggravio della spesa sanitaria e sociale (Guy et al., 2013), non solo in termini di classico *follow-up* clinico ma anche di supporto socio-psicologico. Pertanto appare chiaro come l'insieme di cancro, andamenti demografici e allocazione delle risorse sarà uno dei principali problemi da fronteggiare in salute pubblica nel futuro prossimo.

Anche il mondo della ricerca scientifica si è accorto che gli sviluppi futuri dovranno andare anche in direzioni differenti da quelle usuali, considerando quale importanza possano avere variabili non strettamente cliniche nell'influenzare gli esiti sanitari. Vi è ormai evidenza di come le risorse psicologiche, il supporto sociale e la condizione socio-economica possano influenzare la storia naturale della malattia, a partire dalle esposizioni ai fattori di rischio e dal suo sviluppo sino al trattamento ed al *follow-up* (ad esempio, Bouchardy et al., 2006).

1.2 Gli obiettivi e l'ipotesi alla base dello studio e gli elementi innovativi rilevanti

La ricerca qui descritta si è concentrata su un aspetto particolare della malattia cancro, ossia sulla sopravvivenza che, come vedremo nel capitolo 2, è particolarmente legata alla tempestività e alla qualità delle azioni di diagnosi (ossia in quale stadio e con quale dettaglio viene individuata la malattia) e cura (ossia come il paziente viene preso in carico dal sistema sanitario), nonché alle caratteristiche demografiche e socio-economiche dei pazienti, le quali possono influenzare sia la rapidità con cui si arriva

alla diagnosi, sia la possibilità di accedere e completare correttamente un percorso di cura. Questo, ovviamente, se sono disponibili terapie efficaci per trattare adeguatamente il tipo di tumore da cui il paziente è affetto.

L'obiettivo principale di questa tesi consiste, quindi, nell'esplorare l'esistenza di una relazione fra le caratteristiche socio-economiche dei pazienti e della sezione di Censimento di residenza, le caratteristiche della diagnosi e della cura, e la sopravvivenza alla malattia neoplastica.

Come sarà spiegato in maniera più ampia successivamente, due obiettivi secondari discendono da questo obiettivo primario.

Il primo obiettivo secondario si propone di esplorare se le caratteristiche socio-economiche individuali esprimono aspetti della relazione con la sopravvivenza differenti da quelli espressi dalle caratteristiche socio-economiche della sezione di Censimento.

Il secondo obiettivo secondario, invece, consiste nel valutare la fattibilità di azioni di *record linkage* fra banche dati provenienti da enti sanitari, come i Registri Tumori, e altre provenienti da enti amministrativi territoriali di differente natura, come le Anagrafi comunali o l'Ufficio Statistico regionale, che raccolgono informazioni sulla stessa popolazione.

L'ipotesi alla base di questo studio è che esista una relazione fra le condizioni socio-economiche dei pazienti, delle famiglie e dell'area in cui risiedono e gli esiti della cura (ossia la probabilità di sopravvivenza). Esse possono condizionare il decorso della malattia in molti modi, influenzando ad esempio il momento in cui viene eseguita la diagnosi (precoce o tardiva), l'accesso alle cure e quanto accuratamente vengono seguiti i protocolli terapeutici.

Tale ipotesi, come vedremo in dettaglio nel prossimo capitolo, è ben nota in letteratura ed è stata esplorata a partire dalla fine degli anni '70/inizio degli anni '80.

Possiamo, ad esempio, ricordare lo studio di Dayall et al. (1982) che fu fra i primi a studiare, su 903 pazienti statunitensi malate di tumore alla mammella, le relazioni esistenti fra gruppo etnico di appartenenza, condizione socio-economica e sopravvivenza. I risultati mostrarono come le differenze in sopravvivenza osservate fossero dipendenti dalla combinazione dei fattori socio-economici e dall'appartenenza di gruppo, concentrandosi su come la differente distribuzione delle risorse sembrasse

essere la motivazione principale di tali disuguaglianze. Nel lavoro di Dayall et al. non venne preso in considerazione l'aspetto genetico, che però studi successivi hanno ampiamente dimostrato essere rilevante per l'incidenza e non particolarmente rilevante (o rilevante in maniera discutibile) per la sopravvivenza. In merito si possono ricordare, ad esempio, gli studi importanti di Braun et al. (1995) e Chen et al. (2016) che, studiando il comportamento di diversi tumori in coppie di gemelli, non hanno trovato relazioni significative con la sopravvivenza.

Lo studio di Dayall, benché importante come uno dei punti di partenza per studiare l'associazione fra stato socio-economico ed esiti sanitari, presentava un limite importante, dato dal non considerare gli effetti di interazioni fra variabili cliniche, biologiche e socio-economiche e la sopravvivenza, che in altri studi hanno confinato il ruolo delle variabili socio-economiche a semplici confondenti.

Successivamente molti altri studi¹ in tutto il mondo hanno affrontato tale tematica, utilizzando molteplici metodologie e scelte empiriche di ricerca differenti, ma perseguendo sempre l'obiettivo di esplorare l'associazione fra caratteristiche socio-economiche dei pazienti, aspetti patologici della malattia e caratteristiche della presa in carico dei pazienti da parte dei sistemi sanitari. Solo per citare alcune pubblicazioni di riferimento, possiamo ricordare lo studio di De Cuyper et al. (2003), che analizzò il ruolo combinato di stadio alla diagnosi, caratteristiche del tumore, trattamenti applicati, caratteristiche socio-economiche ed effetti di tutti questi fattori sull'esito della malattia in un gruppo di pazienti affette da tumore della mammella. I risultati mostrarono come, pur prendendo in considerazione tutti gli effetti delle variabili biologiche e cliniche, nonché gli effetti dell'età, gli aspetti socio-economici delle pazienti mantenevano un ruolo statisticamente significativo nell'influenzare lo stadio della malattia al momento della diagnosi (e, quindi, le probabilità di sopravvivenza) e la tipologia di trattamento applicato.

Oppure può essere citato il lavoro di Woods et al. (2006), prima rassegna completa della letteratura di riferimento sul tema, che dimostrò come i soli fattori medici tradizionali non potevano spiegare completamente le differenze osservate in termini di sopravvivenza negli studi considerati, sottolineando gli effetti statisticamente significative delle condizioni socio-economiche dei pazienti e delle aree in cui erano

¹ Cfr. il capitolo 2 di questa tesi per un quadro esteso della letteratura in materia.

residenti.

Quali sono, quindi, gli aspetti innovativi che caratterizzano la presente ricerca, rispetto a quanto già presentato dai lavori esistenti?

Sulla scia di alcuni studi esteri sulla mortalità (ad esempio, Bentley et al., 2008) e sulla sopravvivenza per tumore (ad esempio, Akinyemiju et al., 2013), un primo, importante, elemento innovativo è la possibilità di valutare, su una popolazione ampia, l'esistenza di eventuali differenze fra l'effetto delle caratteristiche socio-economiche individuali e quelle relative alla sezione di Censimento di residenza dei pazienti.

Questo approccio è innovativo e, almeno per l'Italia, esplorato qui per la prima volta nello studio della relazione fra malattia oncologia e condizione socio-economica del malato.

Fino ad oggi, infatti, gli studi italiani di popolazione sulle relazioni fra sopravvivenza ai tumori e caratteristiche socio-economiche dei pazienti si sono prevalentemente svolti attribuendo ai singoli pazienti le caratteristiche dell'area di residenza in cui essi vivono, utilizzando l'area geografica più piccola per cui è possibile ottenere dati ufficiali².

Malgrado molti lavori abbiano dimostrato come questo approccio porti a risultati attendibili (cfr., ad esempio, Cadum et al., 1999), esiste sempre il problema di una possibile distorsione legata alla differenza fra le caratteristiche dell'individuo e quelle della zona in cui vive, che si ipotizzano essere simili (Woods et al., 2005).

Un secondo elemento innovativo è di essere uno studio che, primo nel suo genere in Italia, valuta la fattibilità di applicare con successo procedure di *record linkage* a dati provenienti da diverse fonti amministrative territoriali, che abitualmente non utilizzano in maniera abbinata le proprie informazioni di tipo sanitario, demografico e socio-economico.

Questo lavoro può, quindi, mostrare come possa essere utile e ricco, in termini di ricerca in questo settore, mirare a combinare tali dati, nonché fornire utili indicazioni su come applicare le procedure di *linkage* qui utilizzate pure in altre situazioni di ricerca omologhe, che abitualmente non ne fanno uso.

Un ultimo elemento di innovazione è dato dal metodo e dalla tecnica utilizzata per trattare i dati in esame.

² Per l'Italia, ad esempio, è la sezione di Censimento, le cui dimensioni si aggirano attorno ai 170 abitanti in media.

Fino ad ora, infatti, il rapporto fra sopravvivenza e covariate (sia di tipo medico che di tipo socio-economico) è stato studiato prevalentemente tramite il modello di sopravvivenza di Cox.

Esso permette l'analisi degli effetti fissi delle covariate e in questo settore ha prodotto risultati importanti (ad esempio, Gallo et al., 2012).

Non permette però di lavorare utilizzando il modello della sopravvivenza relativa, particolarmente adeguato quando si utilizzano dati provenienti da un Registro Tumori e si considera un tempo di osservazione piuttosto lungo, e di analizzare gli effetti casuali in una prospettiva di analisi multilivello³ (aspetti che saranno spiegati più dettagliatamente nel capitolo 5 di questa tesi). Non è quindi un modello che possa essere utilizzato in maniera proficua per uno studio come quello qui descritto.

In alternativa, al fine di implementare il modello della sopravvivenza relativa nell'esplorazione degli effetti delle covariate di varia natura, ci si è rivolti ai modelli di sopravvivenza relativa a parametri flessibili (Lambert et al., 2005 e 2009).

Come vedremo più avanti in questa tesi, rispetto al modello di Cox essi permettono una migliore modellizzazione della sopravvivenza nel tempo, in particolare quando si usano dati su ampie popolazioni osservate per periodi di tempo medio-lunghi (superiori ai tre anni). Permettono anche di analizzare più accuratamente gli effetti fissi delle covariate, ma non consentono di considerare gli effetti casuali multilivello.

In questo studio si è quindi deciso di utilizzare un modello innovativo di analisi della sopravvivenza relativa a parametri flessibili, con considerazione degli effetti misti, ossia valutando sia gli effetti fissi delle covariate sia gli effetti casuali multilivello, legati alla sezione di Censimento di residenza (Crowther et al., 2014), ed è una delle prime volte in assoluto che si applica tale metodo di analisi su una simile combinazione di dati e variabili.

Tale modello è stato implementato in un nuovo modulo di Stata, STMIXED (Stata Corp., 2014), ed è stato sviluppato appositamente per questo genere di applicazioni fra il 2014 e il 2015. Il lavoro di analisi in questa tesi è uno dei primi studi che ha potuto testarlo e utilizzarlo in una versione stabile.

³ Solo recentemente, una estensione del modello di Cox ha permesso di valutare gli effetti multilivello, ma anche in questo caso non è implementabile il modello della sopravvivenza relativa (Hastert et al., 2015).

1.3 Il disegno di ricerca

Per quanto riguarda la definizione del disegno di ricerca, si è scelto di considerare i tumori della mammella nelle donne e del colon-retto in entrambi i sessi perché, come vedremo più dettagliatamente nei capitoli 2 e 3, sono due tumori dall'incidenza particolarmente alta, in particolare nella popolazione italiana. Sono caratterizzati dall'aver protocolli di prevenzione, diagnosi e cura particolarmente efficaci, che permettono di ottenere una buona sopravvivenza (AIRTUM Working Group, 2014). In più, sono due tumori per i quali, come è ampiamente riportato in letteratura, sembra esistere un forte legame fra sopravvivenza e caratteristiche socio-economiche dei pazienti.

La scelta della regione Umbria come campo dello studio è stata, invece, determinata da due ragioni:

- una ragione di tipo metodologico, ossia le caratteristiche demografiche e socio-economiche della popolazione umbra sono sufficientemente rappresentative di quelle della popolazione italiana (Vercelli et al., 2013), così come la distribuzione dei due tumori considerati;
- una ragione operativa, ossia la buona integrazione fra le fonti informative necessarie per la raccolta dati, nonché la qualità elevata dei dati stessi.

Le analisi tramite la tecnica sopra illustrata sono state sviluppate presso il Dipartimento di Medicina e Biostatistica del Karolinska Institutet di Stoccolma, in collaborazione con i ricercatori che hanno sviluppato la teoria e le *routine* informatiche relative ai modelli di sopravvivenza relativa a parametri flessibili ed effetti misti (Crowther et al., 2014; Dickman et al., 2015; Lambert et al., 2009).

1.4 Articolazione della tesi

La presente tesi si articola in sette capitoli. Accanto a questo primo capitolo introduttivo, che presenta in sintesi la rilevanza, le motivazioni, gli aspetti innovativi e il disegno di ricerca, il secondo capitolo presenta un ampio quadro teorico e una rassegna della letteratura in cui è inquadrato lo studio, presentando diversi presupposti teorici e modelli che hanno cercato di definire il ruolo delle caratteristiche socio-economiche

nella salute degli individui, in generale, e nella sopravvivenza ai tumori, in particolare. I capitoli 3 e 4 riportano la definizione degli obiettivi e del disegno di ricerca, la descrizione dei dati analizzati, delle fonti utilizzate e della modalità della raccolta dati. In particolare, nel capitolo 4 sono discussi gli aspetti peculiari delle procedure di *record linkage* utilizzate per la costruzione dei *dataset* utilizzati per le analisi. Il capitolo 5 approfondisce le ragioni alla base della scelta del metodo e delle tecniche di analisi, descrivendoli dettagliatamente e riflettendo sugli aspetti innovativi che li caratterizzano, mentre il capitolo 6 riporta i risultati relativi all'analisi dei dati. Chiude la tesi il capitolo 7, in cui sono discussi i risultati ottenuti e i loro limiti; in più, nelle conclusioni, si propongono alcuni suggerimenti sia per potenziali sviluppi futuri di questo lavoro, sia per un'applicazione pratica dei risultati ottenuti.

Capitolo 2

Il quadro teorico e la rassegna bibliografica

Lo studio descritto in questa tesi si inquadra nel filone di ricerca che si occupa di analizzare i rapporti fra caratteristiche socio-economiche e condizioni di salute della popolazione.

Più precisamente, questo lavoro fa parte degli studi che prendono in considerazione se e come le caratteristiche socio-economiche di un individuo e quelle dell'area in cui è solito vivere influenzano le condizioni di salute, con particolare attenzione agli effetti sugli esiti di una determinata patologia (nel nostro caso, la sopravvivenza a due tipi di tumore).

L'obiettivo di questo capitolo è, quindi, definire il contesto teorico in cui è inquadrato questo studio, a partire da una riflessione sugli elementi fondativi che regolano il rapporto fra la salute, la malattia e gli aspetti sociali ed economici, per passare successivamente a descrivere come questi elementi sono stati considerati nello studio dei tumori e dei loro pazienti.

In questa prospettiva, particolare attenzione sarà rivolta a come la letteratura in materia si è espressa sul ruolo che le caratteristiche socio-economiche possono avere nell'influenzare l'andamento delle malattie, in generale, e della patologia oncologica, in particolare, ponendo particolare attenzione a come esse possano essere connesse alla sopravvivenza ai due tumori considerati in questo studio.

2.1 Gli approcci teorici

Le teorie sulla relazione fra fattori socio-economici e condizioni di salute vedono un grande sviluppo fra la fine degli anni '70 e l'inizio degli anni '80, stimolati da alcuni interventi di Wilkinson che, in uno scambio epistolare con l'allora Segretario di Stato inglese, evidenziava come le differenze di classe sociale fossero associate a differenze nelle aspettative di vita e nei rischi di molteplici cause di morte. Da queste osservazioni prese le mosse lo studio conosciuto come "Black Report" (Department of Health and Social Security UK, 1980), che verificò la persistenza di importanti disuguaglianze in

termini di esiti sanitari legate alla condizione socio-economica degli individui, malgrado i notevoli miglioramenti nella copertura sanitaria della popolazione garantita dalla Sanità pubblica.

Questo risultato fu particolarmente rilevante, perché permise di evidenziare come anche un sistema sanitario pubblico non fosse in grado di prevenire tutti i possibili fattori sociali accompagnati agli eventi di malattia. In particolare, Wilkinson et al. hanno notato, in un recente lavoro (2010), come solo grazie al Black Report (e agli studi successivi) fossero emerse le ipotesi che stanno alla base della relazione fra stili di vita, comportamenti a rischio, condizioni socio-economiche e supporto sociale.

A partire da queste osservazioni, Dahlgren e Whitehead (1991) hanno poi provato a fornire un quadro concettuale in grado di legare la condizione di salute con il sistema sociale ed economico esistente in un determinato contesto, prendendo in considerazione gli elementi tradizionali del livello di istruzione, condizione e ambiente professionale, stato civile, caratteristiche del sistema sanitario, ecc. I due autori attribuiscono una forte rilevanza ai fattori relativi alle condizioni socio-economiche, sanitarie e ambientali del contesto sociale generale in cui è inserito l'individuo, perché essi ne condizionano stili di vita, scelte e condizioni materiali. Questi ultimi tre elementi agiscono, secondo gli autori, sulla salute attraverso meccanismi psicologici (come lo stress, l'ansia, ecc.) o concreti (presenza/assenza di supporto familiare e/o sociale).

Successivamente, Wilkinson e Marmot (2003) hanno identificato ulteriori elementi specifici che possono avere un impatto negativo sulla salute, quali la disoccupazione, l'esclusione sociale, le diverse possibilità di accesso a risorse materiali e immateriali. Secondo i due autori, anche questi fattori agiscono sulla salute secondo meccanismi psicologici (basati sulla pressione emotiva che esercitano sugli individui) e materiali (minori risorse da mettere in gioco per supportare la propria salute).

La riflessione sui legami fra disuguaglianze sociali e disuguaglianze in salute è, ovviamente, molto più estesa di quanto sopra riportato e, in sintesi, ha generato una serie di tentativi di spiegazione che si sono poi sviluppati secondo approcci teorici anche profondamente diversi fra di loro.

In sintesi, possiamo individuare due teorie principali:

- le teorie della selezione;
- le teorie della causazione.

1) Le teorie della selezione. Le teorie della selezione si sviluppano a partire dalla prospettiva del darwinismo sociale, che afferma come le differenze genetiche individuali condizionino l'aspetto fisico, psichico ed emotivo. In quest'ottica, le teorie della selezione ritengono che le differenze genetiche siano alla base delle differenze sociali e, di conseguenza, siano le differenze di salute a determinare le differenze sociali.

In realtà, queste teorie sembrano applicarsi male al contesto umano, mostrandosi in grado di spiegare solo le differenze di salute che originano da condizioni alla nascita, come la presenza di disabilità congenite. In questo quadro teorico, però, altre problematiche nelle condizioni di salute dell'individuo, che potrebbero essere ricondotte alle esperienze di vita delle persone, ai comportamenti a rischio, alla storia familiare, vengono comunque analizzate alla luce dell'ipotesi di un'eredità genetica familiare.

L'ipotesi dell'eredità genetica familiare, ad esempio, viene usata per spiegare gli abusi di sostanze. Gli studi in merito, infatti, hanno mostrato come la componente ereditaria possa arrivare a spiegare fino al 70% della varianza dell'utilizzo di stupefacenti (Goldman et al., 2005). L'eredità familiare predispone a una condizione di disagio, che può portare a scegliere di consumare una sostanza con cui, comunque, occorre essere "a contatto" (fattore ambientale); si riconosce quindi che la componente genetica interagisce con quella ambientale (che comprende anche gli aspetti socio-economici e culturali), che però è vista come tendenzialmente immutabile e, quindi, con un effetto meno rilevante (ad esempio, Heat et al., 2002).

Va però salvata l'idea che una condizione fisica di svantaggio influenzi tutte le altre dimensioni della vita e possa portare più facilmente ad un basso status socio-economico dell'individuo.

Se le teorie della selezione hanno il merito di evidenziare gli effetti di uno svantaggio fisico alla nascita o acquisito, in realtà, come discusso da Marmot et al. (2006), la risposta sociale alla malattia è una costruzione sociale e ha degli effetti sullo stato di salute della persona. Questo porta, inevitabilmente, a dover prendere in considerazione fattori che rientrano nel campo di studio della sociologia (Ardisson et al., 2013).

2) Le teorie della causazione. Le teorie della causazione, invece, partono dal

presupposto che la relazione tra disuguaglianze sociali e di salute possa, almeno in parte, essere spiegata dalle scelte comportamentali effettuate dall'individuo. Questo approccio, che tende a dare molta importanza alla responsabilità individuale (Knowles, 1977), trae origine dagli studi sull'epidemiologia dei fattori di rischio, che aveva evidenziato come molte malattie fossero associate a particolari condotte di vita (come, ad esempio, analizzato nel lavoro di Mc Ginnis et al., 1993).

In questa prospettiva, infatti, ogni persona cercherà di utilizzare il capitale umano e sociale che ha a disposizione per prendere decisioni razionali, mirate a massimizzare i propri vantaggi. Il livello di istruzione, ad esempio, rappresenta un certo investimento di tempo e denaro a cui possono corrispondere benefici equivalenti, che sono quindi il frutto di un investimento razionale e libero, frutto della lungimiranza dell'individuo.

In quest'ottica, però, nel campo della sociologia della salute si tende ad attribuire responsabilità rilevanti al malato per l'instaurarsi di uno stato di malattia. Di conseguenza, diventa rilevante favorire azioni di politiche pubbliche (di prevenzione primaria e secondaria) mirate a stimolare ed educare a condotte di vita più salutari, che riducano la presenza di fattori di rischio (Simnet et al., 1999).

Altre teorie legate alla prospettiva della causazione spostano l'attenzione dalle scelte individuali ai contesti (familiari, culturali, politici, storici) nei quali le azioni sono compiute: il contesto definisce e limita le scelte socialmente accettate effettivamente a disposizione degli individui (Krieger et al., 1993).

Secondo questa prospettiva, i comportamenti legati alla salute vanno interpretati come effetti cumulativi delle condizioni imposte dalla struttura sociale, che provoca determinate risposte comportamentali da parte delle persone (Lynch et al., 1997).

I comportamenti di salute degli individui vengono, infatti, associati alla classe sociale dei genitori e affondano le radici nei primi anni di vita; il legame tra alcuni comportamenti e l'appartenenza a una classe socio-economica nell'età giovanile può, quindi, essere spiegato dal fatto che proprio in quegli anni si sono consolidati comportamenti che sono rimasti stabili anche nell'età adulta (Lynch et al., 1997).

La teoria del capitale culturale permette di spiegare l'interazione tra disuguaglianza socio-economica e disuguaglianza di salute; il capitale culturale è, infatti, il prodotto di pratiche culturali, accumulo di conoscenze, interazioni sociali proprie della classe di appartenenza, che formano una disposizione ad agire secondo certe modalità stabili nel

tempo (definita “*habitus*” da Bourdieu) e che comprendono anche, ovviamente, comportamenti che possono influire positivamente o negativamente sulla salute, sulla relazione medico-paziente, sulla capacità/volontà di recepire i messaggi della prevenzione, di prestare attenzione ai segnali del proprio corpo per diagnosi più precoci o di seguire eventuali percorsi terapeutici.

In estrema sintesi, l’*habitus* è definito da Bourdieu come "un sistema di schemi percettivi, di pensiero e di azione acquisiti in maniera duratura e generati da condizioni oggettive, ma che tendono a persistere anche dopo il mutamento di queste condizioni" (Bourdieu, 1983).

L’*habitus*, in quanto sistema di schemi, genera azioni coerenti relativamente al gruppo o classe di appartenenza. L’*habitus* non è dunque né universale, né specifico di un individuo, ma è l’elemento centrale della riproduzione sociale e culturale in quanto è capace di generare comportamenti regolari e attesi, che condizionano la vita sociale degli individui in relazione alla loro classe di appartenenza.

In più, la capacità delle pratiche culturali di modellare i comportamenti individuali è riconosciuta anche dalla teoria delle cause fondamentali (Lieberson, 1985; Phelan et al., 2010), che osserva come le persone con *status* sociale più elevato abbiano sempre a disposizione risorse fondamentali superiori a chi è in posizione sociale più svantaggiata. Queste risorse si concretizzano in denaro, conoscenza, potere, prestigio e rete sociale e interagiscono fortemente fra di loro. Esse influenzano le condizioni di salute di un individuo, poiché sono attive sia a livello del singolo che della collettività in cui è inserito e agiscono in tutti i momenti del percorso che porta dallo stato di buona salute, alla malattia, alla (eventuale) guarigione.

A livello collettivo, i benefici prodotti dalle risorse derivano da quanto i gruppi sociali mettono in gioco o hanno a disposizione per proteggere la salute dei membri del gruppo (Cockerham, 2005 e 2013). Phelan et al. (2010) affermano che gli stili di vita sono, poi, sottoposti a pressioni esterne legate sia alle norme sociali, sia agli interessi di mercato, che possono influenzare i comportamenti anche in maniera contraria alla prevenzione. I gruppi che avranno uno *status* fortemente condiviso e socialmente coeso subiranno meno questo tipo di pressione e potranno mantenere comportamenti più “virtuosi”, protettivi della condizione di buona salute.

L’approccio teorico delle cause fondamentali vede quindi la struttura sociale prevalere

sulla libera scelta dell'individuo, dando minore importanza alle responsabilità individuali e, come la teoria del capitale culturale, pone nella stratificazione sociale l'origine delle disuguaglianze nella salute degli individui.

Anche la teoria della disuguaglianza di Wilkinson (2005) appartiene alla prospettiva teorica della causazione e considera la differenza di reddito come l'origine della disuguaglianza in termini di durata e qualità della vita dei singoli, affermando che tale differenza esercita i suoi effetti non solo concretamente, ma anche a livello psicologico. Oltre agli aspetti pratici dell'avere a disposizione meno risorse, infatti, egli osserva come anche la percezione della differenza contribuisca a ridurre ulteriormente il benessere degli individui svantaggiati, con effetti diretti ed indiretti sulla salute (meno cura di sé stessi).

La disuguaglianza in sé sembra quindi essere l'elemento di stress che opera attraverso un meccanismo di confronto con gli altri e che produce uno stato d'ansia legato alla prestazione sociale (come siamo visti o giudicati). Questo stato di stress è causato dai fatti che includono una minaccia al proprio *status*, mettendo in discussione il proprio sé sociale. Un frequente e prolungato stato di stress ha un effetto importante sul nostro organismo, in particolare sul sistema immunitario (che regola anche gli stati infiammatori legati all'aumento di rischio dell'insorgenza di patologie neoplastiche) e su quello cardio-circolatorio (Dickerson et al., 2004). Questo genere di stress è più frequente e più prolungato se si scende nella scala dello *status* sociale, perché più frequenti saranno le occasioni portatrici di stress e più numerosi i momenti in cui sarà necessario difendere la propria identità sociale (Wilkinson, 2005).

Questa situazione di stress da disuguaglianza sociale, sempre secondo i risultati degli studi di Wilkinson (2005 e 2010), provoca anche altri problemi di salute direttamente misurabili, come problemi di ovulazione nelle donne e disfunzione erettile negli uomini, danneggiamenti di alcune aree del cervello e della relativa funzione cognitiva.

Sempre in questo filone di studio, gli studi di Marmot et al. (1978) hanno trovato come posizioni lavorative con scarso livello di controllo sulle proprie decisioni e azioni causino alti livelli di stress, con conseguenze dirette e indirette sulle condizioni di salute. Questo aspetto e una più ampia trattazione dei fattori socio-economici che influenzano la salute verrà esaminato nel prossimo paragrafo.

2.2 Salute, malattia e fattori sociali ed economici

L'approccio teorico sopra descritto, legato alla teoria della causazione, non si preoccupa tanto di definire le cause della malattia, piuttosto cerca di identificare i fattori che permettono il mantenimento in buona salute degli individui (Gove, 1979; Morgan, 1980; Najman, 1980) e, focalizzandosi sull'aspetto collettivo e, quindi, sulle popolazioni, costituisce la base per le strategie pubbliche di promozione della salute. Come determinanti della salute vengono identificati cinque grandi fattori: l'ambiente sociale ed economico (famiglia e lavoro), l'ambiente fisico, le pratiche di salute personali, la capacità individuale di gestire il proprio stato di salute e i servizi sanitari a cui si può accedere (Labonte, 1993; Locker, 2008). Altri fattori che, successivamente, sono stati citati come determinanti della salute comprendono l'esclusione sociale e la discriminazione, le abitudini alimentari e i trasporti (Wilkinson et al., 1998).

La deprivazione materiale e una mancanza di controllo sulle dimensioni importanti della vita sono, quindi, i temi principali che, in questa prospettiva, devono essere affrontati nella promozione della salute della popolazione (non è un caso che i paesi, in cui il divario di reddito tra ricchi e poveri è ridotto, presentano un tasso di mortalità generale più basso). Di conseguenza, il cambiamento sociale e politico, tra cui la redistribuzione del reddito, può essere necessario per modificare lo stato di salute di gruppi di popolazione di bassa condizione socio-economica.

Sarebbe tuttavia un errore ritenere che il ruolo dei fattori sociali e psicologici come cause di malattia sia stato preso in considerazione solo in tempi moderni. Ad esempio, nel contesto della storia della medicina moderna, l'idea che la malattia possa essere causata da influenze psicologiche è stata parte integrante del lavoro di Freud, che ha spiegato disturbi quali asma e ulcere gastriche come il prodotto di conflitti psicologici irrisolti. L'opera di Freud ha dato origine al concetto che alcune malattie sono condizioni psicosomatiche, mentre altre non lo sono.

Il punto di vista contemporaneo completa il quadro, affermando che anche i fattori sociali, oltre ai fattori psicologici, sono implicati in tutte le malattie, anche se i meccanismi con cui essi influenzano la salute sono complessi e variabili (Maturò, 2008).

Ad esempio, anche se i comportamenti come il fumo sono di tipo individuale, un certo numero di fattori sociali e culturali influenza la possibilità di diventare un fumatore e di

continuare a fumare per un periodo di tempo più o meno prolungato. Questi fattori includono elementi culturali associati al fumo quali l'età adulta, l'attrattiva sessuale e l'emancipazione, la struttura socioeconomica e comunicazionale associata all'industria del tabacco, la legislazione che ne regola i consumi, nonché l'influenza di coetanei o altre persone socialmente significative per l'individuo (Syme, 1986; Glanz et al., 2005). Si ritiene pertanto importante descrivere di seguito i principali fattori sociali e psicologici che influiscono sulla salute e come si è arrivati a verificare che assumono un ruolo fra le cause di malattia in generale.

La maggior parte dei primi studi su fattori sociali e insorgenza di malattia riguardava gli effetti del cambiamento sociale e culturale (ad esempio, industrializzazione e urbanizzazione, migrazione, mobilità sociale, occupazionale e geografica) (Brenner, 1979; Brown et al., 1978; Gove, 1979).

L'esito di malattia studiato inizialmente era quello relativo alle patologie cardiovascolari, poiché erano ritenute malattie tipiche di nazioni industrializzate e urbanizzate, come risultava da studi eseguiti su popolazioni isolate a bassa pressione sanguigna (in apparenza non influenzata dall'età) e i cui tassi di malattie cardiache aumentavano quando queste popolazioni si muovevano verso contesti urbani (Friedman, 1967; Syme et al., 1964).

Un certo numero di studi condotti nel corso degli anni '60 e dei primi anni '70 aveva infatti trovato tassi più alti di malattia tra le persone che avevano cambiato lavoro, luogo di residenza o tipo di vita. Ad esempio, uno studio particolarmente rilevante aveva trovato che individui nati e cresciuti in aziende agricole che si erano trasferiti nei centri urbani, in cui avevano iniziato lavori di tipo operaio qualificato o impiegatizio (comunque riconducibili alla classe media o medio-bassa), avevano più alti tassi di malattia coronarica rispetto a individui paragonabili che avevano continuato a lavorare in un'azienda agricola (Syme et al., 1964). Osservazioni simili sono state fatte in relazione al cancro (a partire da Randal, 1976).

Numerosi fattori potrebbero essere responsabili degli effetti negativi sulla salute, prodotti dai cambiamenti sociali e culturali individuali. Gli effetti avversi possono infatti essere il risultato diretto del cambiamento stesso, oppure un prodotto delle circostanze in cui gli individui si muovono. Uno studio (che negli anni '70 ha tentato di

valutare queste spiegazioni) aveva confrontato i tassi di incidenza di malattie cardiache tra i giapponesi immigrati in California e nelle Hawaii con quelli di uomini giapponesi che ancora vivevano in Giappone (Marmot et al, 1975). Malattie cardiache e mortalità coronarica erano più frequenti tra i residenti in California e meno frequenti in Giappone. Tra quelli che vivevano in California, alcuni avevano adottato cultura e stili di vita occidentali, mentre altri mantenevano uno stile di vita più simile a quello tradizionale giapponese. Il primo gruppo presentava tassi di malattia fino a cinque volte superiori rispetto al secondo, suggerendo come il fattore importante che spiegava il rischio di malattia fosse il cambiamento nell'ambiente in cui queste persone erano residenti.

2.2.1 Il supporto sociale

Uno dei primi studi sul rapporto tra ambiente sociale e sanitario è stato, indubbiamente, lo studio sul suicidio di Durkheim (1897). In questo lavoro, in estrema sintesi, Durkheim ha sperimentato l'uso di metodi statistici per descrivere e spiegare le differenze dei tassi di suicidio tra i diversi gruppi sociali. Durkheim ha spiegato il suicidio in termini di organizzazione sociale di questi gruppi, in particolare nella misura in cui gli individui sono stati integrati nel gruppo, e il modo in cui questa integrazione sia elemento di incoraggiamento o dissuasione dal suicidio. Alti tassi di suicidio sono stati associati con gruppi che avevano livelli molto alti e molto bassi di integrazione sociale, classificando le ragioni del suicidio proprio in base al rapporto fra individuo e società⁴.

Studi più recenti su legami sociali e salute si sono concentrati sul rapporto tra il sostegno sociale e il benessere. Alcuni di questi primi lavori hanno guardato alle differenze in materia di salute derivanti dallo stato civile (Gove, 1979; Morgan, 1980). *Single*, vedovi e divorziati hanno mostrato tassi di mortalità più elevati degli sposati, con differenze molto più marcate per gli uomini che per le donne.

Queste differenze, che erano già state osservate e riportate a partire dalla metà del XIX secolo, si sono mantenute piuttosto stabili nel tempo e attraverso le culture e i sistemi

⁴ Più precisamente si distinguono:

- il suicidio egoistico, causato da una carenza di integrazione sociale;
- il suicidio altruistico, causato da un'integrazione troppo profonda nel tessuto sociale, al punto da suicidarsi per soddisfare l'imperativo sociale;
- il suicidio anomico, tipico delle società moderne, causato dallo scollamento profondo fra valori individuali e della società.

sanitari, con solo una piccola parte delle differenze nei tassi di mortalità che può essere attribuita agli effetti selettivi del matrimonio (Morgan, 1980).

Una possibile spiegazione di queste differenze è che lo stato civile ha un'influenza sugli stati psicologici e sugli stili di vita (Gove, 1979). Gli sposati tendono ad essere più felici e più soddisfatti, sono meno propensi a essere socialmente isolati e hanno legami più solidi. In una società in cui matrimonio e famiglia sono valori centrali della vita, essere sposati dà senso e significato alla vita quotidiana stessa ed è una fonte di sostegno sociale ed emotivo. Questa spiegazione è supportata, negli studi sopracitati, dalle relazioni fra lo stato civile e le cause specifiche di decesso.

Le variazioni nei tassi di mortalità sono, ad esempio, maggiori quando gli stati psicologici o gli stili di vita svolgono un ruolo diretto nella morte, come nel suicidio o nella morte per incidenti, o sono associati a comportamento a rischio come il fumo o il consumo di alcol (Patterson et al., 2004).

Si osservano anche notevoli differenze rispetto a malattie come la tubercolosi o i tumori, in cui i fattori familiari possono influenzare la diagnosi precoce e la volontà/capacità di sottoporsi a cure con la disponibilità di aiuto e supporto (Kroenke et al., 2006; Pinquart et al., 2007).

Uno studio particolarmente rilevante, che ha chiaramente dimostrato come l'integrazione nella comunità abbia un effetto positivo sulla salute, è stato sviluppato da Berkman et al. (1979). Gli autori hanno seguito un campione casuale di adulti in un periodo di nove anni. All'inizio dello studio è stato calcolato, un punteggio per valutare la rete sociale e di supporto su cui poteva contare ogni individuo, sulla base di stato civile, contatti con gli amici e parenti e appartenenza a gruppi sociali e religiosi. Nel corso dei nove anni, le persone con punteggi bassi hanno fatto registrare probabilità di morire più alte rispetto a quelle con punteggio elevato. Dopo il controllo per altri fattori, come il peso, il fumo di sigaretta, l'alcool, il consumo, l'attività fisica, le pratiche di salute e lo stato di salute iniziale, il tasso di mortalità per le persone socialmente isolate è stato di due o tre volte superiore a quello di persone con una rete sociale estesa.

Questi primi risultati sono stati confermati in altri studi longitudinali. Ad esempio, uno studio durato sei anni su uomini in Finlandia ha trovato che il rischio di morte era più alto tra coloro che avevano dichiarato di avere poche persone a cui dare o da cui ricevere sostegno (Kaplan et al., 1994). La mancanza di partecipazione sociale, l'aver

pochi amici e il non essere sposati sono stati associati con la mortalità dopo aver tenuto conto dello stato di salute iniziale, di altri fattori di rischio e di fattori economici, tra cui il reddito.

Altre evidenze empiriche hanno confermato che l'integrazione sociale e il sostegno sociale hanno una grande influenza sulla salute. Studi degli anni '90, ad esempio, hanno scoperto che l'integrazione sociale esercita un effetto protettivo sull'incidenza di infarto miocardico non fatale in uomini di età da 50 anni e oltre (Vogt et al., 1992), mentre l'isolamento sociale è risultato legato ad un aumento del rischio di ictus in uno studio su professionisti sanitari di sesso maschile negli Stati Uniti (Kawachi et al., 1996). Nel corso degli anni, sono anche emerse prove consistenti che dimostrano che la mancanza di sostegno sociale, l'isolamento sociale e il non essere sposati influiscono negativamente sulla sopravvivenza a lungo termine tra gli uomini a seguito di un iniziale infarto del miocardio (Williams et al., 1992). Studi più recenti hanno rilevato che il sostegno sociale migliora le probabilità di sopravvivenza nei pazienti sottoposti a dialisi (Thong et al., 2007), nei pazienti con acuta leucemia mieloide (Pinquart et al., 2007) e tra le donne dopo la diagnosi di cancro al seno (Kroenke et al., 2006). In quest'ultimo studio, le donne risultate socialmente isolate alla diagnosi hanno mostrato un aumento del rischio del 66% per tutte le cause di morte e un aumento del 200% della mortalità per cancro al seno, rispetto alle donne socialmente integrate.

Altri studi hanno trovato che il supporto sociale era legato a livelli di disabilità residua a seguito di un ictus ed è stato correlato all'insorgenza di depressione e grado di disabilità tra i pazienti con artrite reumatoide (Fitzpatrick et al., 1991).

Uno studio che ha seguito per due anni individui con disabilità che vivono da soli ha mostrato che le persone con pochi contatti sociali erano più propense a deteriorarsi nel fisico e nel funzionamento psicosociale, rispetto alle persone con livelli elevati di contatto con gli altri (Patrick et al., 1986).

Considerando che per supporto sociale si intende una piuttosto ampia categoria di elementi, che include assistenza pratica, aiuto finanziario, fornitura di informazioni e consulenza e sostegno psicologico, uno studio di dieci anni su donne anziane in Finlandia ha esaminato l'effetto di sostegno pratico e sostegno emotivo sulla mortalità e ha scoperto che il rischio di morte era di 2,5 volte superiore rispetto a quelle prive sostegno emotivo (Lyyra et al., 2006). In più, da uno studio svedese su grande scala è

emerso che avere diverse fonti di sostegno è risultato essere associato con migliori esiti sanitari a seguito di malattia; tali fonti di sostegno, in ordine di importanza, sono risultate essere il supporto economico, i contatti sociali e la possibilità di discutere di problemi personali (Ostberg et al., 2007).

I molteplici meccanismi attraverso i quali il sostegno sociale ed emotivo migliora o protegge la salute sono però ancora in discussione.

Un'ipotesi riguardo a questo meccanismo è emersa nel contesto di studi relativi all'impatto negativo sulla salute in circostanze di vita stressanti (ad esempio, Antonovsky, 1987; Brunner et al., 2006). Tali atti di supporto sembrano essere un cuscinetto contro gli eventi avversi che altrimenti avrebbero effetti dannosi sulla salute. Nella loro ricerca sulle origini sociali della depressione, Brown et al. (1978) hanno trovato che il supporto sociale era protettivo solo nel contesto di un evento della propria storia personale particolarmente grave. In particolare, il sostegno sociale ha dimostrato di essere associato ad una minore angoscia psicologica e maggiore capacità di gestire l'evento in questione, quando esso era l'insorgenza di una patologia grave come il cancro (Baider et al., 2003).

Infine, sebbene gli studi suggeriscono che il sostegno sociale migliora la capacità di tollerare circostanze stressanti, è anche possibile che esso eserciti un'influenza attraverso percorsi ormonali e neurofisiologici.

2.2.2 Tipologia di occupazione e organizzazione del lavoro

Gli ambienti fisici in cui le persone lavorano sono spesso pericolosi e dannosi per la salute. L'inquinamento atmosferico sul posto di lavoro, l'esposizione ad agenti cancerogeni, il lavorare con macchinari e gli incidenti industriali riscuotono un grande tributo sulla salute dei lavoratori manuali. Tuttavia, anche l'aspetto psicosociale dell'ambiente in cui lavoriamo può avere un impatto negativo sulla salute.

Due modelli che identificano importanti aspetti psicosociali dell'ambiente di lavoro sono il modello di domanda-controllo e il modello di squilibrio del rapporto sforzo-ricompensa (Marmot et al., 1999). Il primo suggerisce che i posti di lavoro che combinano un elevato livello di richieste psicologiche con bassi livelli di controllo e di abilità portano a stress e aumento di rischio di malattie, come la malattia coronarica. Il secondo suggerisce che posti di lavoro che combinano un alto grado di stress con bassi

livelli di guadagno economico o ricompense emotive, quali la sicurezza del posto di lavoro o l'avanzamento di carriera, portano a stress emotivo, stress lavorativo e probabile insorgenza di malattie.

Le prove a sostegno del modello di domanda-controllo vengono da uno studio britannico condotto su impiegati pubblici nella metà degli anni '80, che esplorava i legami tra l'organizzazione del loro lavoro e lo stato di salute (Marmot et al., 1988). Nessuno dei soggetti considerati nello studio viveva in condizioni di povertà, tuttavia erano evidenti differenze nello stato di salute corrispondenti al livello gerarchico professionale. Coloro che erano di livello più basso avevano tassi di mortalità tre volte superiori rispetto quelli di più alto grado, e sempre nei primi si rilevarono più alti tassi di insorgenza di malattie cardiache e più alti tassi di assenze per malattia. Queste differenze vennero direttamente correlate al grado di controllo sul posto di lavoro.

Questo studio ha concluso che la mancanza di libertà nel prendere decisioni sul posto di lavoro, in particolare quando i lavori sono stressanti o psicologicamente esigenti, incrementa sia l'assunzione di comportamenti a rischio, come il fumo, sia la presenza di fattori di rischio fisiologici, come pressione alta e malattie cardiache. Un *follow-up* dei partecipanti allo studio, realizzato un decennio più tardi, ha confermato questi risultati (Kuper et al., 2003); dati provenienti dallo stesso studio hanno indicato che il rischio di insorgenza di *angina pectoris* è stato predetto da ansia e disturbi del sonno (Nicholson et al., 2005).

Studi condotti in Svezia hanno trovato anche un più elevato rischio di infarto miocardico tra gli uomini impiegati in lavori impegnativi a basso controllo, con un'associazione particolarmente evidente tra i lavoratori manuali (Hallqvist, 1998).

Se si prende in considerazione il modello di squilibrio sforzo-ricompensa, diversi studi hanno indicato che esso è responsabile di stress cronico, associato ad un incremento da 2 a 6 volte del rischio di insorgenza di malattie cardiache (Bosma et al., 1998, Siegrist et al. 1990), non spiegabile da altri fattori di rischio biomedici o comportamentali. Lavori più recenti in materia, come lo studio longitudinale sui funzionari britannici, riportano un aumento nel tempo dell'insorgenza di *angina pectoris* negli uomini associabile a tale squilibrio fra sforzo e ricompensa, più comune in dipendenti pubblici di grado inferiore (Chandola et al., 2005). Tuttavia, questo rapporto non è stato osservato nelle donne.

2.2.3 Disoccupazione

Così come l'ambiente fisico e sociale in cui operiamo può avere un effetto negativo sulla salute, anche la disoccupazione può incidere sulle condizioni di salute.

Ci sono due ragioni principali per cui la disoccupazione, concettualmente, potrebbe incidere sulla salute (Marmot et al., 1987). In primo luogo, l'essere disoccupato condiziona pesantemente le condizioni materiali e, in secondo luogo, si tratta di un evento stressante che può diventare cronico e privare un individuo di un ruolo sociale, significativo dell'esistenza quotidiana e del contatto con gli altri.

Due approcci sono evidenti negli studi su disoccupazione e salute ed entrambi sono soggetti a problemi di interpretazione, soprattutto perché è difficile separare gli effetti della disoccupazione dagli effetti di altre condizioni sociali ed economiche (Marmot et al., 1987).

Il primo di questi approcci tenta di dimostrare un'associazione tra tasso di disoccupazione e tasso di mortalità e il modo in cui questi covariano con gli andamenti del ciclo economico. L'importante studio di Brenner (1979) evidenziava già trentasei anni fa come tali tassi seguissero andamenti sufficientemente concordi, particolarmente per le cause di morte improvvise legate allo stress o al disagio individuale.

Il secondo approccio cerca di valutare la salute di persone che sono, o sono diventate di recente, disoccupate. Dal momento che la cattiva salute può portare a disoccupazione e viceversa, tali studi devono essere realizzati con attenzione prima di concludere che la disoccupazione è causa di cattive condizioni di salute. Tuttavia, le ricerche in materia suggeriscono che i disoccupati sperimentano peggiori stati di salute, hanno una maggiore pressione del sangue e sperimentano mortalità più elevata degli occupati (Jin et al., 1996; Montgomery et al., 1999; Turner, 1995). La povertà, lo stress e l'insicurezza, soprattutto per l'alloggio, sono tutti fattori implicati nell'insorgenza dei problemi di salute in seguito a disoccupazione (vedi anche Nettleton et al., 1998).

Anche se una recente rassegna di 104 studi ha confermato che i disoccupati hanno peggiore salute fisica e psicologica rispetto ai loro omologhi occupati, essa ha anche evidenziato che l'entità degli effetti varia secondo la durata della disoccupazione e secondo caratteristiche individuali quali l'età, il sesso e le risorse con cui ogni persona può far fronte a tale situazione (McKee-Ryan et al., 2005). Gli effetti più marcati sono stati osservati tra le persone più giovani e di genere maschile (Artazcoz et al., 2004;

Reine et al., 2004).

Anche altre situazioni, come l'inizio del congedo di maternità o lo stare a casa per accudire i membri della famiglia, sembrano portare alla compromissione del benessere psicologico nel genere femminile (Thomas et al., 2005).

2.2.4 Stili di vita legati alla salute

Ci sono molti comportamenti che possono avere un impatto positivo o negativo sulla salute. Dieta, esercizio fisico, consumo di alcol, fumo, uso di droghe sono esempi che sono stati oggetto di numerose ricerche. Questi comportamenti sono spesso caratterizzati dall'essere il risultato di scelte individuali e della responsabilità personale, anche se è più interessante ai fini di questo studio vederli come il prodotto di circostanze sociali (Jarvis et al., 2006).

La prova che essi sono legati ai contesti sociali si trova nei dati che mostrano come i comportamenti in grado di favorire la salute sono meno comuni nei gruppi soggetti alla povertà e alla deprivazione sociale, mentre i comportamenti che implicano un aumento di probabilità di danni alla salute sono più comuni. Per esempio, la percentuale di fumatori tra uomini e donne è inversamente proporzionale alla classe sociale e all'istruzione (Jarvis et al., 2006).

Chiaramente, la deprivazione materiale e i fattori di stress sociale e personale possono spingere un individuo a fumare. Infatti, i più alti tassi di consumo di tabacco si trovano tra gruppi caratterizzati da combinazioni di questi fattori.

Ci sono anche prove che suggeriscono che smettere di fumare è meno frequente fra i gruppi deprivati, rispetto ai gruppi non deprivati, anche se sembra esserci poca differenza tra questi gruppi rispetto alle motivazioni (Laaksonen et al., 2005). In realtà, come mostra tutta la letteratura in materia, il discorso attorno al fumo in particolare, e ai comportamenti a rischio per la salute in generale, è molto più complesso e si osserva come le trasformazioni siano fortemente correlate anche alle azioni pubbliche di prevenzione primaria e secondaria (Hiscock et al., 2012), che in questa sede non approfondiremo ulteriormente.

2.2.5 La zona di residenza

Molto è stato studiato dei legami tra ambienti sociali e fisici e l'impatto che possono

avere sulla salute, anche con risultati che, a senso comune, possono sembrare scontati. E' però interessante notare come poveri che abitano in quartieri più ricchi hanno una migliore condizione di salute rispetto a persone di simile livello economico che vivono nei quartieri poveri (Blaxter, 1990). Ad esempio, i dati dello studio Whitehall II (Stafford et al., 2003) aveva rilevato che il 12,3% di impiegati di alto livello che vivevano nelle aree più svantaggiate era in cattive condizioni di salute, paragonabili a quelle del 35,9% di dipendenti di basso livello. Tra coloro che vivevano nelle zone meno svantaggiate, la percentuale di individui in cattive condizioni di salute è stata dell'8,8% per gli alti livelli e del 19,7% per i dipendenti di basso grado. Tali effetti di quartiere sono stati trovati per una serie di esiti sanitari, tra cui mortalità e incidenza di malattie cardiovascolari.

MacIntyre et al. (1993) hanno sostenuto che questi risultati mostrano come non solo le caratteristiche ambientali fisiche, ma anche sociali ed economiche dell'area in cui una persona risiede sono in grado di influenzare le condizioni di salute.

Essi hanno considerato due aree di Glasgow simili in termini di struttura dei redditi dei residenti e hanno verificato l'esistenza di tassi di mortalità relativamente bassi nell'area nord-ovest, dove erano disponibili alimenti più sani e più economici, più centri sportivi e ricreativi, servizi sanitari e di trasporto migliori, meno crimine, e tassi di mortalità relativamente alti nell'altra area a sud-ovest, dove queste caratteristiche erano meno presenti in maniera significativa. La conclusione che ne hanno tratto è che fra due persone con caratteristiche personali simili (ad esempio, stesso reddito, dimensione e composizione della famiglia e proprietà dell'alloggio) quella che vive nel nord-ovest sarebbe indiscutibilmente avvantaggiata rispetto al proprio stato di salute fisica e mentale. Uno studio più recente condotto in Scozia e Inghilterra ha identificato come ulteriori aspetti sociali e fisici dell'area di residenza fossero predittivi di scarsa salute (Cummins et al., 2005). Questi erano: scarsa qualità dell'ambiente fisico, più basso impegno politico dei residenti, elevata disoccupazione, basso accesso al trasporto privato, poca disponibilità di trasporto pubblico. Dopo interventi mirati a cambiare queste caratteristiche, come suggerito da tale studio, la salute della popolazione locale è migliorata.

Questi risultati sono particolarmente rilevanti, perché sembrano dare credito al ruolo rilevante che, in molti studi su deprivazione e patologie (in particolare tumori) viene

dato all'area di residenza del malato (quartiere, sezione di Censimento, comune, ecc.), come vedremo meglio successivamente, e come è stato studiato, ad esempio, in Italia da Biggeri et al., 2009; Cislighi, 2009; e all'estero, ad esempio, da DeJardin et al., 2008; Woods et al., 2006.

2.2.6 Coesione di gruppo e capitale sociale

Un altro fattore collettivo, che è associato con la salute, è quello della coesione sociale, intesa essenzialmente come l'estensione della connessione e della solidarietà tra i gruppi di una società (Kawachi et al., 2000); a livello individuale, è definibile come il personale senso di connessione ad una comunità (Patterson et al., 2004).

Comunità coese sono quelle in cui vi è un elevato livello di partecipazione alle attività comuni e alti livelli di appartenenza ai gruppi sociali della comunità stessa (Stansfield, 2006). Elementi importanti della coesione sociale sembrano essere il sostegno dei vicini, le opportunità di interazione con gli altri membri della comunità e la mancanza di paura della violenza e della criminalità.

Recenti evidenze hanno indicato che, sia a livello individuale sia a livello di area di residenza, la presenza e l'assenza di coesione sociale sono associate ai comportamenti a rischio per la salute, come il fumo (Patterson et al., 2004).

La coesione sociale è parte del cosiddetto capitale sociale, ossia di quelle forme associative formali o informali a cui l'individuo appartiene e può mettere in gioco con funzioni di supporto nei momenti critici della vita, in particolare per quanto riguarda le situazioni di malattia. Possono essere, ad esempio, gruppi religiosi, organizzazioni a varie finalità, società sportive associazioni sindacali (Kawachi et al., 1997).

Diversi studi hanno mostrato come il capitale sociale sia associato al tasso di mortalità generale di un'area e, in particolare, ai tassi di morte per cancro e malattie cardiache (Kawachi et al., 1997, (a) e (b); Kawachi et al., 1999; Wilkinson, 1996).

Di conseguenza, interventi che migliorano le caratteristiche di una comunità in modo da promuovere sentimenti di sicurezza e percezione di coesione di un territorio hanno portato a miglioramenti nel campo della salute dei residenti (Halpern, 1995).

La principale conclusione che si può trarre dalle ricerche sopracitate è la conferma che i modelli di salute e malattia sono anche il prodotto di influenze sociali e ambientali. Anche se la salute e la malattia coinvolgono specifici agenti e processi biologici, essi

sembrano inseparabili dai contesti sociali in cui le persone vivono, che condizionano le problematiche incontrate dalle persone nella vita quotidiana e la loro capacità di gestirle.

La modifica di questi ambienti è un modo in cui la salute di una comunità può essere migliorata. Tuttavia, la natura delle comunità e dei determinanti sociali della salute che incarnano sono modellati dalle politiche pubbliche e queste sono influenzate dalle forze sociali, economiche e politiche che influenzano i governi. Lacune nella conoscenza di come queste forze operano e la mancanza di intese professionali e pubbliche dei determinanti sociali di salute sono ostacoli al miglioramento degli ambienti in cui le persone vivono (Raphael, 2006).

2.2.7 Meccanismi biologici influenzati dai fattori socio-economici: alcuni esempi

Anche se le prove che collegano fattori socio-economici e ambientali con la salute sono stringenti, resta la questione sul modo in cui tali fattori operano per influenzare la salute e la malattia.

Benché ci siano ipotesi abbastanza differenziate in merito, le più accreditate affermano che i fattori di stress (che originano dai fattori descritti nei paragrafi precedenti) danno origine a situazioni psicologiche e cambiamenti fisiologici che aumentano la suscettibilità di un individuo alla malattia.

Ci sono prove che suggeriscono che lo stress, o il suo esito patologico nella forma della depressione, porta ad una serie di cambiamenti nel corpo umano. Esso interferisce con il normale funzionamento dei sistemi neuroendocrini, metabolici e immunitari, porta ad un aumento della frequenza cardiaca e della respirazione, alla dilatazione dei vasi sanguigni ai muscoli e ad alterazioni della funzione gastrointestinale. Questi cambiamenti sono ritenuti in grado di causare direttamente malattie o rendere un individuo più incline alla malattia (Brunner et al., 1999).

Ad esempio, lo studio sui dipendenti pubblici britannici, più volte citato in precedenza, ha verificato come lo stress indotto da determinati ambienti sociali sia correlato a processi biologici, come la sindrome metabolica, che aumentano il rischio di malattie cardiache e *ictus* attraverso fattori clinici che caratterizzano tali processi. Questi fattori includono obesità addominale, eccessi di grassi nel sangue, ipertensione, presenza di insulino-resistenza ed elevati livelli di proteina C-reattiva. Coloro che sono sottoposti a

stress cronico da lavoro sono risultati essere due volte più soggetti a tale sindrome rispetto a chi non è sottoposto a tale stress (Chandola et al., 2006). Le persone il cui lavoro è stato caratterizzato da uno squilibrio sforzo-ricompensa e sovra-impegno sono risultati avere livelli di cortisolo e pressione sistolica più alti, in particolare se erano lavoratori di basso livello (Steptoe et al., 2004).

Il legame tra i fattori di stress e la malattia è, quindi, mediato da una serie di elementi che possono aumentare o diminuire la vulnerabilità di un individuo di fronte ad una condizione di stress. Il supporto sociale e variabili psicologiche, come le caratteristiche della personalità, i processi di apprendimento e le strategie di *coping*⁵, interagiscono in modo complesso per influenzare gli esiti sanitari. Inoltre, come l'approccio socio-ambientale suggerisce, risposte comportamentali a fattori ambientali di stress in forma di attività dannose per la salute, come ad esempio il fumo, svolgono un ruolo importante (Najman, 1980).

2.3 Il ruolo dei fattori socio-economici nella malattia cancro

Tutte le precedenti considerazioni sui fattori socio-economici e il loro rapporto con lo stato di salute sono particolarmente rilevanti per comprendere molti aspetti della malattia cancro (Hiatt et al., 2008), in particolare quelli legati alla sopravvivenza.

Il cancro è un gruppo di malattie che impone un pesante fardello per la salute pubblica e, mentre la tendenza secolare all'aumento della mortalità per cancro nei paesi occidentali si è invertita nel 1995 (ad esempio, Howe et al., 2006; Vercelli et al., 2014), la patologia neoplastica rimane comunque la seconda causa di morte e i suoi costi economici per la società sono particolarmente gravosi (Brown et al., 2001). Sono stati, infatti, compiuti progressi impressionanti contro il cancro, non solo legati alle nuove conoscenze sulla genetica e la biologia molecolare o i corrispondenti nuovi approcci terapeutici, ma anche legati alla comprensione delle determinanti sociali ed economiche del cancro (McLeroy et al., 1988).

Tali determinanti sono innanzitutto rilevanti per i loro effetti indiretti attraverso fattori

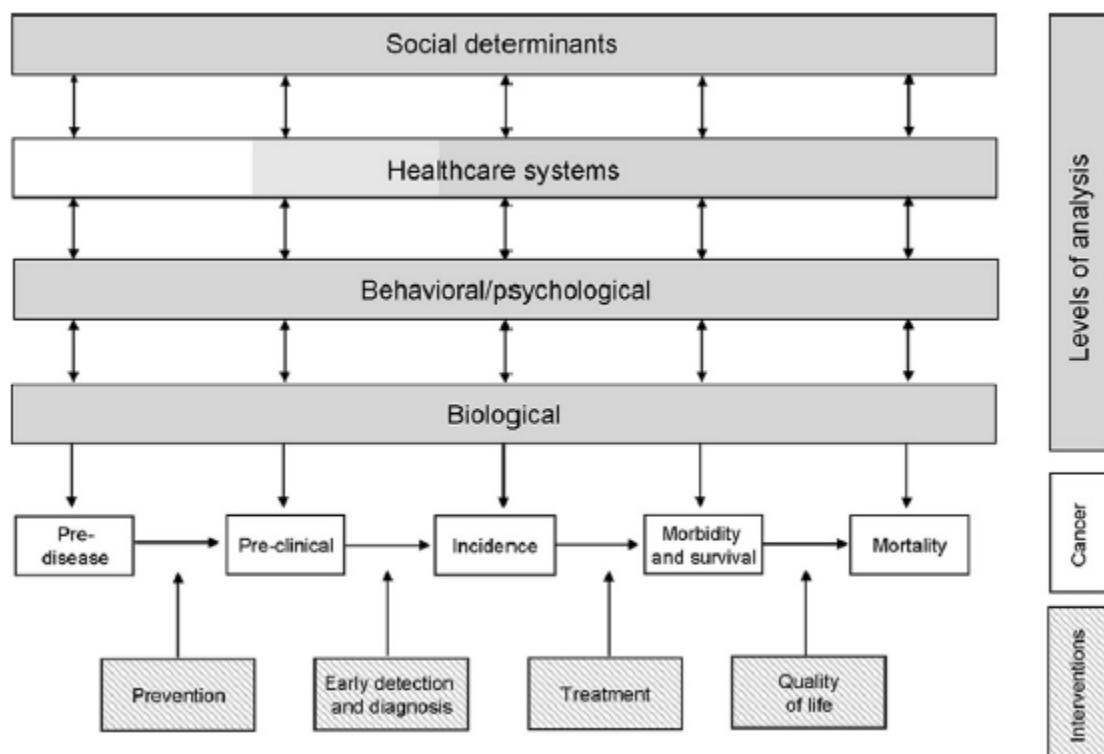
⁵ Per “strategie di *coping*” si intendono le modalità di adattamento con le quali si fronteggiano situazioni stressanti. A seconda dell'esito positivo o negativo il *coping* potrà essere definito funzionale (adattamento) o disfunzionale (aumento dello stress). Lazarus et al. (1991) definirono il *coping* come un insieme di sforzi cognitivi e comportamentali messi in atto per gestire le richieste esterne in relazione alle risorse possedute. Si tratta di un processo dinamico che porta a una continua rivalutazione dell'equilibrio tra ambiente e individuo.

di rischio individuali o comportamentali, come il fumare; oppure, perché interagiscono con fattori biologici e genetici (ad esempio, le interazioni gene-ambiente); o perché sono cause dirette e irriducibili della malattia indipendentemente dalle variabili clinico-terapeutiche considerate (Phelan et al., 2004; Smedley et al., 2001); o per tutte queste ragioni assieme.

Per questa ragione, secondo Hiatt et al. (2008), l'analisi del ruolo delle caratteristiche sociali ed economiche nello sviluppo della patologia tumorale è una problematica che va affrontata in termini di multifattorialità, secondo una prospettiva che potrebbe aiutare nella riportare e diffondere i risultati della ricerca sul cancro all'assistenza sanitaria e alle pratiche di sanità pubblica (Rosenfield et al., 1992; Smedley et al., 2001).

In Figura 2.1 è riportato il grafico concettuale (Hiatt et al., 2008) che cerca di esprimere l'integrazione fra i diversi fattori che condizionano sia l'insorgenza, sia gli effetti della cura (e, quindi, la sopravvivenza) della malattia oncologica.

Figura 2.1 – Determinanti sociali del cancro e livelli di interazione con gli altri fattori



Fonte: figura originale da Hiatt et al., 2008.

Lo schema illustra in sintesi come i determinanti sociali si riferiscono a livelli di analisi

e tipi di interventi differenti da quello sanitario, lungo il “continuum cancro” (Krieger, 2005 e 2008) e come i sistemi sanitari hanno meno probabilità di influenzare l'incidenza del cancro di quanto non possano fare per la sopravvivenza e la mortalità, intervenendo soprattutto nella fase di diagnosi e trattamento.

L'asse orizzontale illustra, in basso, il percorso della malattia, dalla situazione di assenza, attraverso la fase di recepimento dei messaggi di prevenzione, la diagnosi, la sopravvivenza, fino alla fine del ciclo di vita (National Cancer Institute, 2007).

Ogni fase è influenzata da diverse azioni che possono essere messe in atto, a partire dalla possibilità di prevenire la malattia stessa, all'utilizzo di pratiche di diagnosi precoce e alla diagnosi stessa, alla presa in carico per la cura della malattia, fino alle modalità con cui si cerca di recuperare una buona qualità della vita, fino alla fine della stessa.

In questo quadro è però fondamentale il concetto di livelli di analisi, perché cerca di chiarire come i diversi determinanti, fra cui i fattori sociali, influenzano il percorso di sviluppo della malattia sopra descritto, soprattutto la sopravvivenza, interagendo non solo con l'andamento della neoplasia ma anche fra di loro, in una logica che riporta a quanto proposto da Dahlgren e Whitehead (1991).

Come riportato in Figura 2.1, lo schema mostra quattro livelli di analisi sovrapposti allo sviluppo della malattia, che aggiungono il livello sociale ai livelli precedentemente proposti da Anderson et al. (2005):

- le condizioni sociali, da intendere nella più ampia accezione individuale e collettiva, comprendenti quindi sia le caratteristiche dell'individuo che della comunità/popolazione a cui appartiene (Kaplan, 2001);
- l'impatto dei sistemi sanitari;
- i fattori comportamentali e psicologici;
- i meccanismi biologici della carcinogenesi.

I primi tre livelli interagiscono fra di loro e influenzano gli aspetti biologici che stanno alla base dell'insorgenza, dello sviluppo e della conclusione della malattia, le cui opportunità di prevenzione, diagnosi e cura a disposizione del malato sono di conseguenza influenzate da tutti e quattro i livelli sovrastanti.

Gli interventi per ridurre le disparità socio-economiche, di trattamento, ecc. possono

quindi essere introdotti in uno qualsiasi di questi livelli.

Anche se tale schema rappresenta i rapporti fra le diverse componenti che interagiscono con la malattia come un processo lineare, la realtà è, ovviamente, molto più complessa (Vineis, 2006), perché le interazioni multidirezionali collegano influenze biologiche, cliniche e sociali in una più ampia rete di causazione (Krieger, 1994 e 2008).

Ad esempio, i fattori biologici possono influenzare il comportamento e generare la necessità di interventi di assistenza sanitaria. Inoltre, le politiche e la legislazione riguardanti la copertura per l'assistenza sanitaria possono intervenire sui comportamenti individuali e sull'uso dei servizi clinici.

In questo senso, i determinanti sociali sono caratterizzati come i fattori a monte (Link et al., 1998), distali sociali, ambientali, economici e culturali che determinano i comportamenti individuali e di gruppo (Kaplan, 1996; Marmot, 1998; McKinlay, 1979). Secondo Krieger (2005), a partire da questo schema i principali determinanti sociali di cancro risultano essere: provenienza etnica, posizione socio-economica e occupazione, genere, età, titolo di studio e competenza linguistica, copertura sanitaria assicurativa (per le nazioni a copertura privata o mista), stato civile, area di residenza.

Altri studi in materia hanno anche discusso di come le cause e gli eventi a monte influenzino la salute a livello di popolazione (Evans et al., 1994; Pasick et al., 2004; Tomatis, 2001) e, sebbene molte associazioni tra determinanti sociali e conseguenze sulla salute della popolazione a volte possano apparire ovvie, poche relazioni causali sono state rigorosamente verificate.

Capire come i determinati sociali descritti da Krieger sono distribuiti nella popolazione è fondamentale per comprendere le interazioni con il percorso della malattia oncologica. La distribuzione di tali determinanti può essere misurata a vari livelli (ad esempio, individuo, sezione di Censimento, comune, nazione) ed essi definiscono modelli non casuali di comportamenti, correlati agli esiti sanitari del cancro in gruppi di individui dalle caratteristiche simili. In questo modo, possono essere informativi dei meccanismi biologici fondamentali che influenzano (Masi et al., 2010).

A fronte di questi fattori, va poi considerata l'importanza dell'assistenza sanitaria che si pone come un ulteriore elemento chiave per la ricerca e il controllo del cancro (McKinlay, 1979). Un problema cruciale è, infatti, quello di distinguere chiaramente i risultati che sono dovuti a carenze nella fornitura di assistenza sanitaria da quelli esterni

ad essa, in modo che gli interventi possano essere opportunamente mirati.

Molti tumori hanno, infatti, una fase preclinica che inizia quando possono essere prevenuti (tramite la prevenzione primaria) e che si estende fino al rilevamento della malattia. Durante la fase preclinica, l'accesso alle cure sanitarie può influenzare la progressione dei tumori per i quali sono disponibili procedure di diagnosi precoce (ad esempio, mammella, cervice uterina, colon-retto, prostata), ma nel complesso è meno probabile che possa influenzare l'incidenza, laddove sembra invece avere effetti decisamente più rilevanti per la sopravvivenza, che trae vantaggio per molti tumori dalla diagnosi precoce (Phelan et al., 2004).

Questo punto mette in luce un altro problema fondamentale: il cancro non è una sola malattia, con caratteristiche univoche, ma è in realtà un insieme di molte malattie differenti con differenti eziologie. Di conseguenza, l'ambiente sociale può influenzare i diversi tipi di cancro in modi differenti.

Attualmente si considerano circa 80 tipi di neoplasia maligna; tuttavia quattro siti rappresentano circa la metà di tutta l'incidenza del cancro: nel 2012 seno (25%), prostata (15%), polmone (13%) e colon (9%) rappresentavano il 62% di tutte le nuove diagnosi di cancro stimate (WHO, 2015).

La concentrazione di incidenza del cancro tra questi quattro siti ha offerto e offre l'opportunità per studi specifici sulle determinanti sociali. Ad esempio, la mortalità per cancro al polmone è fortemente associata all'uso del tabacco e delle politiche di prevenzione primaria attuate per ridurre la pratica del fumo, principale fattore di rischio. La mortalità per tumore della mammella e del colon è modellata dalla distribuzione dello *screening* nella popolazione. Anche se l'incidenza del cancro al seno è più comune nelle donne di alta condizione socio-economica, la mortalità è più elevata tra le donne di condizione inferiore (Berry et al., 2005; Singh et al., 2003).

Occorre qui precisare un aspetto metodologico importante di questi studi: la maggioranza degli studi sul rapporto tra disuguaglianze e neoplasie si basa sui dati raccolti dai Registri Tumori (descritti dettagliatamente nel capitolo 4 di questa tesi), che permettono di utilizzare dati anche molto dettagliati di tipo sanitario e demografico, ma che sono in genere privi di informazioni sulle caratteristiche socio-economiche

individuali sui pazienti⁶.

Per ovviare a tale problema e procedere con lo studio dei determinanti socio-economici, si è quindi in genere scelto di utilizzare i dati da Censimento delle aree di residenza in cui vivono i malati presenti nei *database*, come *proxy* dei dati individuali (Singh et al., 2003; Barry et al., 2005).

Queste scelte di studio hanno permesso di valutare gli effetti dei differenti gradienti socio-economici d'area sugli esiti delle neoplasie (Adler et al., 1994; Marmot, 1991; Singh et al., 2001; Smith et al., 1996) e hanno anche permesso ai ricercatori di esaminare l'effetto dei fattori socio-economici sugli individui nei contesti in cui vivono e lavorano, sia per quanto riguarda la sopravvivenza ai tumori, sia per la mortalità legata ad altre malattie (Barry et al., 2005; Diez-Roux, 1998 e 2000; Krieger, 1993).

A livello di popolazione, gli studi hanno mostrato che se gli effetti dei gradienti socioeconomici sulla sopravvivenza (o sull'incidenza) ai tumori si mantengono statisticamente significativi dopo l'aggiustamento per i fattori di rischio noti (ad esempio, l'uso del tabacco e altri comportamenti a rischio) e per lo *screening*, vi è sostegno empirico per il valore della ricerca sui percorsi biologici condizionati da specifiche condizioni socio-economiche.

In quest'ottica è particolarmente rilevante il già citato studio di Phelan et al. (2004) che ha confrontato i risultati sanitari di neoplasie per le quali esistono interventi preventivi e terapeutici efficaci, con i risultati di patologie tumorali per i quali non lo sono, trovando forti effetti delle differenze socio-economiche per il primo tipo di tumori.

Gli autori hanno concluso che la causa fondamentale di fondo ha a che fare con l'insieme di risorse accessibili da persone con alto livello socio-economico, anche se questa spiegazione continua ad essere dibattuta. Tuttavia, l'ampia gamma di determinanti sociali che hanno effetto sulla malattia oncologica sottolinea la necessità di studi interdisciplinari atti ad analizzare i ruoli che la biologia, i singoli comportamenti e le determinanti sociali giocano nella formazione di tumori in sedi specifiche (Hiatt, 2006).

Infatti, le disparità osservate hanno motivato lo studio delle influenze a livello sociale sulla eziologia dei tumori. Lo sviluppo dell'epidemiologia sociale ha aperto la strada per

⁶ Fa in parte eccezione la rete dei Registri Tumori statunitensi (SEER), che negli anni più recenti ha iniziato la raccolta di informazioni anche sullo stato civile, l'etnia, il titolo di studio e altre informazioni di tipo socio-economico.

l'analisi multilivello applicata al controllo del cancro.

Tradizionalmente, la salute della popolazione e i fattori sociali sono stati al centro del lavoro di epidemiologi, sociologi, economisti, antropologi, politologi. L'assistenza sanitaria è stata competenza dei ricercatori dei servizi sanitari, economisti, studiosi del comportamento e della comunicazione e medici. I comportamenti umani individuali e i fattori di rischio che mediano la salute sono stati studiati principalmente dagli psicologi. Le discipline scientifiche di base della genetica, biologia cellulare, immunologia e biochimica hanno chiarito percorsi e meccanismi biologici.

L'obiettivo del recente approccio interdisciplinare è quindi trovare un approccio di ricerca collaborativo che permetta di mettere in gioco contemporaneamente tutte queste aree.

L'intensità delle disuguaglianze è, infatti, conseguenza di fattori individuali e contestuali: la stratificazione sociale, dipendente dal contesto economico e lavorativo, influenza le traiettorie di vita personali, dall'infanzia fino alle condizioni di vita e lavoro in età adulta, influenzando inoltre le capacità di valorizzare le proprie opportunità di salute e risorse. Le esposizioni a fattori di rischio e le malattie determinano condizioni di suscettibilità e fragilità, amplificate dalla mancanza di supporto e aggravate da limitazioni nell'accesso ai servizi di prevenzione e cura

2.4 Fattori socio-economici e livello geografico

Come accennato nel paragrafo precedente, nel programmare uno studio sulla relazione tra fattori socio-economici e tumori è molto importante identificare bene sia l'aspetto sociale, sia l'unità statistica (livello individuale o ecologico di micro- o macro-area geografica), sia l'esito sanitario che si vuole analizzare, nonché operare adeguate e specifiche scelte per ciascuno di questi tre punti critici. Numerosi autori suggeriscono di usare più di un indicatore per descrivere i molteplici aspetti sociali della popolazione; frequentemente questi indicatori sono correlati, ma in ogni caso ciascuno di essi ha un valore predittivo indipendente rispetto alla variabile sanitaria investigata (Liberatos et al., 1988).

Si è visto come esista, in letteratura, indicazione di una grande varietà di fattori, indicatori e indici ed ognuno di essi misura differenti aspetti di differenti stati socio-economici. Come accennato, l'intensità delle disuguaglianze varia in base agli indicatori

sociali utilizzati, come conseguenza di diversi tempi e meccanismi di generazione. Occorrono quindi validi principi teorici nell'operare la scelta dell'indicatore, perché indicatori basati su principi erronei e male interpretati possono condurre a false conclusioni. D'altra parte, una teoria corretta nel descrivere lo stato socio-economico non solo porterà ad una migliore comprensione sociale, ma anche ad una più precisa misurazione di come la salute e le malattie siano distribuite all'interno della società.

Un chiaro esempio di come la scelta degli indicatori di status socio-economico possa influenzare i risultati delle indagini epidemiologiche è offerto da due studi sull'associazione tra fattori socio-economici e sopravvivenza di pazienti oncologici in Canada e negli Stati Uniti, che mostrano risultati apparentemente in conflitto fra loro (Boyd et al., 1999; Gorey et al., 2000). Un'analisi successiva ha dimostrato che i differenti risultati nei due paesi erano semplicemente dovuti ad una diversa scelta degli indicatori socio-economici utilizzati per descrivere la popolazione (Zhang et al., 2006).

Alcuni autori hanno quindi investigato quanto il contesto di popolazione e di area geografica, nonché gli indicatori socio-economici scelti, influissero sulle differenze di sopravvivenza riscontrate evidenziando come la dimensione dell'area sembrasse essere più importante della stessa scelta degli indicatori (Woods et al., 2005). Più precisamente, studi come quello di Woods et al. hanno verificato che attribuire agli individui le caratteristiche socio-economiche della loro area di residenza era lecito solo se quest'ultima era di dimensioni sufficientemente piccole (tendenzialmente attorno ai 2.000 abitanti al massimo), al fine di evitare problemi di distorsione fra il livello individuale e quello aggregato d'area.

Altri autori, a loro volta, hanno trovato differenti corrispondenze con la sopravvivenza tumorale a seconda delle variabili *proxy* dello stato socio-economico individuale considerate (Robert et al., 2004). Questo ultimo studio è particolarmente importante per il lavoro svolto in questa tesi, ancorché Robert et al. abbiano preso in considerazione l'incidenza e non la sopravvivenza.

Gli autori, infatti, per verificare se le differenze socio-economiche legate ad un maggiore rischio di sviluppare un tumore si esprimessero a livello individuale o di area di residenza, hanno considerato una coorte di donne ammalate di tumore al seno e un gruppo di donne sane, in un'area del Wisconsin dove era possibile differenziare fra insediamenti urbani e rurali, e raccolto dati sul loro stato civile e sul livello di

istruzione.

I risultati finali dello studio hanno mostrato come, dopo aver controllato per tutte le variabili di tipo clinico e biologico, i fattori socio-economici individuali e di area mantenevano entrambi un'associazione statisticamente significativa con i rischi di sviluppare tale tumore. In più, le variabili individuali e di area sembravano descrivere aspetti differenti del rapporto fra stato socio-economico e rischio di tumore, differenziando il ruolo dei due livelli di analisi.

Questo punto serve per introdurre anche le problematiche legate al più volte citato utilizzo di fattori socio-economici che descrivono la zona di residenza dell'individuo come *proxy* delle caratteristiche dell'individuo stesso, nonché dell'utilità che esse hanno per fornire informazioni sulle disuguaglianze in termini di possibili strategie di azioni mirate per la salute pubblica.

Abbiamo già detto, infatti, dell'esistenza delle disuguaglianze sanitarie all'interno di una nazione e questo problema sta alla base della ripartizione delle risorse sanitarie a livello macro-geografico, ad esempio regionale.

Occorre però ricordare che problemi di equità nella salute possono essere rilevati anche a livelli geografici più piccoli (comuni o singole aree metropolitane, aree montane, ecc.). I fondi destinati al soddisfacimento dei livelli essenziali di assistenza devono basarsi non soltanto sulle caratteristiche e i bisogni a livello aggregato nazionale o regionale, ma anche su quelli delle singole realtà territoriali.

Per questa finalità, occorre passare dall'analisi dei fattori socio-economici a livello individuale a quelli propri dell'area geografica in questione, cioè a misure ecologiche nelle quali le caratteristiche socio-economiche dell'area di residenza identificano il livello socio-economico di tutti i residenti (ad esempio, il tasso di disoccupazione di un'area di Censimento, o la percentuale di genitori *single*; Caranci et al. 2010).

Gli indicatori aggregati territorialmente possono quindi essere utili per due finalità: o rappresentare delle *proxy* dello stato sociale degli individui (in assenza delle disponibilità di indicatori individuali), oppure rendere conto degli effetti di contesto.

Soprattutto nel primo caso bisogna naturalmente tenere conto del possibile *bias* ecologico, cioè della possibilità che l'analisi aggregata non renda correttamente conto delle caratteristiche individuali. Ma se, ad esempio, si ipotizzasse che il ruolo esercitato sulla sopravvivenza dalle caratteristiche contestuali fosse maggiore dei fattori

individuali, risulterebbe più utile usare indicatori aggregati.

Le misure a livello ecologico possono aiutare a valutare meglio le possibilità di accesso ai servizi sanitari, la qualità dei servizi sanitari locali, i fattori ambientali in genere. Alcuni studi, che non hanno rilevato differenze nella sopravvivenza correlate alle caratteristiche socio-economiche individuali, hanno invece scoperto che le differenti caratteristiche delle zone (in cui le persone comprese negli studi vivevano) sembrano suggerire che i fattori contestuali sono in grado di influenzare più dei fattori individuali l'accesso alle cure e la presa in carico.

Una serie di studi pionieristici sul rapporto fra sopravvivenza ai tumori e ruolo delle caratteristiche socio-economiche delle aree di residenza (Berg et al., 1977; Keirn et al., 1985; Linden, 1969; Page et al., 1980) aveva fornito una buona base empirica che verificava tale affermazione.

Sulla base di queste indicazioni si è sviluppato il concetto di “area piccola”, concetto peraltro determinato dalle situazioni contingenti in cui il ricercatore si trova ad operare. E' infatti evidente come la scelta della piccola area sia influenzata dalla disponibilità di dati socio-economici e quindi possa variare a seconda dei paesi in cui vengono svolti gli studi.

Le ricerche sviluppate in diverse nazioni utilizzano aree geografiche che, pur attenendosi al concetto di "area piccola", cambiano ampiamente: dalla sezione di censimento (census tract) utilizzata negli Stati Uniti (5.000 residenti) ed in Canada (4.000 residenti) (Gorey et al., 1997; Potosky et al., 1997; Singh et al., 2003; Wegner et al., 1982) al quartiere (block group, dalla dimensione variabile; Basset et al., 1986); dal codice postale (post code, dalla dimensione variabile; Lamont, 1993) alla sezione elettorale (electoral ward, 5.000 residenti; Coleman et al., 2004; Milner et al., 1987; Whynes et al., 2003) fino al distretto rionale (500 residenti; ad esempio, in Kaffashian et al., 2003; Wrigley et al., 2003), utilizzati in Inghilterra.

In Italia si utilizza la sezione di Censimento, che è un'area molto piccola, avente in media 170 residenti (Caranci et al., 2010).

Tutti questi studi e le considerazioni precedenti tendono a dimostrare che più l'area considerata è piccola e socialmente coesa (in merito, ad esempio, Cislaghi, 2009), più tenderà a ridursi il possibile *bias* ecologico. Se, ad esempio, il reddito medio di una popolazione residente in un'area verrà assegnato come reddito individuale ad ogni

paziente residente in quella comunità, la distorsione sarà minore quanto più l'area geografica sarà piccola e la comunità residente socio-economicamente coesa (Schrijvers et al., 1994).

La maggior parte degli studi, che indagano il peso dei fattori sociali sulla sopravvivenza per tumore, utilizzano fattori aggregati a livello di piccola area di residenza o, meno frequentemente, a livello individuale. Più rari sono invece gli studi che hanno investigato la relazione degli esiti sanitari con indicatori socio-economici di tipo ecologico a livello macro-geografico; si tratta principalmente di lavori su incidenza e mortalità per svariate patologie che hanno utilizzato indicatori economici misurati a livello di nazione.

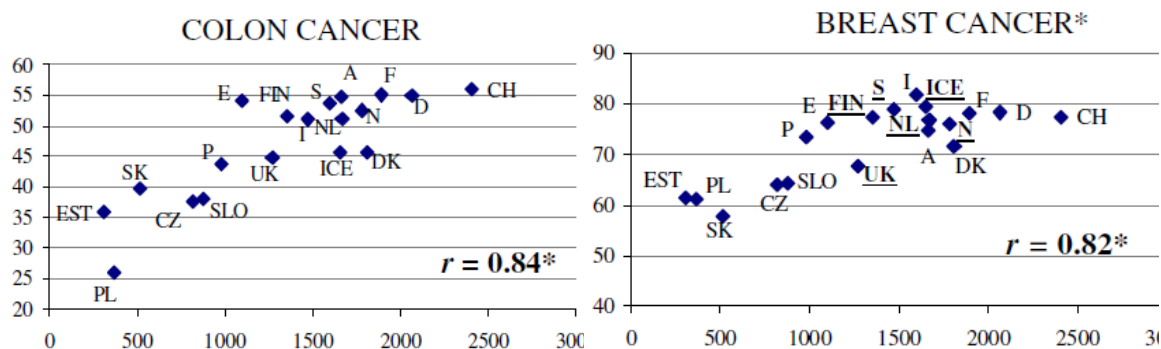
A questo proposito, è stato particolarmente indicativo uno studio olandese sulla correlazione tra standard di vita (prodotto interno lordo pro-capite, accesso a mezzi di trasporto privati, stato occupazionale) e mortalità nei paesi dell'Unione Europea (Mackenbach et al., 1994). I risultati mostravano una forte associazione tra una più alta mortalità generale e più bassi standard di vita, quando l'analisi era sviluppata a livello individuale. Tale associazione, tuttavia, scompariva a livello nazionale a causa di fattori confondenti, correlati sia con l'alta mortalità sia con standard più alti. L'associazione ricompariva, anche se meno forte, quando i fattori, una volta identificati, venivano controllati tramite regressione multipla.

Questa distorsione, dimostrata per la mortalità totale, esiste anche nella relazione tra fattori socio-economici misurati a livello macro-geografico e sopravvivenza per tumore. La correlazione tende ad essere più forte nei gruppi di paesi con un più ridotto livello di standard di vita rispetto a quelli più ricchi. Infatti, sopra un certo grado di sviluppo economico l'associazione tra condizione sociale e salute diventa più debole, perché gli indicatori macro-economici non riflettono appropriatamente le condizioni di vita dei singoli cittadini. La qualità di vita di una singola persona, infatti, non aumenta proporzionalmente al livello degli indicatori macro-economici, come il prodotto interno lordo o la spesa sanitaria totale.

Si è infatti osservato che la correlazione tra sopravvivenza e indicatori macro-economici (prodotto interno lordo, spesa pubblica totale, spesa sanitaria e sociale) nei paesi dell'Unione Europea è rappresentata da una curva molto ripida quando si considerano paesi con una ricchezza medio-alta; la pendenza si attenua e, per i paesi più ricchi, tende

a livellarsi, perché oltre un certo valore di ricchezza la sopravvivenza non cresce più proporzionalmente (vedi Figura 2.2) (Quaglia et al., 2005; Vercelli et al., 2006).

Figura 2.2 – Sopravvivenza relativa a 5 anni per tutti i tumori maligni (a meno di tumori della cute non melanomi) e prodotto interno lordo pro-capite per nazioni europee



Fonte: Vercelli et al., 2006.

Inoltre, come già visto, più ampia è l'area esaminata, maggiore è la disomogeneità delle caratteristiche; in ciascun paese esistono non solo forti differenze geografiche nella distribuzione dei fattori di rischio, con ampie differenze di incidenza e mortalità (AIRTUM Working Group, 2006), ma anche un'alta variabilità di condizioni demografiche, sociali, economiche e di disponibilità e qualità dei servizi sanitari, capaci di influenzare fortemente la sopravvivenza (Teppo, 1984).

Tuttavia, le analisi che valutano gli indicatori epidemiologici, in particolare la sopravvivenza, a livello macro-ecologico rispondono ad alcune necessità. (Liberatos et al., 1988). Primo fra tutti, la possibilità di migliorare la programmazione sanitaria e una più giustificata allocazione delle risorse nelle diverse aree di un paese.

Esistono infatti prove forti di una stretta relazione tra condizioni socio-economiche ed esiti sanitari a livello di macro aree geografiche e gli economisti sanitari hanno dimostrato come la ricchezza in un paese sia una forte determinante di spesa pubblica e privata e delle risorse sanitarie messe a disposizione per la salute pubblica (Kanavos et al., 1999; Tresseras et al., 1992; Waaler et al., 1984). Ampie differenze nei fattori socio-economici, risorse sanitarie ed esiti sanitari sono state osservate fra differenti paesi anche quando sono stati considerati solo paesi appartenenti all'area occidentale

sviluppata (Musgrove et al., 2002).

Studiare la sopravvivenza dei pazienti oncologici a livello macro-ecologico permette quindi di includere tra i possibili determinanti le disuguaglianze nell'organizzazione dei sistemi sanitari che, come osservato da Micheli e collaboratori, presentano anche all'interno dell'Europa fortissime differenze (Micheli et al., 2003), perché entra in gioco il ruolo che le politiche economiche, sociali, di *welfare*, hanno nel moderare l'impatto sulla salute delle condizioni di vulnerabilità.

Studiare tali differenze ed interpretare la correlazione tra la ricchezza di un paese e le caratteristiche del suo sistema sanitario nazionale è essenziale per sviluppare politiche e strategie finalizzate a ridurre le differenze di salute.

In particolare, uno studio italiano ha investigato a livello di nazioni europee la correlazione tra sopravvivenza e fattori socio economici, comprese le principali caratteristiche organizzative dei sistemi sanitari (Verdecchia et al., 2008). Prodotto interno lordo e spesa per la salute influenzavano altamente la sopravvivenza di pazienti oncologici. Paesi con bassa sopravvivenza, come Regno Unito e Danimarca, ma con relativamente elevata ricchezza avevano a disposizione un numero più ridotto di apparecchiature per TAC, utili per la diagnosi precoce. Tra i paesi con prodotto interno lordo più basso, la Polonia mostrava le peggiori prognosi per un gran numero di tumori, essendo anche il paese che dedicava meno risorse alla spesa per la salute.

Paragonare società con differenti standard economici, caratteristiche del *welfare*, abitudini di vita, organizzazioni dei sistemi sanitari e disponibilità di risorse sanitarie può aiutare a comprendere come queste caratteristiche possono essere correlate alle variazioni della sopravvivenza fra paesi.

2.5 Fattori epidemiologici, socio-economici e sopravvivenza dei pazienti oncologici

In letteratura esistono sostanzialmente tre tipi di studi riguardanti l'associazione tra sopravvivenza dei pazienti oncologici e fattori socio-economici.

Il primo gruppo ha come obiettivo quello di analizzare la correlazione esistente tra sopravvivenza e *status* socio-economico e quantificare tramite i rischi relativi la forza di tale correlazione.

Il secondo gruppo indaga se i fattori socio-economici sono in grado di influenzare le variabili cliniche, senza considerare la sopravvivenza. L'obiettivo di questi lavori è

comprendere se un diverso livello di condizione sociale si associa ad un accesso più o meno facile ai servizi sanitari, con un conseguente più o meno avanzato stadio alla diagnosi e appropriatezza di trattamento.

Esiste infine un terzo tipo di lavori, non molto numerosi, che possiedono entrambe le caratteristiche dei due tipi precedenti e si pongono come obiettivo sia l'indagine della relazione tra sopravvivenza e stato sociale, sia l'interpretazione di tale relazione tramite l'analisi delle sue possibili cause.

L'associazione tra fattori socio-economici e tassi di incidenza e mortalità è stata diffusamente studiata e si sono rilevate forti correlazioni specie per alcune sedi (tra cui polmone e vie aeree e digestive superiori negli uomini, mammella e cervice uterina nella donna), non necessariamente con svantaggi per le basse classi sociali (ad esempio, il tumore della mammella tende ad avere una maggiore incidenza nelle donne di più elevata posizione sociale) (Kogevinas et al., 1997).

Le relazioni con la sopravvivenza sono state indagate da metà anni '80, periodo in cui si è sviluppato un notevole interesse in seguito alle differenze rilevate negli Stati Uniti tra differenti gruppi etnici (Young et al., 1984). Si è così compresa la potenziale importanza delle differenze socio-economiche nell'influenzare una diagnosi precoce ed un trattamento appropriato, ossia la completa presa in carico del paziente oncologico.

Sono state dimostrate differenze rimarchevoli in sopravvivenza a seconda dello stato socio-economico in differenti popolazioni. In quasi tutti gli studi si osservano rischi relativi di morte più alti per le classi più svantaggiate. I valori di tali rischi non sono eccessivamente elevati, variando tra 1 e 1,5 a 1. Più frequentemente, il gruppo più deprivato presenta un eccesso di mortalità che varia dal 30% al 50% rispetto al gruppo più ricco. I rischi relativi si modificano, non consistentemente, a seconda del sesso, del paese in cui è stato svolto lo studio e dell'indicatore socio-economico utilizzato.

Per quanto riguarda il tipo di tumore considerato, i più alti rischi relativi di morte sono stati osservati generalmente per i cancro ad intermedia o buona prognosi, come quello della mammella, del colon-retto, della vescica e del corpo dell'utero (Kogevinas et al., 1997; Woods et al., 2006). La sopravvivenza per l'insieme di tutti i tumori maligni è peggiore per i pazienti con peggiore condizione socio-economica in tutti gli studi. Le analisi sui tumori dello stomaco, del polmone, della prostata e dell'ovaio hanno evidenziato, nella maggior parte dei casi, rischi maggiori per i gruppi a più basso livello

socio-economico, sebbene qualche studio non abbia rilevato alcuna associazione.

Gli studi che non hanno rilevato associazioni sembrano presentare due problemi metodologici. In un primo caso, i pazienti sono assegnati, a seconda della loro residenza, ad aree geografiche ampie, con caratteristiche eterogenee degli indicatori socio-economici (Supramanian et al., 1998; Woods et al., 2005). In altri casi la popolazione studiata sembrerebbe troppo piccola per poter ottenere stime significative (Ciccone et al., 2000; Moran et al., 2004).

È particolare invece il caso dei tumori pediatrici, la cui prognosi non sembra, nei pochi studi effettuati, influenzata dal livello socio-economico (Coleman et al., 1999; Schillinger et al., 1999).

Non è comunque facile riuscire a comprendere come lo *status* socio-economico determini differenze nella sopravvivenza ai tumori, perché, come già detto, le condizioni sociali dei pazienti oncologici influiscono sui risultati sanitari attraverso meccanismi e determinanti multifattoriali.

D'altra parte, è complesso verificare quanto lo stato socio-economico influenzi un determinante prognostico intermedio con effetto sulla sopravvivenza e sulle recidive, anche perché spesso vengono utilizzati surrogati e misure *proxy* per definire una condizione socio-economica, non sempre in grado di rappresentare compiutamente tale condizione.

Un singolo e specifico determinante prognostico non spiega spesso tutta la differenza in sopravvivenza tra classi ricche e deprivate. Per eliminare le disuguaglianze dovute a differenze socio-economiche occorre studiare ed identificare le diverse cause di tali differenze e, senza concentrare l'attenzione su un singolo fattore, raccogliere informazioni multiple riguardanti sia la malattia, sia il paziente.

Combinando le informazioni derivate dal secondo e terzo tipo di studi sopra citati, emerge che i possibili determinanti dell'associazione tra deprivazione e peggiore prognosi possono essere divisi in tre differenti classi: caratteristiche del tumore, caratteristiche del paziente e fattori correlati al percorso diagnostico-terapeutico⁷.

Si può innanzitutto ipotizzare che una condizione di deprivazione socio-economica

⁷ Si è visto in precedenza che diversi autori hanno anche verificato che l'organizzazione dei sistemi sanitari gioca un importante ruolo (ad esempio, Auvinen et al., 1997), ma è complesso indagare questo aspetto a livello individuale o micro-ecologico e, come visto sopra, il problema va trattato a livello geografico macro (provincia, regione, nazione; ad esempio, in Lillini et al., 2011; Vercelli et al., 2012 e 2014).

possa influenzare un tempestivo accesso alle cure, sia per un ritardo del paziente nel rivolgersi al medico, sia per un ritardo di chi deve fornire i servizi medici. L'accesso alle cure è un concetto assai complesso che può essere sintetizzato come: "un tempestivo uso di servizi sanitari personali per ottenere i risultati di salute migliori possibili".

Svariati aspetti determinano una maggiore o minore difficoltà di accesso, quali la complessità dell'organizzazione del sistema sanitario, i criteri di allocazione delle risorse, le liste di attesa, le barriere dovute a difficoltà nel trasporto o a difficoltà di comunicazione con gli operatori sanitari, la frammentazione delle cure e la conseguente difficoltà di aderire ai protocolli di trattamento, la frequenza dei contatti con il medico di base per regolari procedure di prevenzione e terapia, la frequenza dell'utilizzo delle strutture sanitarie in regime di emergenza o elettivo.

L'accesso influisce a cascata sia sulle pratiche diagnostiche, sia sulle procedure terapeutiche; la conseguenza più immediata è l'effetto sull'estensione del cancro alla diagnosi, che poi, a sua volta, finisce per determinare la scelta di un trattamento con intento curativo o solamente palliativo.

Essendo lo stadio del tumore alla diagnosi ed il tipo di trattamento i principali determinanti della prognosi di un paziente oncologico, è bene iniziare l'analisi degli eventuali specifici fattori causali proprio dall'estensione della malattia alla diagnosi. Successivamente verranno descritti tutti gli altri elementi che la letteratura ha mostrato agire a livello individuale come determinanti e che sono stati utilizzati nello studio descritto in questa tesi⁸, nell'ordine: grado di malignità, presenza di comorbidità alla diagnosi, trattamenti terapeutici, età e sesso, elementi socio-economici individuali (titolo di studio, stato civile e supporto sociale).

2.5.1 Estensione della malattia alla diagnosi

Le differenze di stadio alla diagnosi spiegano molte delle differenze di sopravvivenza, ad esempio quelle riscontrate in Europa tra differenti paesi partecipanti allo studio EURO CARE; tali risultati concordano sul fatto che, al di là delle caratteristiche socio-economiche, lo stadio sia un forte determinante prognostico sia in studi ospedalieri, sia in analisi su ampia scala a livello di popolazione.

Negli Stati Uniti, già a partire dagli anni '80 e a tutt'oggi, a seguito dell'attenzione rivolta

⁸ Si è già discusso in precedenza del ruolo dei descrittori socio-economici dell'area in cui risiede l'individuo/il paziente, anche essi utilizzati in questo studio.

alle disuguaglianze sanitarie tra le differenti componenti etniche della popolazione, si sono iniziate a studiare le variazioni di prognosi valutando il ruolo giocato dallo stadio nel determinare differenti tassi di sopravvivenza a seconda della razza considerata (Baquet et al., 2000; O'Malley et al., 2003; Roetzheim et al., 2000).

Per quanto riguarda più propriamente sopravvivenza, stato socio-economico e stadio alla diagnosi, esistono in letteratura numerose evidenze della loro stretta correlazione (Hansen et al., 2008).

Ad esempio, i pazienti maggiormente deprivati presentano alla diagnosi malattia più avanzata in molti studi sul cancro del colon-retto (ad esempio, Parikh-Patel et al., 2006; Schwartz et al., 2003; Wrigley et al., 2003), su quello della mammella (ad esempio, Macleod et al., 2000; O'Malley et al., 2003; Rutqvist et al., 2006; Yabroff et al., 2003), su quello della prostata (Liu et al., 2001; Schwartz et al., 2003) e così via.

I risultati differiscono a seconda del tipo di tumore considerato e l'effetto sembrerebbe maggiore per gli organi a media e buona prognosi, meno rilevante per quelli a prognosi cattiva.

In tutti gli studi che hanno incluso nell'analisi dei rischi relativi di ogni condizione sociale anche l'aggiustamento per stadio, si è dimostrato come le differenze nella sopravvivenza non fossero completamente spiegate dall'estensione del tumore, ma esistessero comunque differenze residuali. Questa situazione è apparsa particolarmente chiara, ad esempio, in diversi studi sul cancro della mammella (Kaffashian et al., 2003; Polednak, 2002; Schrijvers et al., 1995), su quello del colon-retto (Polednak, 2001; Hole et al., 2002) e così via.

Alla tempestività della diagnosi contribuisce anche l'adesione ai programmi e messaggi di prevenzione primaria per quei tumori che sono oggetto di *screening*.

Il minore accesso ai servizi di diagnosi viene evidenziato dai dati sulle disuguaglianze di accesso ai servizi di *screening* osservati in molti paesi (Gakidou et al., 2008). Dalle evidenze della letteratura, anche se non numerosissime, la deprivazione si associa ad una minore adesione sia ai programmi di *screening*, sia allo *screening* opportunistico.

Uno studio di popolazione francese ha studiato le caratteristiche dello *screening* per il cancro della cervice uterina. In tale studio, l'età, la densità di popolazione, la disponibilità dei servizi ed il reddito medio-basso erano tutte variabili correlate negativamente con la partecipazione allo *screening* (Challier et al., 2000).

Esistono dati anche per il colon-retto che testimoniano minore partecipazione allo *screening* da parte di soggetti con basso status socio-economico, che arrivano alla diagnosi con uno stadio più avanzato e una conseguente peggiore sopravvivenza (Singh et al., 2004; Whynes et al., 2003).

In Svezia caratteristiche più o meno specifiche di classe sociale svantaggiata, quali disoccupazione, sovraffollamento abitativo e basso reddito, si correlavano con una probabilità minore di aderire allo *screening* per il tumore della mammella nel lavoro di Zackrisson et al. (2004).

Un ampio studio di popolazione statunitense, più di 118.000 partecipanti, ha osservato una correlazione tra basso status socio-economico e più bassi livelli di *screening* per i tumori della mammella e del colon retto ma non della cervice uterina; anche dopo una correzione per fattori individuali, il livello di povertà dell'area di residenza rimaneva associato al non avere mai avuto una mammografia, un esame clinico della mammella, una colonscopia/sigmoidoscopia o un test del sangue occulto (Schoutman et al., 2006).

2.5.2 *Biologia del tumore*

E' stato ipotizzato che differenze nelle caratteristiche biologiche (come il grado di malignità) del tumore possano essere correlate con differenti condizioni socio-economiche. Gli studi in proposito non sono numerosi e, soprattutto, non raggiungono risultati univoci.

Il cancro della mammella è forse il tumore più studiato sotto questo punto di vista; morfologia e grado di differenziazione non sembrano spiegare differenze in sopravvivenza dovute alla classe sociale e neppure lo stato recettoriale è risultato variare in modo chiaro a seconda del livello di deprivazione (Carnon et al., 1994; Macleod et al., 2000; Miller et al., 2002).

2.5.3 *Ruolo delle comorbidità*

Si è ipotizzato che, tra le caratteristiche legate allo stato del paziente, le comorbidità, ossia le patologie differenti dal tumore presenti alla diagnosi, possano avere un effetto nel mediare l'influenza della classe sociale sulla sopravvivenza.

In effetti, esistono diversi lavori che confermano una maggiore frequenza di comorbidità nei pazienti con una peggiore situazione socio-economica (ad esempio,

Dalton et al., 2007; Paterson et al., 2002; Schrijvers et al., 1997).

Uno studio di popolazione sul tumore della mammella che ha coinvolto circa 26.000 soggetti in Danimarca, con parità di accesso alle cure e largo utilizzo di linee guida per il trattamento, ha rilevato che il rischio di morte, già più alto nelle pazienti deprivate, si alzava ulteriormente in presenza di condizioni patologiche concomitanti (Dalton et al., 2007).

L'ipotesi è che soggetti appartenenti a classi sociali più basse abbiano un'esposizione più forte ai fattori di rischio, sia quantitativamente che qualitativamente (in particolare a fumo, alcol, caratteristiche della nutrizione). Tali differenze possano comportare un carico maggiore di patologie croniche al momento della diagnosi di tumore.

Anche per le comorbidità vale quanto già discusso precedentemente per lo stadio: gli studi che hanno utilizzato e utilizzano metodi di regressione multivariata non riescono, aggiustando per comorbidità e altri fattori, a spiegare tutta la differenza prognostica dovuta ai fattori socio-economici (Polednak, 2001; Reyes-Ortiz et al., 2006; Villingshøj et al., 2006).

2.5.4 *Trattamento terapeutico*

La scelta del tipo di trattamento risulta, almeno teoricamente, un fattore causale importante della variabilità dei tassi di sopravvivenza per stato sociale.

In generale, gli studi considerati riportano risultati simili a quelli già visti per lo stadio: la scelta della terapia non è in grado da sola di annullare le differenze in sopravvivenza per classe sociale.

Gli studi più frequenti riguardano il tumore della mammella; durante gli ultimi due decenni si è notato un miglioramento della sopravvivenza ed una diminuzione della mortalità, principalmente dovuti alle campagne di *screening* introdotte in numerosi paesi europei negli anni '90, ed al miglioramento dei trattamenti, fondamentalmente delle terapie adiuvanti (Berry et al., 2005; Lemmens et al., 2005; Sant et al., 2006).

Alla luce di tutto ciò non sono incoraggianti i risultati circa le differenze di trattamento per condizione socio-economica, che indicano un minor utilizzo di terapie adiuvanti nelle pazienti deprivate, affette da tumore della mammella. Queste ultime ricevono meno frequentemente la radioterapia dopo interventi di chirurgia conservativa o chemioterapia per linfonodi positivi; sono inoltre sottoposte meno frequentemente a

linfadenectomia e più in generale a trattamenti radicali o con intento curativo (McGinnis et al., 2000; Osborne et al., 2005; Quaglia et al., 2011).

Anche per il tumore del colon-retto è stata osservata una diminuzione di terapie adiuvanti nei pazienti deprivati (Cicolallo et al., 2005; Lemmens et al., 2005).

Spesso il correggere per stadio e tipo di trattamento non annulla le differenze in sopravvivenza per condizione sociale (Hole et al., 2002; Polednak, 2001 e 2002), mentre in alcuni casi le differenze si annullano (Munro et al., 2004; Wrigley et al., 2003).

I lavori che riscontrano differenze residue considerano un numero di siti tumorali molto più vasto (testa e collo, melanoma, laringe, collo dell'utero, mammella e colon-retto), mentre le differenze si annullano completamente solo in studi sui tumori del colon-retto e della vescica.

Uno studio sui tumori del sistema nervoso centrale ha evidenziato nei pazienti più deprivati uno stadio più avanzato ed un uso meno frequente di chirurgia e radioterapia (Chang et al., 2005). Per il cancro della prostata si è osservato nelle classi sociali più basse una minore frequenza di asportazione radicale chirurgica, con preferenza di radioterapia o di nessun intervento con intento curativo (Denberg et al., 2005). Allo stesso modo pazienti con tumori del polmone, del colon-retto e della mammella, appartenenti a classi più deprivate vengono sottoposti meno frequentemente ad intervento chirurgico, rispetto alla controparte più benestante (Blais et al., 2006; Pollock et al., 1998).

È possibile che le differenze in sopravvivenza ancora evidenti dopo correzione per tipo di trattamento siano dovute, almeno in parte, alle modalità di somministrazione della terapia, influenzando la qualità delle cure.

Questo assunto sembrerebbe confortato dal fatto che in certi studi i pazienti con più bassa classe sociale vengono meno di frequente inviati a specialisti o indirizzati a centri oncologici di riferimento (Herbert et al., 2002). In un ampio studio di popolazione svolto nel Regno Unito, i pazienti più deprivati erano sottoposti più frequentemente ad interventi chirurgici di emergenza piuttosto che elettivi e afferivano a centri con un basso numero di casi curati abitualmente (Pollock et al., 1998).

Secondo altri autori, donne con cancro della mammella appartenenti ad una più alta classe sociale avevano una probabilità maggiore di essere trattate da chirurghi

specializzati (Kingsmore et al., 2003 e 2004).

Un fattore correlato fortemente alla buona prognosi dei pazienti affetti da tumore del colon-retto è sicuramente la chemioterapia adiuvante e il completamento dei cicli a dosi adeguate.

Un ampio studio statunitense di popolazione, basato su dati raccolti dai registri tumori afferenti al SEER, ha rilevato che pazienti deprivati, donne e vedovi di entrambi i sessi hanno maggiore probabilità di non completare i cicli di chemioterapia e di avere, conseguentemente, tassi di sopravvivenza più bassi (Dobie et al., 2006).

È difficile comunque, se non attraverso le vie indirette appena citate, rendere ragione della qualità delle cure poste in essere, essendo la qualità stessa influenzata da un insieme di fattori difficili da controllare, legati al paziente, alla disponibilità di risorse mediche avanzate e all'esperienza di tutti i differenti soggetti implicati nell'intero percorso diagnostico-terapeutico.

2.5.5 Fattori demografici: età e sesso

Nell'ampio contesto dei fattori socio-economici occorre considerare gli indicatori demografici, focalizzando l'attenzione su età e sesso. Si tratta di due fattori spesso non considerati a fondo, poiché tanto evidenti quanto semplici, ma estremamente importanti nel determinare differenze prognostiche, in modo particolare per quanto concerne le neoplasie.

Il cancro è una malattia cronica, il cui maggiore fattore di rischio è rappresentato dall'età, aumentando esponenzialmente sia i tassi di incidenza sia quelli di mortalità dopo i 50 anni. In Europa, ad esempio, nel 2012 l'incidenza per tutti i tumori combinati andava, negli uomini, da 462 casi per 100.000 nel gruppo di età tra i 50 e 54 anni a 2.209 nel gruppo fra 70 e 74 anni, mentre nelle donne si passava da 488 a 1.085. Sempre nel 2010, più del 70% delle morti per cancro si verificava in pazienti anziani ultrasessantacinquenni (WHO, 2015).

In più, le donne di età più avanzata, specie se di bassa istruzione, tendono ad arrivare con maggiore ritardo alla diagnosi (ad esempio, si veda il lavoro iniziale di Montella et al., 2001).

Studi di epidemiologia descrittiva su popolazioni molto numerose di macro-regioni europee hanno ormai dimostrato, con larga evidenza, che non solo esistono differenze in

sopravvivenza per età tra giovani e molto anziani, ma che tali differenze riguardano anche segmenti di popolazione con età avanzata e molto simile (Quaglia et al., 2007; Vercelli et al., 1998 e 2000).

Si è osservato un *gap* prognostico molto elevato anche paragonando gruppi di età oltre i 55 anni: praticamente per tutte le sedi tumorali i pazienti compresi fra 65 e 84 anni presentano probabilità di morte molto più alte dei pazienti del decennio precedente (tra 55 e 64 anni).

I rischi relativi appaiono sempre più alti a un anno dalla diagnosi, ma anche a cinque anni si registrano valori molto elevati. Alcuni dati, riferiti ai tumori più frequenti, appaiono particolarmente allarmanti. Per il cancro della mammella i rischi relativi di eccesso di morte nelle anziane rispetto alle adulte di mezza età sono di 2,6 volte più alti a un anno dalla diagnosi e 1,5 volte più alti a cinque anni dalla diagnosi. Anche per il tumore della prostata i soggetti anziani hanno un eccesso di rischio rispetto agli adulti più giovani pari a 2,1 e 1,5 a uno e cinque anni dalla diagnosi rispettivamente.

Gli anziani di entrambi i sessi presentano un rischio maggiore del 30-40% anche per i tumori del colon e del retto, a cinque anni dalla diagnosi (Quaglia et al., 2007).

Anche il sesso gioca un ruolo importante quale determinante prognostico, in particolare se considerato in stretta connessione con l'età alla diagnosi.

Le donne di tutte le età presentano tassi di sopravvivenza quasi sempre maggiori della controparte maschile (Micheli et al. 1998), vuoi per ragioni biologiche, vuoi per una maggiore attenzione che il sesso femminile possiede per il proprio corpo ed i suoi fenomeni fisiologici e patologici. Sembra però che tale vantaggio prognostico parzialmente diminuisca studiando pazienti anziane ultrasessantacinquenni; le donne con età compresa fra 65 e 84 anni hanno rischi relativi di morte, rispetto alle donne più giovani fra 55 e 64 anni, molto maggiori degli uomini anziani rispetto agli adulti di mezza età. Questo accade per tutti i cancri combinati e per ognuna delle sedi tumorali maggiori. In particolare, i tumori ginecologici e della mammella mostrano i rischi maggiori compresi tra 2,3 e 2,8 a un anno dalla diagnosi e tra 1,5 e 2,5 a cinque anni (Quaglia et al., 2007).

Oltre ai classici fattori chiamati in causa per spiegare le differenze in sopravvivenza dovute all'età, quali problemi diagnostici che conducono ad uno stadio più avanzato, problemi correlati al trattamento o alle comorbidità (Bouchardy et al., 2003; Fentiman et

al., 1996; Lavelle et al., 2007), occorre tenere in considerazione tutti i fattori socio-economici che caratterizzano gli individui.

Appare logico supporre che il supporto sociale acquisisca maggiore importanza in riferimento agli anziani, in quanto in grado di modificare pesantemente le abitudini e gli stili di vita, quale ad esempio l'approccio nei confronti delle strutture sanitarie.

Studi anche datati su popolazione ampie hanno mostrato un effetto indipendente sulla sopravvivenza anche dopo correzione per altri fattori (Goodwin et al., 1987, Reynolds et al., 1990); altri due studi non hanno confermato tale influenza, perché l'effetto era mediato dalle difficoltà di accesso alle cure e da un più avanzato stadio alla diagnosi (Cassileth et al., 1985; Milner et al., 1987).

Lavori più recenti hanno evidenziato correlazioni molto alte tra la sopravvivenza dei pazienti anziani, i fattori macro-economici, le risorse sanitarie ed alcuni indicatori demografici quali stato civile e numerosità delle famiglie (Quaglia et al., 2005; Vercelli et al., 2006). È più probabile che un paziente anziano ricerchi il consiglio medico con ritardo rispetto all'inizio dei sintomi, oltre che per motivi legati alla presentazione della malattia, anche per problemi sociali, psicologici e cognitivi che conducono, specie per le donne e per i tumori ginecologici, ad uno stadio più avanzato alla diagnosi e a terapie adiuvanti/palliative e non curative (Ferrante et al., 2000; Kant et al., 1992; Potosky et al., 1987, 184-185).

2.5.6 L'istruzione

Quando ci si vuole rendere conto della capacità dei pazienti di valorizzare le proprie opportunità di salute, il livello di istruzione rappresenta, da un punto di vista metodologico, una variabile sociale più affidabile, per certi aspetti, rispetto ad indicatori quali occupazione e reddito, e questa ipotesi di studio è presente già nei primissimi lavori che hanno utilizzato tale indicatore per evidenziare differenze di tipo socio-economico negli indicatori epidemiologici (come nel lavoro sulla mortalità negli Stati Uniti di Kitagawa et al., 1973).

Il livello di istruzione è, infatti, un indicatore meno soggetto a variazioni nel tempo e correlato con lo stile di vita anche molti anni dopo il periodo scolastico. Il livello educativo raggiunto può comportare, attraverso vari fattori intermedi, una differenza in mortalità e sopravvivenza legata alla deprivazione. Uno dei primi studi a questo

riguardo identificò un rischio più elevato di mortalità per cancro del polmone e bassi livelli di istruzione; tale svantaggio si annullava dopo aver controllato per comorbidità, stadio alla diagnosi e tipo istologico (Stavraki et al., 1987).

Diversi autori hanno riscontrato un effetto positivo e indipendente del grado di scolarità, sia sulla sopravvivenza generale, sia su quella di gruppi di pazienti affetti da specifiche malattie (Guralnik et al., 1993; Kogevinas et al., 1997; Marmot et al., 1987).

In particolare, alcuni studi focalizzati su educazione e stato cognitivo hanno dimostrato un'associazione con la sopravvivenza di pazienti oncologici, anche dopo avere effettuato aggiustamenti per stadio e trattamento, suggerendo ulteriori meccanismi di azione.

Due studi svedesi, condotti a livello nazionale, hanno riscontrato, anche in un paese con ottimo accesso alle cure e universale copertura degli *screening*, una stretta correlazione tra l'educazione raggiunta e la sopravvivenza; soggetti laureati avevano tassi di sopravvivenza maggiori del 40% di quelli di individui con meno di nove anni di scolarità (Hussain et al., 2008, (a) e (b)).

Un basso livello educativo può comportare problemi nella capacità di utilizzare i servizi sanitari a disposizione, sia in fase di prevenzione che di cura e di *follow-up*. Conoscenza e consapevolezza possono influenzare non solo il tempo intercorrente tra i primi sintomi e la prima visita, ma anche il successivo decorso.

2.5.7 Il supporto sociale nella sopravvivenza al cancro

Il supporto sociale può essere definito in due modi differenti: una definizione oggettiva che descrive il supporto come rete di relazioni tra individuo (o gruppi) e contesto (disponibilità di parenti, frequenza di contatti con amici o familiari, organizzazioni, associazioni); una definizione soggettiva, forse meno valida da un punto di vista epidemiologico, che descrive il supporto attraverso la percezione personale e la valutazione della qualità e dell'adeguatezza delle relazioni sociali.

Nel caso dei malati di cancro, le due definizioni vanno riferite a supporto emozionale e aiuto materiale tangibile (Robert, 1998). Purtroppo esistono problemi metodologici che rendono difficile studiare questo fattore, quali la difficoltà a studiare gruppi di pazienti sufficientemente numerosi o la difficoltà di misura oggettiva, tenendo nel contempo conto dei fattori clinici.

In un ampio studio di popolazione si era osservato che un insufficiente supporto sociale,

identificato dallo stato civile dei pazienti, è un fattore di rischio per una peggiore prognosi, a causa di una ritardata diagnosi e problemi di scelte terapeutiche (Goodwin et al., 1987).

Infatti, le relazioni sociali possono avere influenza già a livello di partecipazione a *screening* per cancro della mammella e della cervice (Kang et al., 1993; Suarez et al., 1994); una mancanza di rapporti sociali è legata a problemi di adesione corretta alla terapia, particolarmente negli anziani, che comporta una minore probabilità di ricevere un trattamento con finalità curativa per limitato accesso ai trasporti o mancanza di assistenza durante i cicli di radio e chemioterapia (Goodwin et al., 1993; Ryan et al., 1989). Sembrerebbe che anche il follow-up possa essere influenzato dalla rete di relazioni, costituita innanzitutto dal partner, dai figli e dal supporto informale (Mor et al., 1992).

Spesso il supporto sociale è risultato essere più scarso per individui a basso reddito o comunque più genericamente deprivati e tali fattori possono interagire nella loro influenza sulla sopravvivenza e contribuire alle differenze osservate nella prognosi (Macleod et al., 2004).

Il *network* sociale e certi tratti specifici della personalità sembrano essere associati con una prognosi favorevole anche dopo controllo per fattori prognostici clinici, come già risultava dagli studi di Hislop et al. (1987) e Waxler et al., (1991) sulle donne affette da tumore al seno.

Il supporto sociale potrebbe essere quindi considerato una sorta di modificatore d'effetto nelle associazioni tra fattori clinici e prognosi della malattia.

Una dimensione del supporto sociale che ha un'evidente influenza sulla salute è lo stato civile e molti studi hanno descritto l'effetto dello stato civile sulle condizioni di salute.

E' infatti noto che soggetti vedovi o mai sposati hanno un rischio generale di mortalità 2,3 volte più alto rispetto a soggetti sposati (ad esempio, Ho et al., 1991), mostrando come coloro che sono sposati hanno anche una migliore sopravvivenza una volta colpiti da tumore (Harvei et al., 1997; Kravdal, 2000; Polednak, 2002), forse perché sperimentano un più tempestivo accesso alle cure, come sembrava suggerire già lo studio di Ayanian et al. (1993) su donne con tumore al seno negli Stati Uniti.

Si è osservato come lo stato psicologico dei pazienti oncologici possa influenzare la loro prognosi; stati depressivi misurati secondo scale psicologiche sono stati, infatti,

associati ad una maggiore probabilità di recidive e ad una più bassa sopravvivenza a cinque anni dalla diagnosi (Watson et al., 1999).

Più in generale, infine, si osserva come, nell'ambito della consapevolezza del proprio stato di salute e dell'informazione sanitaria, due ricerche su donne affette da cancro della mammella (appartenenti a differenti livelli socio-economici) hanno riscontrato una minore capacità di identificare sintomi legati al tumore da parte di soggetti con un minore supporto sociale e, d'altra parte, una maggiore possibilità di ottenere notizie da parte di specialisti ospedalieri, accompagnata a minori preoccupazioni per problemi familiari e sanitari in donne di migliore condizione socio-economica (Grunfeld et al., 2002; Macleod et al., 2004).

2.6 Cancro, identità e effetti sulla qualità della vita

E' infine necessario fare cenno alla riflessione, che discende da molte delle osservazioni sviluppate in questo capitolo, fra l'interazione che esiste tra malattia oncologica ed identità, dal momento della diagnosi in poi, e quali riflessi essa abbia sulla qualità della vita dopo la fine delle cure.

Anche se non è un argomento strettamente sviluppato nello studio descritto in questa tesi, esso è stato affrontato e ritenuto rilevante in molti studi, che hanno verificato come vi siano forti interazioni fra le risorse sociali ed economiche, il modo in cui la diagnosi viene recepita e l'esito finale delle cure, nel momento in cui ad una persona è diagnosticato un tumore (Harwood et al., 2003; Hoffman et al., 2013).

Poche esperienze di vita possono essere fonte di maggiore paura e angoscia che ricevere una diagnosi di cancro, con le minacce alla propria *routine* di vita che essa comporta e le decisioni, trattamenti e sfide che si attivano nei malati e nelle loro famiglie. E, spesso, l'aspetto biomedico, mirato a trovare la strategia di cura più efficace, ha la meglio sulla considerazione del benessere psicosociale del paziente (Rowland, 2008).

Questo, malgrado gli studi sullo stress emotivo (che indubbiamente questa esperienza genera) abbiano mostrato che anche esso ha un ruolo sul modo con cui si affronta la malattia e sulla sopravvivenza a lungo termine (Hoffman et al., 2013).

Una volta diagnosticato, il tumore ha infatti il potenziale per influenzare quasi ogni aspetto della vita di un individuo: fisico, psicologico, interpersonale, professionale e spirituale (Aziz, 2007).

Rispetto agli individui che non hanno avuto il cancro, i sopravvissuti hanno più probabilità di avere problemi psicologici, peggiore salute, limitazioni funzionali (come l'essere in grado di guidare o lavorare) (Hewitt et al., 2003). Gli individui diagnosticati in età più giovane (prima dei 45 anni) sono risultati a più alto rischio di sperimentare problemi psicologici, che possono persistere nel tempo (Arndt et al., 2004; Costanzo et al., 2009; Hewitt et al., 2003).

In un sondaggio su larga scala su malati di cancro, 73% dei quali avevano terminato le cure da due o più anni, Wolff (2007) ha rilevato che oltre il 70% degli intervistati ha riferito di sperimentare la depressione a causa dell'esperienza della malattia, il 60% ha riferito problemi di relazione e l'83% ha riferito diminuzione del reddito. Oltre un quarto degli intervistati ha indicato che non avevano risorse sufficienti per soddisfare le loro necessità pratiche ed emotive (ad esempio, problemi legati al lavoro) e più della metà ha indicato che era più difficile affrontare le conseguenze emotive e pratiche del cancro piuttosto delle conseguenze mediche.

Eppure, nonostante questi elementi chiaramente negativi, molti intervistati hanno anche dimostrato una notevole capacità di recupero. Il 62% per cento aveva riferito di stare attualmente vivendo in buona salute e il 47% aveva dichiarato che l'esperienza del cancro aveva migliorato la loro vita in qualche modo (Wolff, 2007).

Nel suo studio, Rowland (2008) prende poi atto della resilienza dei sopravvissuti al cancro, il senso di crescita personale e di maestria che possono provare coloro che sopravvivono, nonché la necessità di comprendere meglio i fattori che possono garantire il loro benessere post-malattia. Allo stesso modo, Feuerstein (2007) e Foster et al. (2011) citano la necessità che i modelli teorici si concentrino sugli aspetti positivi della sopravvivenza al cancro.

E' interessante, a questo proposito, uno studio realizzato da Kaiser (2008) sul significato del concetto di identità per donne sopravvissute ad un tumore al seno.

Lo studio ha coinvolto 39 donne statunitensi che avevano completato le cure da tre a diciotto mesi prima e le cui storie sono state raccolte tramite interviste in profondità. Il punto focale dell'intervista è se esse si sentissero "sopravvissute", nel senso che avevano superato una sorta di ordalia da cui erano uscite segnate e, in qualche modo, cambiate. Nonostante il significato positivo attribuito al termine "sopravvissute", molte intervistate hanno dato un senso alterato al termine, o addirittura lo hanno respinto.

In particolare, la definizione di “superstite” era sentita aliena da coloro che lottavano con la minaccia di recidiva, o che affermavano di non ritenere la loro esperienza del cancro abbastanza grave da meritare questo titolo, o che desideravano vivere in maniera riservata l’esperienza della malattia e della guarigione.

Questi risultati mostrano come il vissuto della malattia sia fortemente individualizzato e mediato dalla propria storia di vita e di come andrebbe adeguatamente approfondito lo studio di tale percezione e interpretazione, per fornire un adeguato supporto a chi, dopo la fine delle terapie, avrebbe bisogno di recuperare, anche da questo punto di vista, qualità della vita e identità personale.

Malgrado questo, un recente sondaggio ha stimato che meno del 2% del *counselling* psicologico è impegnato nella ricerca sul cancro (Raque-Bogdan et al., 2013) e questo può rivelarsi un problema sulla sopravvivenza a lungo termine o, meglio, sulla sua qualità.

Se, infatti, gli interventi psicosociali non possono influenzare in modo ottimale i tassi di sopravvivenza, è però stato verificato da diversi studi come essi possano avere un grande potenziale per promuovere il recupero del benessere emotivo e gli altri aspetti dell’adattamento alla fine della malattia, consentendo in tal modo ai sopravvissuti di recuperare piacere e produttività dalla propria vita, compromesse dalle incertezze e interruzioni imposte dall’esperienza del tumore (Chow et al., 2006).

Questo potrebbe anche avere un effetto protettivo rispetto a eventuali riprese di malattia, anche se su questo punto si può parlare ad oggi solo di ipotesi e non di studi che le comprovino.

Capitolo 3

Il disegno della ricerca

Il capitolo si propone di presentare gli obiettivi dello studio e le motivazioni che stanno alla base delle scelte empiriche che sono state prese per realizzarlo, ossia quali sedi di tumore sono state analizzate, quali individui, quale periodo e quale area geografica sono stati considerati.

3.1 Gli obiettivi della ricerca

Questa tesi si pone due tipi di obiettivi: un obiettivo principale, che vuole rispondere al quesito di ricerca presentato nel primo capitolo, e due obiettivi secondari di natura metodologica. I due tipi di obiettivi sono fortemente interconnessi fra di loro; il raggiungimento degli obiettivi di tipo metodologico è, infatti, diretta conseguenza delle operazioni necessarie per il conseguimento dell'obiettivo principale.

Gli obiettivi possono, quindi, essere così sintetizzati:

- l'obiettivo principale è indagare la relazione fra le caratteristiche socio-economiche dei malati di tumore e la sopravvivenza;
- il primo obiettivo secondario è valutare se esista o meno una differenza nell'eventuale effetto che gli indicatori socio-economici individuali e quelli relativi all'area di residenza⁹ dei malati hanno sulla sopravvivenza. In questo modo si vuole anche esaminare l'esistenza e le caratteristiche della fallacia ecologica, ossia della distorsione (o *bias* ecologico) che si può verificare quando l'associazione che si osserva tra le variabili aggregate non rappresenta necessariamente l'associazione esistente a livello individuale;
- il secondo obiettivo secondario è condurre uno studio di fattibilità del *record linkage* fra banche dati messe a disposizione da fonti amministrative differenti. Nello specifico, tale studio ha unito i dati amministrativi di tipo sanitario (ossia le caratteristiche dei dati patologici, diagnostici e terapeutici della malattia) con i dati demografici e socio-economici individuali presenti nelle Anagrafi comunali

⁹ In questo studio l'area di residenza dei malati è espressa dalla sezione di Censimento.

e con i dati aggregati per sezione di Censimento della regione in cui è stato svolto lo studio, provenienti dal Censimento Nazionale della Popolazione e delle Abitazioni.

In merito all'obiettivo principale, questo lavoro intende verificare l'esistenza di associazioni fra le caratteristiche demografiche e socio-economiche dei pazienti, quelle della sezione di Censimento in cui sono residenti, le informazioni relative alle caratteristiche della malattia, dello stato alla diagnosi e delle terapie utilizzate (ossia le variabili indipendenti), e la sopravvivenza per due specifiche sedi tumorali, il colon-retto in entrambi i sessi e la mammella per le donne (ossia la variabile dipendente). Raggiungere questo tipo di obiettivo è rilevante, perché può fornire informazioni importanti su elementi che, sia direttamente sia indirettamente, influenzano l'esito finale della malattia (ad esempio, Akinyemiju et al., 2013; Bentley et al., 2009).

Utilizzare in un unico modello tutte le variabili sanitarie e socio-economiche permette, infatti, di investigare più attentamente gli effetti indiretti di quest'ultime e, di conseguenza, evidenziare il ruolo che hanno nell'influenzare le condizioni del paziente alla diagnosi (stadio ed estensione del tumore, comorbidità) e le terapie a cui viene sottoposto (De Cuyper et al., 2009; Quaglia et al., 2011).

Il primo obiettivo secondario intende valutare se esistano differenze significative fra gli indicatori socio-economici a livello individuale e a livello d'area di residenza correlati alla sopravvivenza e come possano essere interpretate tali differenze.

Come abbiamo già visto nel capitolo precedente, infatti, la quasi totalità degli studi basati sulla combinazione di dati epidemiologici e socio-economici¹⁰ assume che le caratteristiche socio-economiche dell'area di residenza di un individuo siano simili a quelle dell'individuo stesso, generando così la possibilità di un *bias* ecologico (che si verifica quando le caratteristiche dell'area non descrivono adeguatamente l'individuo o descrivono aspetti non individuali dell'eventuale associazione fra sopravvivenza e caratteristiche socio-economiche) (Woods et al., 2005). Individuare differenze fra i due tipi di caratteristiche socio-economiche permetterà, quindi, di valutare l'esistenza effettiva di tale *bias* e di riflettere sulla necessità di estendere gli studi in materia anche

¹⁰ Ossia studi che, come il presente, prendono in considerazione dati provenienti da fonti esistenti certificate, come i gli uffici statistici degli enti pubblici e i registri di patologia.

a livello individuale, oltre a continuare a migliorare e approfondire gli studi a livello aggregato.

Infine, il secondo obiettivo secondario è legato alle procedure operative di *linkage* da utilizzare per ottenere i *dataset* necessari allo svolgimento delle analisi.

Tale obiettivo si propone di arrivare a costruire e validare procedure di *record linkage* fra le informazioni disponibili presso differenti enti territoriali, che abitualmente non fanno confluire i loro dati in unico database (nel nostro caso: il Registri Tumori regionale umbro, le Anagrafi comunali, l'Ufficio Statistica della Regione Umbria). Per realizzare questo studio è stato, infatti, necessario applicare specifiche procedure di raccolta e unione di tali dati.

Le caratteristiche ed il successo di queste procedure, di cui parleremo ampiamente nel capitolo 4 di questa tesi, permettono di fornire utili indicazioni su come potrebbero essere applicate in altre situazioni di ricerca omologhe, che abitualmente non ne farebbero uso ma che potrebbero trarne un rilevante valore aggiunto in termini di risultati più ricchi e più accuratamente interpretabili.

3.2 Le sedi tumorali e i soggetti inclusi nello studio

Per la realizzazione di questo studio sono stati scelti i tumori della mammella nelle donne e i tumori del colon-retto in entrambi i generi.

I criteri che hanno portato a questa scelta sono fondamentalmente tre e sono legati alla particolare rilevanza che queste neoplasie hanno all'interno dei principali sistemi sanitari pubblici europei, in particolare in quello italiano:

- sono i due tumori più diffusi nella popolazione italiana;
- entrambi sono caratterizzati da una buona probabilità di sopravvivenza e, di conseguenza, rappresentano due tipi di tumore particolarmente rilevanti in termini di gestione complessiva dei pazienti da parte del sistema sanitario;
- è stato verificato, da studi italiani e stranieri¹¹, che le caratteristiche socio-economiche dell'area di residenza dei pazienti sono associate a differenze significative nella sopravvivenza di questi due tumori.

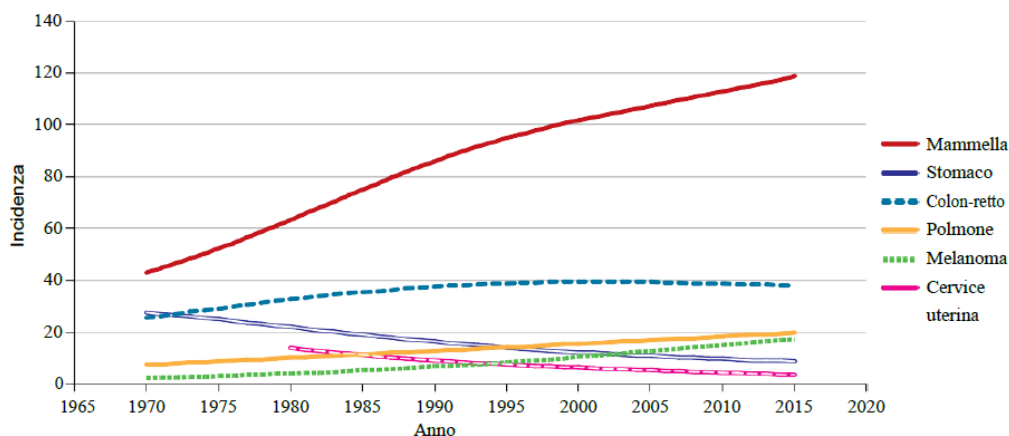
¹¹ Come descritto nel capitolo 2 di questa tesi.

Secondo i dati più recenti il tumore alla mammella è il cancro più diffuso nella popolazione femminile italiana, rappresentando il 29% di tutti i nuovi tumori diagnosticati ogni anno (AIRTUM/AIOM, 2014) e pari a circa 46.900 nuovi casi stimati (dati ITACAN 2013, AIRTUM, 2014).

In più, considerando il totale dei nuovi tumori annui per entrambi i generi assieme, il cancro alla mammella si pone al secondo posto con il 13% e, poiché il numero di tali neoplasie fra gli uomini è decisamente trascurabile (1.100 nuovi pazienti, ossia circa il 2%, al 2013), tale patologia è di particolare rilievo per valutare lo stato di salute della popolazione femminile.

L'andamento dell'incidenza di questo tumore nel tempo è marcatamente in crescita: si è infatti passati da un tasso di incidenza di 43 su 100.000 individui nel 1970 ad una stima al 2015 di 122 per 100.000 nel 2015 (Rossi et al., 2013), con un *trend* temporale in crescita costante (Graf. 3.1).

Graf. 3.1 – Tassi di incidenza per sede di tumore in Italia nel periodo 1970-2015. Età 0-99 anni. Donne.



Fonte: Rossi et al., 2013.

L'incremento dell'incidenza è particolarmente marcato ed è dovuto a molteplici fattori, fra cui è sicuramente importante l'invecchiamento della popolazione, fenomeno particolarmente rilevante in Italia e in buona parte del mondo occidentale. Va anche considerato che, come nella maggior parte dei paesi europei, dalla fine degli anni '90 per il tumore alla mammella sono attivi in buona parte delle regioni italiane programmi di *screening* organizzato che, grazie al miglioramento delle tecniche di diagnosi per

immagini, hanno sicuramente contribuito ad aumentare il numero di casi incidenti, diagnosticati anche in fase molto precoce (Paci et al., 2011).

Non tutto l'incremento, che era già in crescita marcata a partire dalla metà degli anni '70, può però essere spiegato con tali ragioni. Fattori di rischio legati a stili di vita non salubri (fumo, obesità), esposizioni ambientali/professionali e mutamenti nelle abitudini riproduttive sono stati identificati come elementi altrettanto rilevanti per l'incidenza (ad esempio, Dossus et al., 2014; McTiernan, 2003).

Questi dati sono confermati anche a livello europeo, dove il tumore alla mammella è di gran lunga il cancro più diffuso nel genere femminile, arrivando anche qui al 29% di tutte le neoplasie incidenti, per un numero stimato di nuovi casi annui pari a 458.718 nuove pazienti, e il più diffuso anche considerando entrambi i sessi, risultando il 13% di tutti i casi di tumore (WHO, 2015). Anche i *trend* temporali di incidenza si presentano crescenti come in Italia, benché con un incremento meno marcato.

I tumori del colon-retto, sempre secondo i dati più recenti, sono i più diffusi nella popolazione italiana, rappresentando il 14% di tutti i nuovi cancri diagnosticati ogni anno per entrambi i sessi, (AIRTUM/AIOM, 2014), il che vuol dire circa 54.600 nuovi casi stimati (dati ITACAN 2013, AIRTUM, 2014).

Essi sono i secondi per incidenza nelle donne (13% di tutti i tumori diagnosticati ogni anno), per un totale di circa 23.200 nuovi casi ma, come è evidenziato dal grafico 3.1, sembrano avere arrestato il proprio andamento crescente all'inizio degli anni 2000 ed essersi attestati attorno ad un tasso di incidenza di 40 su 100.000 individui.

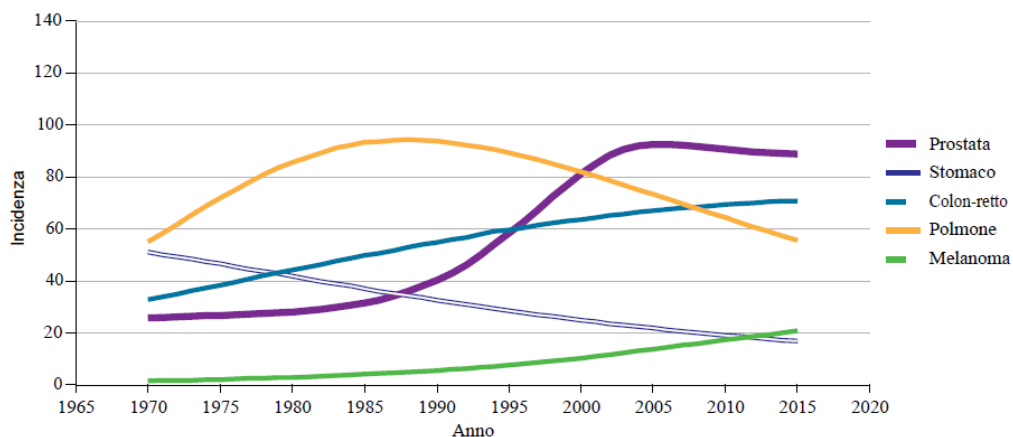
Per quanto riguarda la popolazione maschile la questione è più complessa. I due tumori più diffusi per incidenza sono quelli alla prostata (20%) e al polmone (15%), con le neoplasie del colon-retto al terzo posto (14%). Vi è però un problema relativo all'accuratezza statistica e diagnostica della misura dell'incidenza del cancro alla prostata, legato al ruolo del test del PSA che sembra individuare un eccessivo numero di casi fra falsi positivi e tumori quiescenti che, molto probabilmente, non avrebbero mai causato veri problemi alla persona (Quaglia et al., 2003; Rossi et al., 2013). E' quasi certo, quindi, un problema di sovrastima del numero dei nuovi casi, già da tempo soggetto ad un ampio dibattito nazionale ed internazionale (su questo tema si veda, ad esempio, Vismara Fugini et al., 2011; Canadian Agency for Drugs and Technologies in

Health, 2013). E' quindi difficile studiare in maniera attendibile il ruolo che le caratteristiche socio-economiche possono avere nella sopravvivenza degli uomini a cui è stata diagnosticata tale neoplasia.

Il tumore al polmone, invece, è per il genere maschile il secondo per incidenza e il primo per numero di decessi provocati (AIRTUM/AIOM, 2014); esso presenta una bassissima sopravvivenza legata alla mancanza di terapie efficaci, pertanto l'influenza delle disuguaglianze socio-economiche individuali sulla sopravvivenza stessa è praticamente nulla (Lillini et al., 2011) e non è quindi una sede neoplastica rilevante per questo specifico studio.

In più, se si osservano i *trend* temporali di incidenza per gli uomini (Rossi et al., 2013), si nota come il tumore al colon-retto sia quello che presenta un incremento continuo dal 1970 ad oggi (Graf. 3.2).

Graf. 3.2 – Tassi di incidenza per sede di tumore in Italia nel periodo 1970-2015. Età 0-99 anni. Uomini.



Fonte: Rossi et al., 2013.

Anche in Europa i cancri coloretali rivestono un ruolo importante sul totale dell'incidenza delle neoplasie: essi sono il secondo tumore nelle donne (13%), per un totale di 205.323 nuovi casi annui, e il terzo negli uomini (13%), per un totale di 241.813 nuovi casi annui (WHO, 2015). Considerando i due generi assieme, questa patologia è al secondo posto fra tutti i tumori (13%), per un totale di 447.136 nuovi casi

annui. Anche i *trend* temporali di incidenza presentano andamenti simili a quelli italiani.

Un ulteriore elemento di interesse per queste due patologie è che entrambi sono caratterizzate da un marcato incremento dell'incidenza al crescere dell'invecchiamento della popolazione (AIRTUM/AIOM, 2014) e, poiché tale processo è fortemente presente nella popolazione italiana, ecco che questi due tumori sono fra quelli destinati a rivestire un impatto crescente su tanti aspetti del sistema sanitario pubblico e sullo stato di salute della popolazione (Christensen et al., 2009; Crocetti et al., 2012; Vercelli et al., 2013).

A fronte di questi numeri, la sopravvivenza in Italia per queste due patologie è considerata buona quando viene messa a confronto con i risultati europei.

La sopravvivenza relativa a cinque anni dalla diagnosi per il tumore alla mammella nelle donne è, infatti, pari all'85% mentre per il colon-retto è pari al 60%. La sopravvivenza relativa a cinque anni dalla diagnosi per il colon-retto nei maschi è invece pari al 58% (dati ITACAN 2013, AIRTUM, 2014).

Nel confronto con l'Europa, la sopravvivenza per tumore alla mammella femminile in Italia si pone al quarto posto, dietro Francia, Finlandia e Svezia, e con valore superiori alla media europea (85% vs. 81%); per quanto riguarda il colon-retto, la sopravvivenza per donne e uomini è simile a quella dei paesi economicamente più rilevanti (i paesi nordici, la Germania, la Francia e la Svizzera) ed è superiore alla media europea (60% vs. 58% nelle donne e 58% vs. 56% negli uomini) (Allemani et al., 2015).

In più, la tendenza storica è di una sopravvivenza in crescita per entrambe le neoplasie; i dati AIRTUM mostrano infatti come la sopravvivenza sia passata da un valore del 78% all'inizio degli anni '90 ad un valore attuale dell'85% con *trend* tutt'ora crescente per la mammella nelle donne, e un andamento crescente simile per il colon-retto sia nei maschi che nelle femmine.

Questi risultati sono stati resi possibili da una serie di miglioramenti continui nelle tecniche diagnostiche (che permettono analisi sempre più precoci), nelle terapie (che in molti casi permettono approcci curativi sempre più efficaci con azioni sempre meno invasive) e nelle cure riabilitative post-terapia. Un ulteriore pratica che ha contribuito a

migliorare la sopravvivenza è l'introduzione nel sistema sanitario dello *screening* per migliorare l'individuazione precoce di tali neoplasie (ossia, la prevenzione secondaria). Allo *screening* sono poi state associate campagne di promozione di comportamenti virtuosi negli stili di vita, che possono contribuire a ridurre l'insorgenza stessa della malattia (ossia, la prevenzione primaria).

Malgrado tutto questo, la mortalità per questi tumori è particolarmente rilevante. Nel 2011, ad esempio, si sono registrati 19.077 decessi per neoplasie del colon-retto in entrambi i sessi e 11.959 decessi per neoplasia della mammella femminile, numeri che portano questi due tumori ad essere rispettivamente il primo per mortalità nelle donne (mammella) e il secondo in entrambi i sessi (AIRTUM/AIOM, 2014).

E' quindi importante acquisire più informazioni sul perché la mortalità resta così rilevante, malgrado le opzioni diagnostiche e terapeutiche abbiano permesso di raggiungere i buoni livelli di sopravvivenza sopra citati.

Come abbiamo già visto nel capitolo precedente, gli studi sull'argomento hanno ampiamente verificato come età e stadio della malattia alla diagnosi siano i due principali fattori che condizionano la sopravvivenza ai tumori per i quali esistono protocolli di prevenzione, diagnosi e cura efficaci, poiché possono condizionare le scelte di trattamento (ad esempio, Abadi et al., 2014; Dialla et al., 2015; Quaglia et al., 2011).

In più, la letteratura ha ampiamente verificato come le caratteristiche socio-economiche dell'area di residenza dei pazienti, associate ai pazienti stessi, influiscano sulla sopravvivenza, perché condizionano la possibilità di avere una diagnosi precoce e/o di seguire in maniera accurata le cure necessarie per superare la malattia. Di tutto ciò il capitolo precedente ha riportato numerosi esempi e qui si può ricordare il recentissimo studio di Woods et al. (2016), che ha dimostrato l'impatto della deprivazione socio-economica sull'adesione ai programmi di *screening* per il tumore della mammella e di conseguenza, sulla diagnosi precoce; oppure il lavoro di Le et al. (2008), che ha verificato come le differenze socio-economiche nei malati di cancro al colon-retto di differente provenienza etnica condizionano sia la tipologia delle cure che la sopravvivenza.

Nel presente studio si sono, di conseguenza, scelti i principali indicatori socio-economici individuali che, da letteratura, sono risultati rilevanti per descrivere le

relazioni di cui sopra, ossia la condizione familiare, espressa dalla variabile “Stato civile”, il livello di istruzione, espresso dalla variabile “Titolo di studio” e la situazione lavorativa, espressa dalla variabile “Condizione professionale”.

Similmente, basandosi su quanto è stato fatto in passato in letteratura, sono state considerate tutte le variabili che, nel Censimento 2001 della Popolazione e delle Abitazioni, riportano le informazioni socio-economiche di ogni sezione di Censimento relative a struttura familiare, livello educativo, lavoro, invecchiamento, supporto sociale, caratteristiche delle abitazioni.

I soggetti coinvolti nella ricerca sono tutti gli individui con diagnosi di tumore al colon-retto e tutte le donne con diagnosi di tumore alla mammella presi in carico dal sistema sanitario pubblico della regione Umbria dal 2001 al 2010, con *follow-up* al 31/12/2012, presenti nel Registro Tumori della regione Umbria. Si è scelto di focalizzarsi sugli anni 2001-2010, perché i dati raccolti in tale periodo sono maggiormente attendibili e verificabili, in particolare per quanto riguarda le informazioni socio-economiche e di posizionamento geografico. Non erano disponibili, nel periodo in cui è stato sviluppato questo studio, i dati di incidenza successivi al 31/12/2010.

Nel capitolo 4 verranno descritte più diffusamente le caratteristiche dei dati sui pazienti messi a disposizione dal Registro Tumori umbro. Sempre in tale capitolo saranno discusse le caratteristiche delle variabili socio-economiche individuali e di quelle relative all'area di residenza del paziente.

3.3 L'area geografica

Si è scelto di condurre lo studio sulla regione Umbria, per due ragioni:

- una ragione di tipo metodologico: tale regione presenta caratteristiche socio-economiche, condizioni di salute della popolazione e sviluppo del sistema sanitario simili a quelle riscontrabili a livello medio nazionale, mentre la popolazione tende ad essere leggermente più anziana della media nazionale (Vercelli et al., 2013). Questo punto è rilevante, non tanto perché sia obiettivo di questa tesi generalizzare direttamente i risultati all'intero territorio nazionale, ma perché si propone come elemento di confronto per futuri studi simili su altre aree, o su tutta l'Italia;

- una ragione di tipo pratico: il sistema informativo regionale mette a disposizione dati di elevata qualità per quanto riguarda gli aspetti sanitari, in particolare per quanto riguarda il Registro Tumori regionale. In più, l'esistenza di pratiche di raccolta dati condivise fra i diversi enti rende più agibile che altrove collegare informazioni provenienti da fonti amministrative differenti (nel presente caso da Registro Tumori, Anagrafi comunali e Servizio Statistico regionale).

Per quanto riguarda la ragione metodologica della scelta, esaminiamo alcuni elementi fondamentali che la giustificano.

Le caratteristiche demografiche e socio-economiche generali della popolazione regionale sono sufficientemente simili a quelle italiane. In particolare, per quanto riguarda quelle abitualmente associate in letteratura alle disuguaglianze nello stato di salute, l'Umbria si pone su posizioni comparabili con quelle del territorio nazionale. Ad esempio, nel 2011 le persone in possesso del solo titolo di scuola dell'obbligo sono il 77,1% in Umbria contro il 77,5% a livello nazionale, mentre i laureati sono il 12,4%, contro l'11,2% del dato nazionale (ISTAT, 2013).

Il confronto fra i dati umbri e quelli nazionali sull'occupazione e sulla struttura familiare rivelano una situazione analoga (Tab. 3.1 e Tab. 3.2).

Tab. 3.1 – Alcuni indicatori del mercato del lavoro e dell'occupazione: Umbria vs. Italia, dati 2011.

| Indicatori | Umbria | Italia |
|---------------------------|---------------|---------------|
| Occupati nell'agricoltura | 3,2% | 3,7% |
| Occupati nell'industria | 30,2% | 29,1% |
| Occupati nei servizi | 66,5% | 67,2% |
| Tasso di disoccupazione | 7,0% | 8,0% |

Fonte: ISTAT, 2013.

Tab. 3.2 – Caratteristiche dello stato civile e del nucleo familiare: Umbria vs. Italia, dati 2011.

| Indicatori | Umbria | Italia |
|--|---------------|---------------|
| Celibi/Nubili | 38,7% | 41,2% |
| Coniugati/conviventi | 51,0% | 49,3% |
| Separati/Divorziati | 8,4% | 7,5% |
| Vedovi | 1,9% | 2,0% |
| Numero medio di componenti il nucleo familiare | 2,4 | 2,4 |
| Famiglie di 1 componente | 30,6% | 29,4% |
| Famiglie di 2 componenti | 26,8% | 27,4% |
| Famiglie di 3 componenti | 21,6% | 20,4% |
| Famiglie di 4 componenti | 15,1% | 16,1% |
| Famiglie di 5 e più componenti | 5,9% | 6,7% |

Fonte: ISTAT, 2013.

La distribuzione delle età restituisce, invece, l'immagine di una popolazione regionale più anziana della media nazionale (23% di persone di 65 anni e più in Umbria contro il 21% nazionale nel 2011), ma con un andamento crescente perfettamente sovrapponibile all'andamento dell'invecchiamento in Italia (ISTAT, 2013).

Questo è un primo criterio di scelta rilevante per due ragioni: la popolazione europea, in generale, ed italiana, in particolare, sta subendo un progressivo processo di invecchiamento (Christensen et al., 2009; Lutz et al., 2008) che porterà a situazioni problematiche per i sistemi sanitari pubblici e per gli effetti di disuguaglianza che ne deriveranno.

In più, come già sottolineato nel paragrafo precedente, i tumori in generale e le due sedi scelte per questo studio in particolare crescono in incidenza al crescere dell'età, rendendo più complessa la situazione da gestire per il servizio sanitario, per le famiglie e gli individui.

Un altro criterio che ha guidato la scelta della regione Umbria è l'analogia fra la situazione sanitaria regionale e quella nazionale: sono stati infatti valutati alcuni

indicatori che permettessero di valutare se la regione scelta fosse sufficientemente simile all'intera nazione anche da questo punto di vista.

Ad esempio, nel 2011 la spesa sanitaria media pro-capite è stata di 1.798 euro all'anno, contro i 1.816 euro nazionali, per una percentuale rispetto al prodotto interno lordo del 9,3% contro l'8,9% italiano (dati 2011 da ISTAT, 2013); la spesa pubblica per la salute è simile: 7,6% in Umbria contro il 7,2% nazionale del prodotto interno lordo (dati 2011 da ISTAT, 2013).

Anche gli andamenti storici di tali indicatori per la regione Umbria sono molto simili a quelli calcolati per l'intera nazione (Graf. 3.3).

Inoltre, gli indicatori sullo stato di salute della popolazione sono molto simili fra Umbria e Italia nel suo complesso: ad esempio, nel 2011, il tasso di mortalità per tutti i tumori è pari al 30 per 10.000 vs 29 per 10.000 (dati 2011 da ISTAT, 2013), con andamenti nel tempo fra di loro sovrapponibili (ISTAT, 2013).

Anche i cosiddetti comportamenti a rischio della popolazione (ad esempio, l'abitudine al fumo o l'essere sovrappeso) sembrano distribuiti in maniera simile a quanto accade nella popolazione italiana, come si osserva nella tabella 3.3.

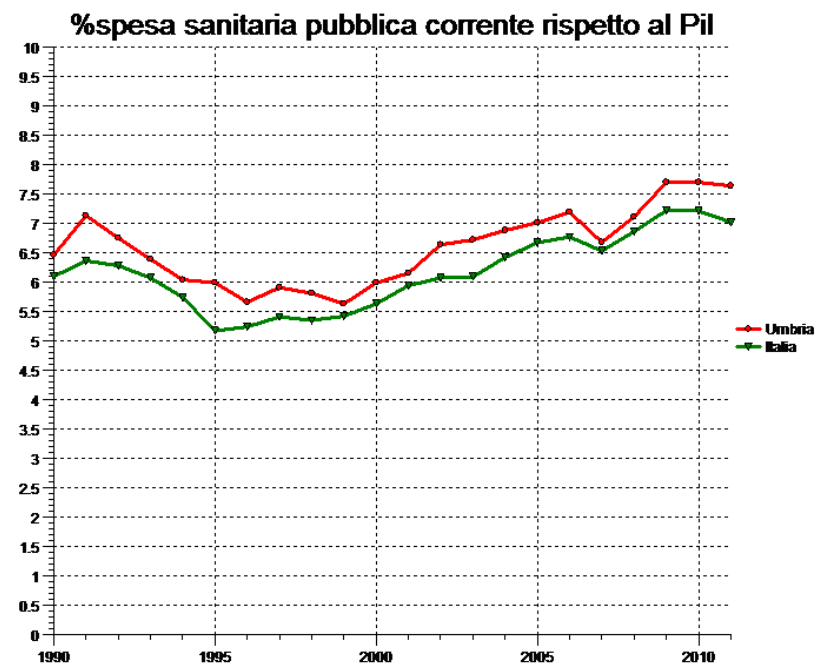
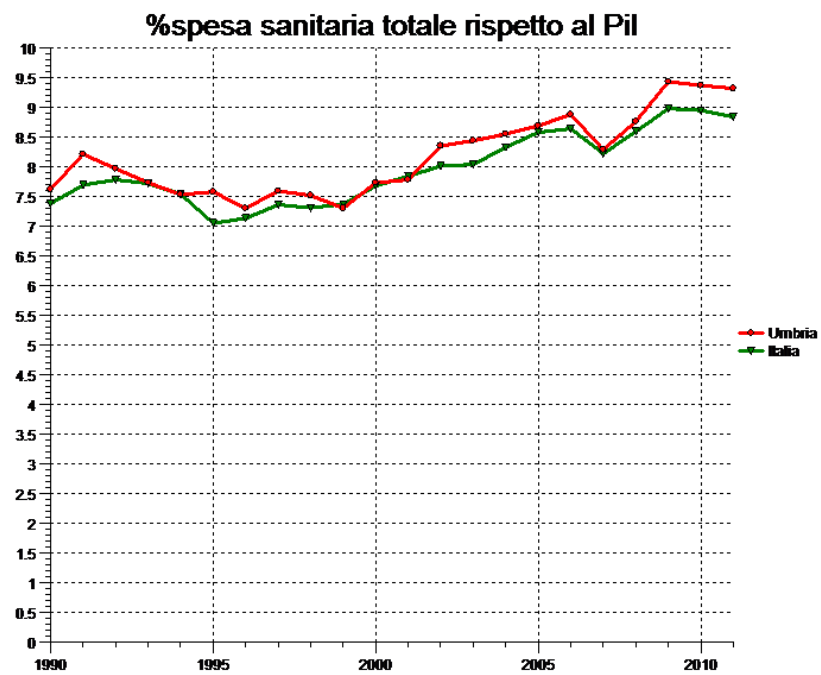
Tab. 3.3 – Indicatori di comportamento a rischio per la salute: Umbria vs. Italia, dati 2011.

| Comportamento a rischio | Umbria | Italia |
|------------------------------------|---------------|---------------|
| Fumatori (15 anni e più) | 21,8% | 22,5% |
| Grandi fumatori (15 anni e più) | 7,4% | 6,9% |
| Persone sovrappeso (18 anni e più) | 37,7% | 36,9% |
| Persone obese (18 anni e più) | 11,2% | 10,1% |

Fonte: ISTAT, 2013.

Sono stati scelti questi indicatori di comportamento a rischio per la salute come ulteriore criterio di scelta dell'area geografica da studiare, perché sono ritenuti essere fattori che aumentano significativamente il rischio di incidenza dei due tumori considerati in questo studio (ad esempio, Coups et al., 2007; Dossus et al., 2014; McTiernan, 2003).

Graf. 3.3 – Andamenti della spesa sanitaria totale e pubblica (% del prodotto interno lordo) - Umbria vs. Italia, dati 2011.



Fonte: ISTAT, 2013.

Se si prendono, infine, in considerazione i due tumori oggetto di questo studio, la sopravvivenza relativa standardizzata per età si pone ad un livello molto simile alla media nazionale senza differenze statisticamente significative (AIRTUM, 2014), come evidenziato nella tabella 3.4.

Tab. 3.4 – Sopravvivenze relative standardizzate percentuali a 5 anni dalla diagnosi: Umbria vs. Italia, dati relativi al quinquennio 2000-2004.

| Sedi | Sopravvivenza (%) e intervallo di confidenza (al 95%) | |
|-----------------------|--|---------------|
| | Umbria | Italia |
| Colon-retto – Maschi | 58 (56-61) | 58 (58-59) |
| Colon-retto – Femmine | 62 (59-64) | 60 (59-61) |
| Mammella – Femmine | 87 (85-89) | 85 (84-86) |

Fonte: database ITACAN dell'Associazione Italiana Registri Tumori (AIRTUM), 2014.

Per quanto riguarda la ragione pratica della scelta, l'Umbria presenta una situazione particolarmente favorevole per questo tipo di studi¹².

Il Sistema Informativo regionale, ossia l'insieme degli enti e soggetti pubblici che raccolgono, verificano, conservano ed elaborano i dati amministrativi su tutta la regione, è caratterizzato da un buon livello qualitativo, in particolare per quanto riguarda i dati del Registro Tumori regionale. In più, fra i diversi enti e l'Università degli Studi di Perugia esistono e sono operative convenzioni che permettono la ricerca e il controllo incrociato delle informazioni, aumentando così il livello qualitativo e la possibilità di realizzare studi che coinvolgano più fonti contemporaneamente.

Come vedremo dettagliatamente nel prossimo capitolo, queste caratteristiche regionali non solo hanno permesso la realizzazione di questo studio tramite la raccolta dati da più fonti, ma hanno anche garantito il successo delle procedure di *linkage* fra i dati provenienti da *database* amministrativi differenti, che solitamente non si “parlano” fra di loro, pur avendone la possibilità.

¹² Almeno nel periodo in cui è stato realizzato questo studio.

Capitolo 4

La costruzione del *dataset*

Al fine di indagare l'impatto delle caratteristiche socio-economiche degli individui sulla sopravvivenza ai tumori, è necessario disporre di un *dataset* che raccolga tali informazioni, possibilmente sia a livello individuale sia a livello di area di residenza, e contemporaneamente disponga di informazioni altamente dettagliate sulle caratteristiche della malattia (quando si è manifestata, come è stata diagnosticata e quali erano le sue caratteristiche alla diagnosi, come è stata curata e quale è stato l'esito finale, ecc.).

A conoscenza dell'autore, però, non esiste un simile *dataset*, né a livello nazionale, né a livello locale; nella prima fase di questo studio è stato quindi necessario creare tale *dataset*.

I dati utilizzati per la sua costruzione sono stati raccolti presso tre diverse fonti informative:

- il Registro Tumori umbro, che ha messo a disposizione le informazioni dettagliate di tipo anagrafico, patologico, diagnostico e terapeutico per tutti i pazienti del sistema sanitario regionale affetti da tumore alla mammella (solo donne) e da tumori al colon-retto (sia uomini, che donne);
- l'Ufficio Statistica della Regione Umbria, che ha fornito i dati relativi al Censimento Nazionale 2001 della Popolazione e delle Abitazioni, per tutte le sezioni di censimento umbre;
- le Anagrafi comunali umbre, che hanno fornito le informazioni individuali relative a stato civile, titolo di studio e condizione professionale dei pazienti.

L'obiettivo di questo capitolo è illustrare queste tre fonti, le loro caratteristiche istituzionali e funzionali e le caratteristiche dei dati che hanno fornito, descrivendo come si è costruito il *dataset* utilizzato nelle analisi con particolare attenzione alla procedura di *data linkage* utilizzata.

4.1 I Registri Tumori e il Registro Tumori umbro

I Registri Tumori sono strutture che si occupano di raccogliere, archiviare, analizzare e interpretare i dati patologici, diagnostici, clinici, demografici relativi a persone affette da tumore.

Attraverso una complessa procedura, che si basa sull'utilizzo di dati provenienti da diverse fonti (come, ad esempio, le cartelle ospedaliere, l'Anagrafe sanitaria, ecc.), essi raccolgono informazioni complete su tutti i nuovi casi di tumore e aggiornano continuamente le informazioni riguardanti i pazienti già inseriti nel proprio archivio, con particolare attenzione allo stato in vita di ogni individuo.

Più precisamente, i Registri raccolgono informazioni su tutti i tumori maligni insorti nella popolazione osservata seguendo, per ogni individuo malato, la sua evoluzione nel tempo. La casistica eleggibile è data da:

- tutti i tumori dal comportamento maligno;
- tutti i tumori (anche benigni e a comportamento incerto) delle sedi soggette a *screening*;
- i tumori a comportamento incerto, classificati come non sufficientemente definiti o rilevati solo tramite certificato di morte.

Infine, poiché i Registri operano su un'area territoriale ben definita, la residenza ufficiale al momento della prima diagnosi di un tumore è la condizione per l'inclusione del caso nell'archivio di un determinato Registro. Essa viene controllata presso le Anagrafi comunali o attraverso archivi specifici, come l'Anagrafe degli assistiti ASL, che attingono direttamente alle Anagrafi comunali stesse.

L'attività dei Registri è quindi fondamentale ai fini delle politiche di programmazione socio-sanitaria, sia a livello nazionale che locale. Infatti, i dati dei Registri forniscono elementi utili per:

- conoscere la dimensione reale e la distribuzione della patologia tumorale nell'area territoriale di competenza;
- identificare i fattori legati all'insorgenza della patologia oncologica (studi descrittivi spazio-temporali, studi eziologici di coorte e caso-controllo);
- valutare l'efficacia delle misure preventive e terapeutiche che sono state prese per contrastare la patologia oncologica (possibilità di misurare, in

termini di cambiamento della frequenza della patologia tumorale, gli interventi di prevenzione primaria o di prevenzione secondaria).

Malgrado la loro rilevanza informativa, però, in Italia non esiste una normativa nazionale unica che sta alla base della fondazione dei Registri. Invece, la loro costituzione è regolamentata a livello locale e la loro istituzione prevede solitamente una stretta collaborazione fra assessorati, strutture universitarie e/o ospedaliere.

Per capire il ruolo e l'importanza dei Registri Tumori per i sistemi sanitari e per il mondo della ricerca scientifica, è necessario avere presente un quadro sintetico di come sono nati e si sono sviluppati. Fino ai primi del '900 l'epidemiologia dei tumori era basata esclusivamente sulle statistiche di mortalità e solo nel 1936, nel tentativo di riflettere sul ruolo dell'incidenza, nacque il primo Registro Tumori nel Connecticut, seguito nel 1940 dalla regione del Saskatchewan in Canada, dalla Danimarca nel 1942 e, nel 1945, dalle regioni del sud-ovest in Gran Bretagna. Già dai primi lavori di queste realtà emersero affidabili statistiche di incidenza, ipotesi eziologiche (come l'inquinamento industriale) e i primi studi analitici. Tali risultati confermarono la fattibilità e l'utilità dei Registri Tumori.

Diversi studi di rilevanza storica (ad esempio Cornfield, 1951; Blalock, 1964; Clemmesen, 1965 e 1974) dimostrarono l'utilità dei dati raccolti dai Registri, fornendo per la prima volta informazioni precise e dettagliate sulle patologie tumorali. Questo li portò ad essere considerati, quindi, come uno strumento valido e attendibile di quantificazione dei problemi della sanità pubblica, che proprio in quegli anni disegnava i propri servizi sanitari (in Nord Europa) o modernizzava i propri organismi di programmazione e valutazione federale (in Nord America) (Zanetti, 2007).

Proprio per questo motivo, i Registri Tumori passarono negli anni '70 e '80 da poche unità ad alcune decine, anche se limitatamente ad aree geografiche e nazioni più ricche (l'istituzione e il mantenimento di un Registro Tumori è un costo rilevante per il Sistema Sanitario).

Essi iniziarono anche ad essere coordinati dalle istituzioni e organizzazioni scientifiche che si occupavano (e si occupano) dello studio e della valutazione della malattia neoplastica¹³.

Nella storia della medicina questo è un passaggio unico, perché le specializzazioni che si occupano di altre patologie importanti, come le malattie cardiovascolari o metaboliche, non hanno conosciuto simili dinamiche di costruzione di validi e duraturi sistemi informativi così estesi.

In Italia, il primo Registro Tumori si è costituito a Varese nel 1976, seguito dopo pochi mesi dal Registro di Parma, entrambi operanti a livello provinciale. Da allora si è assistito ad un incremento costante nel numero dei Registri, fino ad arrivare a 40 Registri, distribuiti su tutto il territorio nazionale e operanti principalmente a livello provinciale. Solo Friuli Venezia-Giulia, Umbria e Veneto hanno completa copertura regionale, mentre Torino, Milano e Napoli operano a livello comunale.

Allo stato attuale, si può quindi osservare una copertura ampia e diffusa a livello italiano, europeo e mondiale della popolazione da parte dei Registri Tumori, organizzati in associazioni scientifiche nazionali e internazionali che certificano e garantiscono i criteri e i parametri di qualità per la raccolta, la conservazione, l'elaborazione e la diffusione di dati.

Per quanto riguarda l'Italia, i Registri Tumori sono organizzati nell'Associazione Italiana dei Registri Tumori (AIRTUM) e raggiungono una copertura complessiva della popolazione osservata pari al 52% circa dell'intera popolazione italiana (dato AIRTUM, 2015), a cui si può aggiungere un ulteriore 22% circa legato a Registri Tumori di recente attivazione e di cui si sta attualmente valutando la qualità operativa.

L'adesione di un registro all'Associazione è particolarmente rilevante, perché permette di partecipare al risultato più importante conseguito da AIRTUM: una banca dati a cui affluiscono tutti i dati raccolti dai Registri Tumori (sia di popolazione che specialistici) che afferiscono all'Associazione¹⁴.

¹³ Ad esempio, il Surveillance, Epidemiology, and End Results Program (SEER) statunitense o l'International Agency for Research on Cancer (IARC) internazionale.

¹⁴ La banca dati AIRTUM è il grande archivio nazionale, ospitato presso l'Istituto per lo Studio e la Prevenzione Oncologica di Firenze, in cui confluiscono tutti i dati sui tumori raccolti dai singoli registri accreditati. Attualmente vi sono depositate informazioni su oltre due milioni di malati di tumore (2.324.564 tumori maligni diagnosticati nel periodo 1967-2011) e oltre un milione di deceduti per tumore (1.071.414 decessi avvenuti nel periodo 1967-2011) (<http://www.registri-tumori.it/cms/it/bancadati>).

Questa banca dati costituisce la più ricca e verificata fonte di informazioni epidemiologiche sui tumori a livello italiano, contribuisce alle banche dati internazionali ed è uno strumento preziosissimo per finalità di ricerca, sorveglianza sanitaria, ecc.

L'adesione all'AIRTUM da parte di un Registro comporta che tutte le attività di raccolta, controllo e archiviazione dei dati seguano procedure standardizzate piuttosto impegnative da attivare, specialmente nella fase di fondazione del Registro¹⁵, che permettono:

- di verificare direttamente o indirettamente l'esaustività della raccolta dei casi;
- di raccogliere per ciascun paziente (caso) un ampio insieme predefinito di informazioni al miglior livello qualitativo possibile;
- di garantire la standardizzazione dei metodi di raccolta e validazione dei dati (e conseguentemente dei risultati) tra i diversi Registri e nel corso del tempo.

4.1.1 Funzionamento dei Registri Tumori

Per spiegare le caratteristiche e il funzionamento di un Registro Tumori può essere utile descrivere cronologicamente le varie scelte a disposizione e le azioni che devono essere poste in atto dal momento in cui ne viene proposta la fondazione fino all'inizio e al mantenimento della sua attività quotidiana.

La prima fase è la scelta del tipo di Registro. Esistono, infatti, fondamentalmente due tipi di Registri Tumori:

- i Registri di popolazione, che estendono la loro attività a tutta la popolazione di una definita area e a tutte le malattie oncologiche;
- i Registri specializzati, che considerano una fascia di popolazione (Registri infantili) o una malattia specifica (Registri specializzati d'organo o di patologia).

Per entrambi i tipi, l'attività di registrazione si fonda su un processo in cui vengono consultate tutte le fonti disponibili per reperire nuovi casi e classificarli al meglio.

Questa prima considerazione sulle caratteristiche della raccolta dati porta direttamente al secondo tipo di scelta, ossia il modello operativo.

I modelli operativi utilizzati abitualmente sono due (Ferretti et al., 2007):

¹⁵ D'altronde queste procedure sono necessarie per garantire la qualità dell'informazione di un Registro Tumori, che viene comunemente considerata come un "gold standard" per chi opera in area sanitaria.

- registro manuale o semiautomatico; dopo una prima fase di acquisizione dei casi da una o più fonti di dati, si procede al controllo delle segnalazioni multiple; le fasi di acquisizione e controllo dei dati si effettuano per lo più tramite l'attività manuale degli operatori, ma a volte si utilizzano anche procedure di *record linkage* parzialmente informatizzate;
- registro automatizzato; già dalla prima fase di acquisizione i dati sono processati automaticamente con *record linkage* automatico che segnala casi dubbi o discrepanti sui quali condurre successivamente verifiche sia automatiche che manuali.

In genere, i Registri automatizzati operano su bacini di utenza regionali per i quali le procedure manuali rappresenterebbero un carico di lavoro non gestibile.

Una volta definiti questi punti, è necessario inquadrare il Registro Tumori come un organo del flusso informativo (ossia dell'insieme di enti e istituzioni che, a diverso livello territoriale - nazionale, regionale, provinciale, comunale - producono, verificano e distribuiscono dati e informazioni) in cui verrà inserito. Vanno ossia coinvolte anche a livello normativo le istituzioni locali a cui esso farà riferimento.

Questo passaggio è fondamentale, perché di fatto permette al Registro di passare alle fasi successive, fino ad iniziare l'attività di raccolta dati e riguarda sia i Registri manuali (e semiautomatici), sia i Registri automatici.

Più precisamente, si attua una pianificazione del flusso informativo del Registro che l'ente o gli enti locali (Regione, ASL, Università, ecc.) dovranno giudicare e approvare, anche da un punto di vista normativo. Per raggiungere questo obiettivo, vengono valutate le modalità di acquisizione e gestione dei dati tramite uno studio di fattibilità che analizza approfonditamente (Ferretti et al., 2007):

- le motivazioni all'istituzione di un nuovo Registro Tumori, la sua contestualizzazione e la sua valenza all'interno del territorio;
- la popolazione interessata in termini di dimensioni e caratterizzazione socio-demografica;
- il ritorno informativo atteso e il suo contributo nei confronti del sistema sanitario.

E' fondamentale che vengano anche analizzate le risorse disponibili in termini di:

- effettiva disponibilità delle fonti dei dati;
- struttura operativa di acquisizione, registrazione, codifica e analisi dei dati;
- adeguatezza dello strumento informatico per la registrazione dei casi;
- archivi sanitari consultabili su carta e/o con accesso informatico;
- rapporti con anagrafi comunali ed eventuale possibilità di accesso informatico;
- quantificazione del costo reale (in termini di attività lavorativa, costi finanziari e tempi di acquisizione dei dati) di ogni metodo di registrazione e di ogni fonte informativa;
- qualità e livello di completezza dei dati che possono essere garantite;
- tempi previsti per l'attivazione delle procedure di accreditamento, ove previsto, e per la messa a disposizione dei dati alla banca;
- sistemi di analisi e di controllo qualità previsti e relativa periodicità.

Terminato lo studio di fattibilità e la valutazione di inserimento del Registro nel flusso informativo, la costituzione del Registro deve essere recepita e approvata dall'amministrazione a cui il Registro stesso afferisce (ad esempio, la regione o l'università).

L'atto di costituzione deve essere recepito dall'amministrazione regionale ai fini del riconoscimento giuridico, oltre che per consentire la strutturazione del flusso informativo e la stabilità nel tempo del gruppo di lavoro.

Prima dell'avvio della fase operativa, ogni Registro deve produrre un quadro delle risorse umane necessarie, un organigramma e un protocollo operativo adattato alle specifiche esigenze locali.

Stabilito il quadro generale delle risorse umane necessarie al funzionamento, viene strutturato l'organigramma. Esso prevede una figura centrale di coordinamento e supervisione (direttore del Registro) che rappresenta il Registro nei confronti degli altri enti ed istituzioni ed è giuridicamente responsabile della garanzia di confidenzialità e protezione del dato, problema estremamente delicato viste le complicazioni legate alla normativa su *privacy* e sensibilità del dato, e definisce il protocollo operativo.

Al direttore del Registro riporta il gruppo di registrazione, composto da personale tecnico, che ha il compito di gestire la raccolta delle informazioni e, assieme al direttore, è l'unico ad avere l'autorizzazione a trattare i dati personali degli individui.

Il personale tecnico può essere affiancato da esperti esterni (clinici, patologi, informatici, statistici, epidemiologi) in base alle necessità specifiche.

Dopo la definizione e l'accettazione dell'organigramma e l'attivazione della figura del direttore, viene costruito il protocollo operativo, ossia il documento ufficiale di funzionamento del Registro che definisce la maggior parte degli aspetti fondamentali delle attività del Registro stesso, quali, ad esempio, il tracciato del *dataset* (quali sono le variabili obbligatorie e le variabili facoltative), le modalità di controllo di qualità e di revisione della casistica, le modalità di conservazione degli archivi informatici e cartacei (comprese le procedure di *backup* e di sicurezza della struttura), ecc.

In questa fase di sviluppo di un Registro Tumori, a fianco del protocollo operativo viene anche progettata la relazione di scambio informativo con i Registri di Mortalità, siano essi locali (più precisamente regionali, quando sono presenti nell'area in cui opera il Registro) o nazionale (ossia gestito da ISTAT).

Nel momento in cui inizia la propria attività, ogni Registro aderisce a convenzioni circa gli aspetti fondamentali della registrazione dei tumori stabilite a livello internazionale dall'International Agency for Research on Cancer (IARC), dall'International Association of Cancer Registries (IACR) e dallo European Network of Cancer Registries (ENCR) (Jensen et al., 1991; Esteban et al., 1995; Parkin et al., 1995; ENCR, 1995).

Ogni Registro è tenuto ad allinearsi a tali regole tecniche, ma può sviluppare eventuali ulteriori regole interne in relazione a peculiarità del flusso informativo o a particolari necessità locali. In questo ambito vanno ricondotte, per esempio, le attività di registrazione di variabili o informazioni non previste a priori nel *dataset* della Banca Dati nazionale AIRTUM, ma ritenute necessarie per quelle attività di supporto all'attività epidemiologica e di ricerca clinica che sempre di più coinvolgono i Registri (valutazione degli *screening*, *follow-up*, verifica delle familiarità in genetica oncologica, caratteristiche socio-economiche dei pazienti, ecc.) (Krieger, 2001; Glaser et al., 2005; Goodman et al., 2005).

4.1.2 I dati e le fonti dei dati dei Registri Tumori

Per poter svolgere la propria attività, i Registri Tumori si appoggiano a tutta una serie di fonti dati ufficiali inserite nei flussi informativi locali e nazionali.

Esse vengono consultate in diversa maniera, a seconda del grado di informatizzazione che presentano. Vale a dire che, a seconda dei casi specifici, le informazioni da reperire presso le diverse fonti possono essere recuperate in maniera automatica o semi-automatica (consultazione degli archivi direttamente *on line*, trasmissione da parte degli archivi esterni dei dati in formato elettronico, ecc.), oppure tramite consultazione diretta di materiale cartaceo (conservato presso gli archivi delle fonti o trasmesso alla sede del Registro).

Le fonti dati utilizzate dai Registri sono specificate nella fase di pianificazione del flusso informativo e nello studio di fattibilità; esse sono principalmente di tre tipi (Ferretti et al., 2007):

- fonti anagrafiche, ossia Anagrafi comunali e Anagrafe sanitaria;
- fonti sanitarie primarie (diagnostiche, cliniche, ecc.), come ad esempio le schede di dimissione ospedaliera e gli archivi di anatomia patologica;
- fonti sanitarie secondarie (o complementari) rispetto alle precedenti, come ad esempio i dati relativi agli *screening* o alle esenzioni da *ticket* per ragioni oncologiche.

I dati anagrafici essenziali che i Registri devono archiviare per ciascun caso sono: cognome, nome, sesso, data di nascita, comune di nascita, comune di residenza, indirizzo civico di residenza, data di morte, luogo di morte, luogo di emigrazione (variabile facoltativa, ma raccomandata per il *follow-up*), data di emigrazione (variabile facoltativa, ma raccomandata per il *follow-up*) e codice fiscale. Quest'ultimo dato è fondamentale, perché è la chiave primaria attorno a cui ruotano tutte le successive operazioni di *linkage*.

Le fonti anagrafiche di riferimento sono principalmente le Anagrafi dei comuni appartenenti al territorio coperto dall'attività di registrazione. La fonte anagrafica permette di stabilire le caratteristiche dell'identità di ogni caso considerato, nonché la

sua appartenenza territoriale che, si ricorda, lo rende o meno eleggibile per la registrazione nel momento in cui si rileva l'insorgere del tumore. In più i dati delle Anagrafi comunali permettono di minimizzare le copie e le duplicazioni di dati, che sono la causa della maggior parte degli errori anagrafici.

Accanto alle Anagrafi comunali i Registri utilizzano, come fonte di supporto, l'Anagrafe sanitaria dell'ASL del territorio di riferimento. Essa risulta utile in caso di cessazione dell'assistenza (di cui fornisce le motivazioni) o di erogazione dell'esenzione dal *ticket* sanitario (per i casi a non sufficiente evidenza di malattia), permettendo quindi di supportare il *follow-up* di ogni caso.

Una volta definite le fonti di tipo anagrafico, un Registro Tumori deve acquisire le fonti primarie di tipo clinico-sanitario, indispensabili per la raccolta delle informazioni sulla patologia. Queste fonti possono essere collegate tra loro grazie al fatto che ognuna di loro riporta almeno una parte delle informazioni raccolte presso le fonti anagrafiche (codice fiscale e/o nome e cognome e/o indirizzo di residenza, ecc.). I dati anagrafici risultano essere così il cardine attorno a cui ruota la possibilità di costruire *record* di informazioni complete per ogni caso registrato.

In genere, le fonti primarie clinico-sanitarie possono essere riassunte in:

- schede di dimissione ospedaliera,
- cartelle cliniche,
- archivi di anatomia, istologia e citologia patologica,
- Registri di Mortalità.

La scheda di dimissione ospedaliera (SDO) raccoglie in sintesi le informazioni principali su ciascun ricovero di un paziente. Essa è articolata in tre parti: una relativa all'identificazione del paziente stesso e alla sua posizione all'interno del Servizio Sanitario Nazionale, una riguardante il ricovero (reparto e data di ammissione, trasferimenti, tipo e reparto di dimissione) e una corrispondente alle patologie diagnosticate e trattate. Questo flusso informativo, automatizzato a livello regionale a partire dagli anni '90, rappresenta una fonte primaria di acquisizione dei dati, generalmente di facile accessibilità e in grado di consentire una buona ricostruzione del dato storico di ogni paziente, anche proveniente da ricoveri fuori regione. L'archivio

SDO è una fonte primaria per i registri automatizzati, ed è essenziale per i registri manuali e semiautomatici per la ricerca delle cartelle cliniche da consultare.

Una seconda fonte informativa clinico-sanitaria è costituita dalle cartelle cliniche. Essa è fondamentale sia per il reclutamento dei casi/pazienti sia per il controllo dei dati, soprattutto per ciò che concerne la definizione precisa della data di incidenza e per la disponibilità dell'anamnesi patologica remota (tumori multipli, precedenti interventi, ecc.). Quando l'anamnesi riporta informazioni nosologiche complete (data e luogo di diagnosi e trattamento, sede e tipo istologico, trattamenti eseguiti), esse sono da utilizzare per la registrazione del caso e della data di incidenza.

Ulteriori fonti informative primarie sono gli archivi di anatomia, istologia e citologia patologica di ospedali e ASL, in quanto forniscono un'informazione insostituibile sia per quanto concerne le caratteristiche e l'estensione di malattia (stadio), sia per quanto riguarda la data di incidenza della patologia neoplastica. La fonte anatomo-patologica è importante, perché rappresenta l'unica risorsa per il recupero dei dati autoptici in grado di confermare un sospetto diagnostico non risolto in vita o di segnalare neoplasie senza evidenza clinica.

L'ultima fonte primaria è rappresentata dai Registri di Mortalità, attraverso i quali le Aziende sanitarie locali custodiscono i dati contenuti nelle schede ISTAT di morte.

La disponibilità dei dati di mortalità è, infatti, fondamentale per ogni Registro in quanto, integrata con i dati di incidenza, permette:

- di completare il *follow-up* sullo stato in vita dei pazienti registrati nel *database*;
- di verificare luogo e causa di morte negli eventi di decesso, specialmente per i casi con informazioni non particolarmente esaustive;
- di rintracciare casi precedentemente sfuggiti alla rilevazione di incidenza e di ripercorrerne la storia (anche utilizzando le eventuali indicazioni sulla durata della malattia), stimando queste perdite attraverso gli indicatori forniti dai certificati di morte;
- di correggere eventuali errori anagrafici;
- di effettuare misure e stime di sopravvivenza e prevalenza.

Esiste poi tutta una serie di fonti informative sanitarie secondarie (o complementari) che possono fare parte dei flussi del Registro, sia in maniera continuativa sia limitata alla risoluzione di particolari bisogni o problemi informativi.

In estrema sintesi, esse sono (Ferretti et al., 2007):

- Registri tumori infantili (qualora operativi sullo stesso territorio);
- Registri tumori ospedalieri e specifici d'organo (qualora presenti sullo stesso territorio);
- Registri regionali dei mesoteliomi;
- Registri nazionali di patologia;
- commissioni regionali per l'autorizzazione del rimborso delle cure all'estero;
- unità operative di cure palliative, *hospice*, assistenza domiciliare integrata;
- servizi di radioterapia;
- servizi di diagnostica per immagini;
- servizi di laboratorio analisi;
- servizi di oncologia;
- centri *screening*;
- archivi di esenzione *ticket* per pazienti affetti da patologia oncologica;
- medici di medicina generale e pediatri di libera scelta;
- *network* degli oncologi pediatrici.

Per abbinare le informazioni provenienti dalle diverse fonti si utilizzano procedure di *record linkage* deterministico, che prevedono diverse chiavi di *link* gerarchiche (ad esempio, codice fiscale, cognome, nome, sesso e data di nascita).

La chiave primaria utilizzata è il codice fiscale, seguita da combinazioni di altre variabili chiave, come cognome e nome, data di nascita, indirizzo di residenza, codice dell'Anagrafe sanitaria.

Prima di utilizzare queste combinazioni, vanno però eliminati possibili elementi di ambiguità (come spazi, apostrofi, accenti fra cognome e nome) e valutare la presenza nelle diverse fonti di eventuali errori. Esistono infatti tutta una serie di errori specifici, in genere legati alle procedure di *data entry*, che tendono a ripetersi nel tempo.

Il tipo specifico di errore in determinate variabili chiave può, infatti, causare problematiche specifiche di sfasamento e mancato abbinamento fra dati. Gli errori più comuni meritano un breve approfondimento, perché sono stati presi in considerazione anche nella fase di *record linkage* e controllo di qualità dei dati di questo studio.

1) Errore nel Cognome. Può essere dovuto alla presenza di preposizioni, accenti, apostrofi, doppio cognome, cognome straniero, inserimento del cognome da coniugata. In generale il *linkage* funziona se si utilizza il codice fiscale, che riduce questo tipo di errore. Tuttavia non sempre tutte le fonti contengono il codice fiscale. In questo caso, occorre comunque correggere il cognome per arrivare ad una forma unica e condivisa fra le fonti, pena la perdita di informazioni nel *linkage*.

2) Errore nel Nome. Un tipo di errore è connesso alla sussistenza o all'assenza di un doppio nome; le regole dell'anagrafe di stato civile sono infatti mutate nello scorso decennio, ed escludono il doppio nome se vi è la separazione con una virgola, ma i pazienti tendono a registrarsi secondo loro consuetudine. In altri casi è la dicitura congiunta o disgiunta a trarre in inganno (ad esempio, Gian Franco, Gian-Franco, Gianfranco). Altri tipi di errore riguardano spesso i nomi stranieri (errori ortografici) e nomi usati sia al maschile sia al femminile (Andrea) o con diversa accentazione (Élia, Elia) che possono causare erronee attribuzioni di genere.

3) Errore nella Data di nascita. Sono possibili errori materiali in fase di immissione del dato, specialmente in caso di copiatura (errori di battitura). Ulteriori problemi dipendono dal formato della data (italiana, anglosassone) adottata da sistemi diversi (quando gli *output* di un programma sono utilizzati come *input* a cascata della fase successiva) e, frequentemente, dalla diversa formattazione della data (riguardo le cifre del secolo) che rendono i passaggi tra i vari *software* di archiviazione delle informazioni ad alto rischio di errore (con generazione di casi mancanti).

4) Errore nell'Età. Alle difficoltà che ovviamente possono generarsi dai problemi delle date si somma, nel calcolo automatico delle età alla diagnosi, l'effetto dell'arrotondamento, che può provocare disomogeneità delle classi di età calcolate in situazioni diverse (di norma è preferibile il troncamento all'età in anni compiuti).

5) Errore nel Luogo di nascita. E' il dato che insieme a cognome, nome, sesso e data di nascita costituisce l'insieme minimo di informazioni per l'identificazione anagrafica. Alcuni problemi possono derivare dai sistemi automatici di codifica del comune (come

avviene per le schede di dimissione ospedaliera). Le anomalie sono in genere legate a comuni che sono entrati a far parte di una nuova, o di altra, provincia, o comuni che sono stati soppressi definitivamente o temporaneamente nei decenni passati (unificati con altri e poi ricostituiti, o che hanno cambiato nome).

6) Errore nel Comune di residenza. Accanto alle problematiche esaminate al punto precedente, non sono infrequenti altri tre tipi di errore: attribuzione come residenza del capoluogo di provincia, errata classificazione per incompletezza della denominazione del comune (ad esempio, Villanova d’Asti attribuito a Villanova Biellese), attribuzione come comune di residenza del luogo di domicilio o di una frazione del comune (generalmente in documentazioni cartacee).

7) Errore nel Codice fiscale. E’ proporzionalmente più raro degli altri errori e, in genere, legato al *data entry*. Può essere corretto per interpolazione con i dati del Ministero delle Finanze.

Per ovviare al problema dei mancati abbinamenti, le strategie utilizzate consistono principalmente nella revisione delle chiavi di *linkage* per ridurre il numero di errori, utilizzando ulteriormente sia procedure proprie del *record linkage* di tipo deterministico che, più raramente, di tipo probabilistico (Dal Maso et al., 2001).

4.1.3 La qualità del dato

In precedenza si è detto che i Registri Tumori producono dati che possono essere considerati un “*gold standard*” per l’alta qualità che esprimono rispetto alla realtà che permettono di osservare.

Questa qualità viene raggiunta grazie al controllo di quattro criteri (Bray et al., 2009; Parkin et al., 2009): comparabilità, completezza, validità e tempestività dei dati.

1) La comparabilità è il criterio che valuta se e quanto le procedure di codifica e classificazione di un Registro sono coerenti con le linee guida nazionali e internazionali. Requisiti fondamentali per la comparabilità sono la standardizzazione delle pratiche relative alla classificazione e codifica di nuovi casi, nonché la coerenza delle definizioni di incidenza. Questi requisiti sono analizzati tramite quattro parametri: il sistema utilizzato per la classificazione e la codifica delle neoplasie, la definizione di incidenza

e la definizione della data di incidenza, la distinzione tra tumore primario (nuovo caso) e recidiva o metastasi di uno già registrato, la ricodifica di tumori rilevati in individui asintomatici.

La comparabilità dei dati è quindi misurata tramite un esame approfondito delle *routine* di registrazione, con indicazioni particolareggiate circa le norme e le definizioni che sono state seguite, in accordo con i parametri di cui sopra.

2) La completezza è la misura in cui tutti i tumori incidenti che si verificano nella popolazione sono inclusi nel *database* del Registro. I tassi di incidenza e di sopravvivenza saranno tanto più vicini al loro valore reale, quanto più le procedure di reperimento dei casi potranno garantire la massima completezza. Trattandosi di un criterio fondamentale, numerosi sono i metodi che vengono utilizzati per valutarlo. In sintesi, vengono utilizzati metodi semi-quantitativi e quantitativi.

I metodi semi-quantitativi forniscono un'indicazione generale della completezza di un Registro, ma non sono in grado di quantificare i casi mancanti. Sono: la valutazione della stabilità dell'incidenza nel tempo fra popolazioni, la forma delle curve di incidenza età-specifiche, l'andamento della mortalità, le caratteristiche e le modalità operative delle fonti che forniscono le informazioni di ogni caso, la verifica istologica della diagnosi.

I metodi quantitativi permettono di ottenere una valutazione numerica del livello di completezza. Sono: l'accertamento dei casi di Registro tramite una ulteriore comparazione delle fonti (per rintracciare eventuali casi andati perduti), il ripetere le procedure di *linkage*, la verifica tramite i certificati di morte. La scelta dei metodi di controllo dipende dalla disponibilità e dalle caratteristiche delle fonti, dalla disponibilità di personale e tempo, dalla sostenibilità dei costi.

3) La validità (o accuratezza) si riferisce alle caratteristiche dei casi, come le caratteristiche biologiche della malattia, lo stadio alla diagnosi, ecc. Essa misura la percentuale di casi le cui caratteristiche sono reali e non attribuite per errore nel corso delle procedure di registrazione. Ad esempio, viene calcolato per quanti casi la diagnosi è corretta, lo stadio del tumore all'incidenza è corretto, ecc.

La validità dipende dalla precisione delle fonti informative iniziali e dal livello di competenza di chi si occupa delle procedure di codifica e ricodifica. Parkin et al. (1995) descrivono quattro metodi che forniscono indici numerici di validità e quindi permettono confronti con altri Registri o, all'interno di uno stesso registro, confronti temporali e/o fra sottoinsiemi di casi.

Essi sono: ripetizione della valutazione e ricodifica, valutazione dei criteri diagnostici (verifica istologica, certificato di morte), analisi delle informazioni mancanti, coerenza interna dei metodi di registrazione.

Per questo criterio sono stati definiti, a livello nazionale e internazionale, valori di soglia (periodicamente aggiornati) il cui superamento conferma il livello di validità dei dati di un Registro.

4) Per quanto riguarda il criterio della tempestività, va detto che non esiste una definizione formale, sebbene in termini generali essa sia definita come la rapidità con cui un Registro può raccogliere, elaborare e comunicare dati sufficientemente affidabili e completi.

Smith-Gagen et al. (2005) hanno definito un indicatore temporale di disponibilità del dato come l'intervallo tra la data della diagnosi (o la data di incidenza) e la data in cui il caso diventa disponibile per finalità di ricerca. Per molti registri, quest'ultima data coincide con quella presa come riferimento per calcolare le statistiche annuali per un rapporto ufficiale.

Questo indicatore misura, di fatto, la rapidità con cui il Registro completa le operazioni di aggiornamento del proprio *database*.

4.1.4 Il Registro Tumori umbro

Il Registro Tumori umbro è operativo dal 1994 ed è stato il primo Registro Tumori italiano a garantire una copertura regionale; dal 1995 il Registro afferisce all'AIRTUM. Esso svolge tutte le attività di documentazione e ricerca tradizionali e partecipa alle attività di ricerca delle principali associazioni nazionali e internazionali citate in precedenza.

A partire dalla sua istituzione, il Registro è sempre stato gestito dal Dipartimento di Specialità Medico-Chirurgiche e Sanità Pubblica, Sezione di Sanità Pubblica,

dell'Università di Perugia, che fornisce gli spazi, il personale e le apparecchiature, ed è sempre stato finanziato dal Servizio Sanitario regionale.

Il Registro si caratterizza per livelli di comparabilità, completezza e validità fra i più alti in Italia (AIRTUM/AIOM, 2014); rispetto agli altri registri, si contraddistingue per la tempestività e la qualità del dato prodotto (ad oggi, il Registro ha completato il *follow-up* di tutti i casi al 31/12/2013, risultato assolutamente di rilievo nel panorama nazionale).

Questi risultati sono stati resi possibili grazie al lavoro degli ingegneri afferenti al registro che hanno sviluppato un'innovativa architettura informatica (Stracci et al., 2010) che ha permesso una perfetta integrazione operativa nella Rete Oncologica regionale.

Tale sistema permette:

- l'ottimizzazione dell'integrazione fra fonti anagrafiche e fonti sanitarie primarie e secondarie, rendendo più semplice e immediata l'immissione dei dati da parte di tutti gli utenti abilitati (ad esempio, gli operatori del Registro, oppure gli impiegati delle diverse ASL che collaborano con esso), indipendentemente da dove viene effettuato l'*input* stesso dei dati e di garantire il funzionamento in rete con un elevato grado di sicurezza e in base a protocolli standard;
- la geocodifica automatica di tutti i casi contenuti nel database del Registro Tumori (ossia, il posizionamento di ogni caso in termini di indirizzo di residenza, coordinate geografiche, sezione di censimento corrispondente), con una precisione della procedura automatizzata pari al 98,8% (rendendo molto più agevole la procedura manuale di recupero dell'informazione per i pochi casi mancanti);
- l'integrazione informatica completa con il Registro di Mortalità umbro, garantendo in questo modo un'alta tempestività dell'aggiornamento dei dati relativi allo stato in vita dei casi registrati.

Queste procedure sono state oggetto di pubblicazioni scientifiche internazionali (Bianconi et al., 2010 e 2012) e si stanno implementando in altri Registri Tumori, sia nazionali che esteri.

Accanto alle attività di registrazione ordinaria, il Registro si occupa anche dello studio di alcuni tumori particolarmente frequenti nella popolazione regionale tramite l'istituzione di protocolli operativi specifici che prevedono una raccolta di informazioni particolarmente dettagliate su tutti gli aspetti della diagnosi, delle caratteristiche patologiche, del trattamento e dell'esito della malattia oncologica stessa. La raccolta di tali dati così dettagliati è nota come "studio ad alta risoluzione". I tumori interessati sono quelli del colon-retto e della mammella (indagati da questo studio), della cervice uterina, della prostata e i melanomi.

Va ricordato che le raccolte dati ad alta risoluzione non fanno parte delle attività ordinarie di un registro tumori e, in genere, sono sviluppate solo per rispondere ad interessi di ricerca che nascono principalmente dall'area disciplinare clinica.

Inoltre, gli studi e le raccolte dati ad alta risoluzione sono limitati a campioni specifici delle popolazioni osservate dai Registri; nel caso umbro, invece, grazie alla struttura della raccolta dati sopradescritta, per le patologie studiate con il protocollo ad alta risoluzione sono presi in considerazione tutti i pazienti che hanno manifestato le neoplasie sopra citate.

Vi è poi un'ulteriore attività di ricerca portata avanti dal Registro umbro di particolare rilievo per questa tesi, sia tramite collaborazioni nazionali, sia tramite collaborazioni internazionali: l'esplorazione della relazione tra caratteristiche socio-economiche dei pazienti, fattori di rischi di tumore, accesso ai servizi sanitari (con particolare attenzione agli *screening*), qualità dei trattamenti e sopravvivenza.

4.1.5 I dati del Registro Tumori umbro

I dati del Registro sono a disposizione per attività di monitoraggio e ricerca scientifica e possono essere richiesti da chiunque sia ufficialmente inserito in un ente/istituto che svolge tali funzioni e/o sia iscritto all'AIRTUM. Per accedere ai dati del Registro, è stato necessario presentare al direttore del Registro una richiesta informale, contenente una breve descrizione dello studio e le sue finalità, e il curriculum vitae scientifico. Trattandosi di uno studio che non coinvolge direttamente il paziente in procedure di tipo sanitario, non è stato necessario richiedere il nulla osta del Comitato Etico interno. I dati

sono stati forniti in formato di archivio Microsoft Access, in seguito trasformato in matrice dati SPSS e Stata.

Le informazioni presenti nel database del Registro sono sostanzialmente: le caratteristiche della diagnosi (ad es., la data della diagnosi, lo stadio e l'estensione del tumore, l'età del paziente alla diagnosi), le modalità tecniche della diagnosi (ad es., se il tumore sia stato diagnosticato tramite esame citologico, radiologico, endoscopico), le caratteristiche demografiche e del follow-up (ad es., il codice fiscale dell'individuo, l'indirizzo di residenza, la sezione di censimento, lo stato in vita dell'individuo all'ultimo controllo e l'eventuale data di morte), le informazioni relative ai soggetti del sistema sanitario intervenuti (ad esempio, l'ospedale o il centro di cura in cui è stato diagnosticato il tumore, o in cui sono state applicate le eventuali terapie)¹⁶. Queste informazioni sono disponibili per tutti i tumori e non solo per quelli studiati con un protocollo ad alta risoluzione.

1) **Le caratteristiche demografiche.** Nelle caratteristiche demografiche sono presenti:

- Identificativo anagrafica: è un codice numerico che identifica univocamente i pazienti presenti nel database, rispetto al tipo di tumore diagnosticato e alla data di inserimento nel database del Registro; è un valore univoco per ogni paziente;
- Nome, Cognome, Sesso, Data e Comune di Nascita, Codice fiscale, Comune di Residenza: le informazioni provengono dal flusso informativo sanitario e, tramite il codice fiscale, sono verificate con l'anagrafe comunale corrispondente;
- Codice sanitario: è il codice dell'Anagrafe sanitaria, che permette l'identificazione univoca di ogni utente del Sistema Sanitario Nazionale;
- Indirizzo di Residenza e Sezione di Censimento 2001): l'indirizzo di residenza (comprensivo di numero civico) al momento della diagnosi (verificato con le anagrafi comunali) permette l'associazione con la sezione di censimento tramite gli atlanti che ISTAT mette a disposizione di comuni e regioni; per tale associazione vi è doppia verifica: la sezione corrispondente all'indirizzo è dapprima ricavata dagli atlanti comunali e poi verificata dall'atlante a disposizione dell'Ufficio Statistico Regionale.

2) **Le caratteristiche della diagnosi.** In questa sezione sono presenti tutte le informazioni con cui la malattia si è presentata alla diagnosi:

¹⁶ In Appendice, è riportato in tabella l'elenco completo delle informazioni presenti nel *dataset* (Tab A.1).

- Data di Prima Diagnosi ed Età alla Prima Diagnosi: è la data a cui viene attribuita l'incidenza del tumore ed è, quindi, l'informazione fondamentale da cui discendono tutti gli andamenti temporali delle informazioni successive di cura ed esito sanitario. Da questa data il paziente entra ufficialmente nelle statistiche epidemiologiche non solo della banca dati del Registro, ma anche della Banca Dati Nazionale AIRTUM. L'età è calcolata automaticamente a partire dalla data di nascita;
- Lateralità e Codice ICD: sono elementi identificativi del tumore; la lateralità identifica la posizione del tumore nell'organo considerato, mentre il codice ICD identifica la malattia secondo la classificazione universale delle malattie (*International Classification of Diseases*; più precisamente, *International Statistical Classification of Diseases, Injuries and Causes of Death*) stilata dall'Organizzazione mondiale della sanità (OMS-WHO) e obbligatoriamente utilizzata in tutto il mondo; essa presenta una serie di sotto-classificazioni legate alle peculiarità della malattia oncologica;
- Y, T, N, M: sono i codici con cui vengono identificate una serie di caratteristiche del tumore; più precisamente, Y indica se la classificazione della malattia è stata o meno effettuata durante (dopo il primo ciclo) o dopo una terapia multimodale in grado di modificare il decorso della patologia stessa, T indica con categorie crescenti da 0 a 4 l'estensione della malattia (più una categoria per i tumori *in situ*, e una per i tumori dall'estensione non nota¹⁷), N indica con categorie crescenti da 0 a 4 l'estensione di metastasi nei linfonodi della regione corporea interessata dal tumore (più una categoria per le estensioni regionali¹⁸ non conosciute), M identifica invece la presenza o assenza di metastasi a distanza (più una categoria per situazioni non conosciute). Le variabili identificate come Lettera T, N, M sono sotto-classificazioni più approfondite di tali caratteristiche;
- Grado e Stadio del tumore: il grado di aggressività (o malignità) del tumore esprime la pericolosità del tumore per la sopravvivenza del paziente. Il grado è classificato con

¹⁷ Per questa e le variabili N e M, avere valori non noti equivale in genere a pazienti deceduti, le cui informazioni provengono esclusivamente dal certificato di morte, che non le riporta.

¹⁸ Le regioni anatomiche utilizzate per la classificazione TNM sono le seguenti: CAPO E COLLO (ESCLUSA TIROIDE), TIROIDE, ESOFAGO, STOMACO, INTESTINO TENUE, COLON E RETTO, CANALE ANALE, FEGATO, COLECISTI, DOTTI BILIARI EXTRAEPATICI, PANCREAS, AMPOLLA DI VATER, POLMONE, PLEURA, TESSUTI MOLLI, OSSA, CUTE, MAMMELLA (LN DI I LIVELLO), VULVA, VAGINA (TERZO MEDIO E INFERIORE), VAGINA (TERZO SUPERIORE), COLLO E CORPO DELL'UTERO, OVAIO, PENE, PROSTATA, TESTICOLO, RENE, PELVI RENALE URETERE, VESCICA URINARIA, URETRA, TUMORI OFTALMICI, TUMORI PEDIATRICI.

categorie crescenti da G1 (tumore ben differenziato, con prognosi potenziale migliore) a G4 (tumore indifferenziato, dal comportamento potenzialmente più letale) più una categoria X per i gradi non conosciuti. Lo Stadio è l'informazione che sintetizza l'estensione della malattia, classificata in 4 categorie di estensione crescente, più una categoria per le estensioni non disponibili.

3) **Le modalità tecniche della diagnosi.** Le variabili di questa sezione sono tutte relative alle procedure diagnostiche a cui, potenzialmente, può essere sottoposto un paziente per identificare la presenza di un tumore e valutarne le caratteristiche iniziali, di cui al punto precedente. Tutte le variabili sono espresse in forma dicotomica, ossia forniscono l'informazione se la diagnosi è stata o meno raggiunta e confermata con le diverse modalità. Ogni paziente può essere stato sottoposto ad una o più metodiche di diagnosi, potenzialmente a tutte.

4) **Le caratteristiche del *follow-up***¹⁹. Le variabili di questa sezione sono ovviamente condizionate dalla data in cui è possibile avere il *follow-up* completo dei casi nel *dataset*: più ci si allontana dalla data di diagnosi, più aumenta la probabilità di decesso dei pazienti, non necessariamente a causa del tumore in questione. Le variabili di *follow-up* sono quindi fondamentali per il calcolo della sopravvivenza in termini di probabilità e di anni potenziali di vita. Le informazioni presenti sono:

- Stato in vita e Data di Fine *Follow-up*: questa variabile ci informa se, alla data considerata di fine *follow-up*, il paziente è vivo o deceduto. La data di fine *follow-up*, in realtà viene periodicamente aggiornata e, nel caso specifico di questo studio, ad oggi il *follow-up* completo per tutti i pazienti del Registro in generale, e dei due tumori considerati in particolare, è 31/12/2012;

- Data di Morte, Cause e Comune di Morte: in caso di decesso del paziente, queste variabili forniscono tutte le informazioni necessarie per l'utilizzo della mortalità all'interno degli indicatori di Registro. Esse sono ottenute tramite i Registri di Mortalità, come descritto in precedenza e tramite controllo diretto degli operatori sulle altre fonti concorrenti al flusso informativo, al fine di migliorare la qualità del dato e correggere eventuali errori legati al Certificato di Morte;

¹⁹ Si ricorda che per "*follow-up*" si intende il periodo per cui un individuo è seguito dall'attività del Registro Tumori, che ne riporta, ad esempio, lo stato in vita, la situazione della malattia, ecc. Teoricamente, esso dura fino alla fine della vita dell'individuo, ma in pratica non sempre ciò è possibile, perché dipende da molteplici fattori, non tutti controllabili dal Registro.

- Codice Sanitario: è il codice univoco dell'Anagrafe Sanitaria, associato ad ogni utente del Sistema Sanitario Nazionale.

5) Il ruolo operativo del Sistema Sanitario Nazionale. Le informazioni generali relative ai principali attori del Sistema Sanitario Nazionale sono:

- Medico di Base e ASL di riferimento: viene indicato il codice del medico di base e della ASL a cui il paziente fa riferimento;

- Tipo di assistito: viene indicato il tipo di assistito rispetto alle modalità di carico economico del paziente sulla Sanità Regionale. La classificazione è nazionale ed è basata sul criterio dell'appartenenza territoriale che differenzia esenzioni, pagamenti dovuti dal paziente, ecc. (Deroga interna, Domiciliato, Non residente in deroga, Residente, Residente in deroga, Straniero residente, Trasferito). Un'ultima categoria (Deceduto) indica i casi del Registro che sono arrivati alla registrazione già deceduti o deceduti prima di qualsiasi intervento;

- Ospedale, Data di Ricovero, Data di dimissione, Stato alla Dimissione: le informazioni riportate in queste variabili riguardano gli ospedali in cui, di volta in volta, è stato ricoverato il paziente dal momento della diagnosi alla fine della cura, le date in cui è stato ricoverato ed è stato dimesso, le diverse diagnosi alla dimissione.

Le informazioni relative alla raccolta dati ad alta definizione sono, invece, sintetizzate nella tabella 4.1, che presenta le sezioni in cui sono state organizzate e raggruppate.

Tabella 4.1 – I gruppi di variabili della parte ad alta risoluzione del dataset.

| Mammella | Colon-Retto |
|------------------------|------------------------|
| Caratteristiche Tumore | Caratteristiche Tumore |
| Citologia | Citologia |
| Biopsia | Biopsia |
| Linfonodi | Procedura Endoscopica |
| Linfonodi Sentinella | Istologia |
| Chirurgia | Chirurgia |
| Chemioterapia | Pezzo Operatorio |
| Immunoterapia | Linfonodi |
| Ormonoterapia | Marcatori |
| Radioterapia | Chemioterapia |
| | Immunoterapia |
| | Radioterapia |

Le variabili rilevate per ognuna di queste sezioni sono molto numerose e riguardano tutti gli aspetti descrittivi, diagnostici e terapeutici della malattia²⁰.

Gli aspetti descrittivi della malattia includono informazioni di interesse prettamente clinico, come ad esempio la focalità o i recettori di estrogeni nel tumore della mammella.

Gli aspetti diagnostici e terapeutici riportano le procedure effettuate in maniera dettagliata, il luogo in cui sono stata effettuate (ad esempio, quale tipo di chirurgia è stata applicata, in che centro di cura, ecc.), l'esito conseguito, ecc.

Per gli anni su cui si focalizza questo studio, tutte le sopracitate variabili (e quelle presentate in maniera completa in appendice) hanno subito il controllo di qualità, grazie al supporto operativo dei tecnici del registro e possono quindi essere considerate utilizzabili nello studio. I controlli che hanno utilizzato le informazioni relative a nome, cognome, codice fiscale, indirizzo di residenza sono stati operati esclusivamente dai tecnici del Registro, autorizzati ad utilizzare tali informazioni sensibili dagli emendamenti regionali alla normativa sulla *privacy*. Tali informazioni, infatti, non sono visualizzabili direttamente dall'autore di questo studio.

²⁰ L'elenco completo delle variabili relative alla raccolta dati ad alta risoluzione è riportato in Appendice (Tab. A.2.1 e Tab. A.2.2)..

4.2 Il Censimento Nazionale della Popolazione e delle Abitazioni

Oltre ai dati del Registro Tumori umbro, questo studio si avvale anche dei dati del Censimento Nazionale della Popolazione e delle Abitazioni condotto dall'ISTAT nel 2001.

Questi dati sono stati messi a disposizione dall'Ufficio Statistico della Regione Umbria e sono composti da più variabili di quelle messe pubblicamente a disposizione dall'ISTAT sul proprio sito Internet.

I dati del Censimento sono rilevanti per questo studio, perché forniscono informazioni dettagliate sulle condizioni demografiche e socio-economiche dell'area in cui risiedono i soggetti considerati. Esso coinvolge, infatti, la totalità delle famiglie e dei cittadini residenti in Italia e le informazioni raccolte riguardano le caratteristiche socio-economiche individuali (come, ad esempio, lo stato civile, il titolo di studio, la condizione professionale), familiari (come, ad esempio, il numero di individui presenti nel nucleo familiare o l'età media dei componenti il nucleo familiare) e le condizioni delle abitazioni (come, ad esempio, il titolo di godimento, il numero di stanze, la presenza di servizi).

Purtroppo però, per ragioni legate alla normativa sulla *privacy*, i dati a livello individuale non sono accessibili, ma sono invece accessibili quelli a livello di sezione di censimento, ossia l'unità territoriale di popolazione più piccola, riconosciuta a livello ufficiale (170 abitanti in media, con variazione da 0 a 3.386).

Il Censimento è realizzato ogni 10 anni da ISTAT, in collaborazione con gli enti locali (comuni, province, regioni) e sotto la supervisione del Ministero dell'Interno²¹. Nel 2001 la raccolta dei dati è avvenuta tramite questionario, fornito per l'autocompilazione a tutti gli Italiani obbligati per legge a rispondere e la cui restituzione è verificata dai comuni²².

Per la sua realizzazione, il Censimento si avvale di una struttura organizzativa complessa, che vede una stretta collaborazione fra ISTAT, gli Uffici Statistici regionali e gli Uffici di Censimento comunali. Questi soggetti assieme organizzano tutte le attività preparatorie alla rilevazione, strutturano la mappa delle aree (le sezioni di Censimento), coordinano le attività di rilevazione, definiscono le modalità di

²¹ Decreto legislativo 322/89.

²² Legge n. 144 del 1999 e Regolamento di esecuzione dei Censimenti generali, D.P.R. 276/2001, pubblicato nella Gazzetta Ufficiale n. 159 dell'11 luglio 2001.

conservazione dei dati e organizzano il sistema dei controlli per la verifica della qualità del dato.

I dati del Censimento, infatti, prima di essere pubblicati vengono sottoposti a tutta una serie di controlli di qualità che ne certificano l'alto livello di affidabilità.

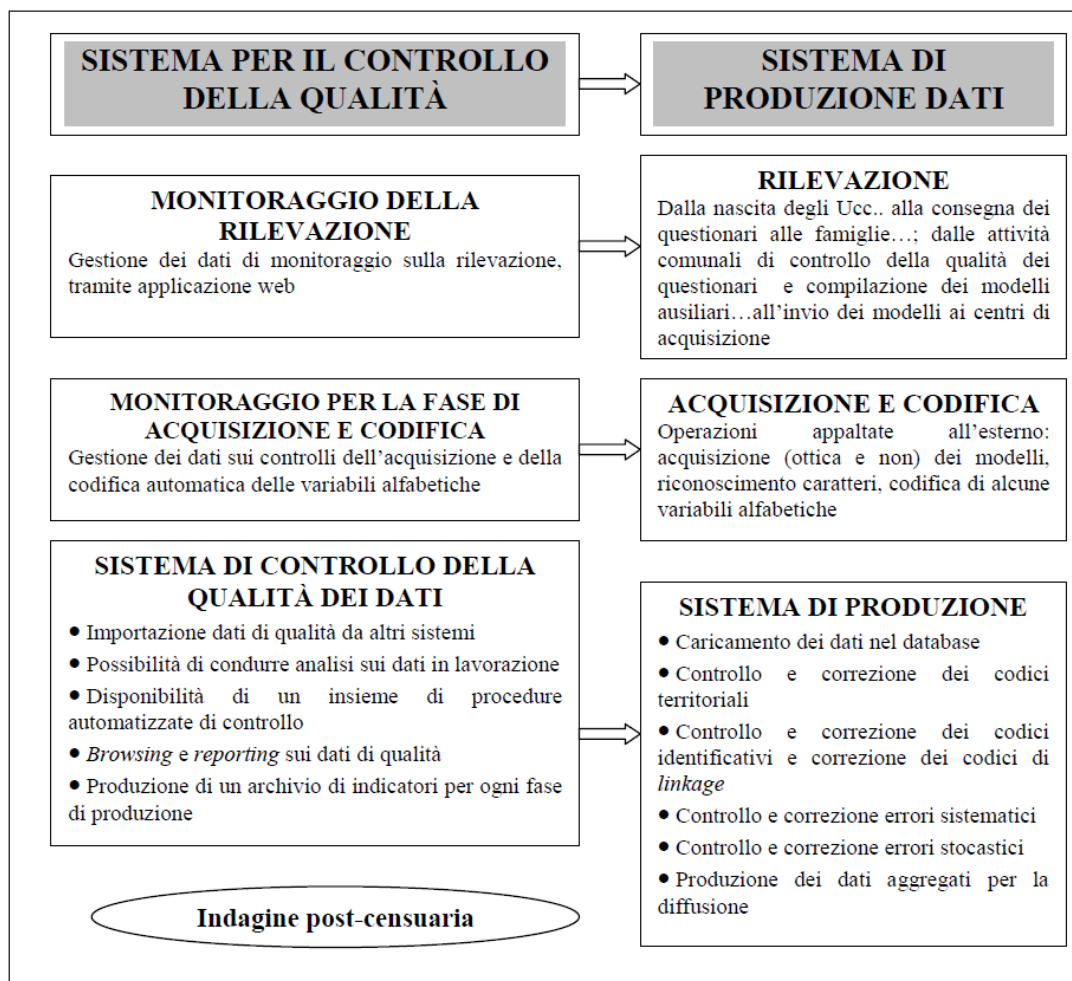
Tale controllo della qualità (attendibilità e coerenza dei dati rilevati, validità dei risultati) si sviluppa a partire dalla fase di rilevazione dei dati e termina con la loro pubblicazione e si focalizza su tre aspetti fondamentali (Panizon, 2009):

- il monitoraggio di tutte le fasi, sia sui dati che sulle procedure, in modo da garantire interventi tempestivi ed efficaci nella gestione delle situazioni anomale;
- la visione integrata di tutte le fasi censuarie, realizzabile con uno scambio di informazioni fra i diversi sistemi di monitoraggio ed elaborazione dei dati;
- la registrazione degli eventi, dei dati derivati dai controlli e dei risultati delle procedure in un "archivio di qualità", per consentire una valutazione finale dell'errore e fornire indicazioni sintetiche sulla qualità dei dati.

Senza entrare nei dettagli delle procedure seguite, la figura seguente (Fig. 4.1) riporta il sistema con cui è stato sviluppato e applicato il controllo di qualità del dato, integrato strettamente nel sistema di produzione stessa del dato.

Le aree di analisi nel sistema di controllo della qualità vengono individuate sulla base delle possibili tipologie di errori (Chieppa et al., 2001); per ciascuna di queste aree sono stati previsti sia controlli automatizzati sia controlli manuali.

Fig. 4.1 – Il sistema di controllo di qualità del dato.



Fonte: Panizon, 2009.

Le tipologie di errori considerate sono le seguenti sei.

- 1) **Errori nel calcolo dei primi dati di Censimento (dati provvisori)**. Vengono effettuate analisi sulla validità/attendibilità dei primi risultati per avere livelli di qualità necessari alla diffusione dei dati; gli Uffici Statistici delle Regioni e gli uffici regionali ISTAT sono gli attori di tali controlli.
- 2) **Errori attribuibili a smarrimento o alterazione di modelli di rilevazione**. Vengono effettuate analisi di controllo sulle quantità di questionari acquisiti, confrontando i dati stimabili dai dati ottenuti tramite il monitoraggio della rilevazione e della fase di acquisizione. Qui intervengono gli uffici comunali e provinciali interessati.

3) **Errori di copertura.** Vengono confrontati i dati censuari e dati che derivano da altre fonti demografiche. Questo controllo è sviluppato per lo più a livello di Regione e di ISTAT.

4) **Errori nella identificazione delle unità censite.** Vengono effettuate analisi sui totali e sulle relazioni tra le diverse unità (ad esempio, edifici e abitazioni, abitazioni e famiglie); qui il controllo è affidato agli uffici comunali.

5) **Errori nei valori relativi alle singole variabili rilevate.** Viene analizzata la distribuzione delle variabili; ISTAT e Regioni sviluppano tale controllo.

6) **Errori introdotti dalle procedure di elaborazione.** Vengono analizzate le modifiche introdotte sui dati e queste analisi sono gestite da ISTAT.

Il completamento di queste fasi di controllo permette la distribuzione agli enti territoriali, in particolare alle Regioni, del *database* finale. Essi possono analizzarlo e produrre interpretazioni/pubblicazioni solo in accordo con ISTAT e per le aree territoriali di loro competenza. Possono altresì fornire agli enti locali e alle istituzioni che fanno parte del sistema statistico regionale tale *database*, anche in anticipo sui tempi di pubblicazione ISTAT, per le finalità previste dagli enti stessi.

Ad esempio, alla fine delle operazioni del Censimento, viene effettuato il confronto fra i dati censuari raccolti e le informazioni presenti nell'Anagrafe municipale, in modo da permettere la revisione dell'Anagrafe²³.

Vengono mantenuti nell'Anagrafe gli iscritti che sono stati censiti come abitualmente dimoranti nel proprio territorio e cancellati coloro che, in successivi accertamenti, non sono stati rintracciati. Operando correttamente e tempestivamente questa revisione anagrafica, i comuni possono così disporre di archivi aggiornati e corrispondenti alla situazione reale della propria popolazione. Questi archivi, continuamente aggiornati con i dati su nascite, immigrazioni, morti ed emigrazioni, garantiscono nel periodo intercensuario la disponibilità di informazioni indispensabili per la gestione economica e sociale del territorio.

²³ Secondo l'Art. 46 del D.P.R. n. 223 del 30 maggio 1989, Regolamento anagrafico di esecuzione della Legge n. 1228 del 24 dicembre 1954 sull'ordinamento delle Anagrafi della popolazione residente.

4.2.1 I dati umbri del Censimento della Popolazione e delle Abitazioni

E' stato possibile avere a disposizione direttamente il *database* regionale per sezione di censimento dei dati relativi al Censimento Nazionale 2001, grazie al fatto che il Registro Tumori umbro è parte della Rete Oncologica regionale ed è, di fatto, un organo sanitario della Regione. Trattandosi di dati pubblici e rilasciati in formato aggregato (per sezione di censimento), non è presente alcuna problematica relativa alla *privacy* e, per ottenerli, è stato sufficiente presentare una richiesta motivata all'Ufficio Statistico regionale, simile a quella presentata al Registro Tumori.

Il database contenente le informazioni per ogni sezione di Censimento è stato trasmesso via *e-mail* dall'Ufficio Statistico della Regione Umbria al sottoscritto e, in copia, al Registro Tumori, in formato Microsoft Excel e trasformato in matrice dati SPSS e Stata. I dati messi a disposizione dalla Regione hanno un vantaggio specifico rispetto a quelli che mette liberamente a disposizione ISTAT sul suo sito: la Regione può fornire una serie di variabili che altrimenti non sono disponibili al pubblico. Queste variabili riguardano in misura minore alcuni aspetti delle caratteristiche di popolazione ma, più diffusamente, aspetti delle caratteristiche abitative. Ad esempio, sono state fornite informazioni su alcuni aspetti relativi alle caratteristiche di proprietà, alle condizioni strutturali dell'immobile e alla tipologia di servizi igienici e di riscaldamento (questi ultimi di interesse per gli aspetti di salute ad essi connessi).

La tabella seguente mostra una sintesi delle principali variabili dal Censimento, per aree tematiche, che vengono usate nello studio (Tab 4.2).

Tab. 4.2 – Sintesi delle variabili relative al Censimento 2001.

| Aree tematiche | Esempi |
|---|---|
| Codice Provincia | Il numero di codice che identifica la Provincia in tutti gli archivi ISTAT. Per la Provincia di Perugia, ad esempio, è 054. |
| Codice Comune | Il numero di codice che identifica il Comune in tutti gli archivi ISTAT. Per il Comune di Assisi, ad esempio, è 054001. |
| Comune | Il nome del comune per esteso, in accordo alla denominazione ufficiale. |
| Sezione di Censimento | Il numero di sezione di Censimento. |
| Caratteristiche demografiche e di composizione del nucleo familiare | Età (quante persone di una certa età sono presenti in ogni sezione di Censimento), residenza italiana o estera, numerosità del nucleo familiare, età media del nucleo familiare, ecc. |
| Stato civile | Quanti celibi/nubili, sposati, separati, ecc. sono presenti in ogni sezione di Censimento. |
| Livello di istruzione | Quanti individui con titolo di media superiore, inferiore, laurea, ecc. sono presenti in ogni sezione di Censimento. |
| Differenti aspetti della condizione lavorativa | Occupati, disoccupati, tipologie dell'occupazione, ecc. |
| Indici di sintesi ISTAT | Indice di vecchiaia, indice di ricambio, indice di dipendenza strutturale, tasso di attività, tasso di occupazione, tasso di disoccupazione. |
| Caratteristiche delle abitazioni | Dimensioni, struttura, tipo di proprietà, tipo di utilizzo, condizioni strutturali, caratteristiche dei servizi presenti nell'abitazione, ecc. |

L'unità di analisi è la sezione di Censimento e i dati di ogni sezione sono forniti in valori assoluti e percentuali (ad esempio, quanti individui dal titolo di studio elementare vi sono in una data sezione e che percentuale rappresentano rispetto al totale dei residenti in tale sezione).

Le caratteristiche demografiche e la composizione del nucleo familiare identificano le caratteristiche dominanti in ogni sezione; in questo studio esse risultano utili per descrivere similitudini e differenze di tipo demografico fra i casi del Registro e il totale dei residenti delle medesime aree in cui essi vivono.

Stato civile, livello di istruzione e condizione lavorativa sono le variabili d'area che, in questo studio, rivestono l'interesse maggiore: come già visto nel quadro teorico, sono le informazioni che permettono, sia a livello d'area che a livello individuale, di descrivere le disuguaglianze in salute rispetto alla condizione socio-economica. Lo stato civile da

Censimento è codificato in: Celibi/Nubili, Coniugato/a, Separato/a di fatto, Separato/a legalmente, Divorziato/a, Vedovo/a. Il livello di istruzione è codificato in: Non sa leggere e scrivere, Sa leggere e scrivere, Licenza di scuola elementare, Licenza di scuola media inferiore, Diploma di scuola secondaria superiore, Diploma non universitario post maturità, Diploma universitario, Laurea. Per quanto riguarda la situazione lavorativa, sono codificate sia la situazione di occupazione/disoccupazione (quindi occupati, disoccupati, in cerca di prima occupazione, ritirati, inabili al lavoro, studenti, casalinghe), sia le tipologie di occupazione (imprenditori, liberi professionisti, dipendenti, lavoratori in proprio, soci di cooperative, coadiuvanti familiari).

Gli indici di sintesi forniscono una lettura immediata delle principali caratteristiche socio-economiche delle singole sezioni. L'indice di vecchiaia è il rapporto percentuale fra la popolazione di 65 anni e più e con età inferiore o pari a 14 e descrive il peso della popolazione anziana. L'indice di ricambio è il rapporto percentuale tra coloro che stanno per lasciare il mondo del lavoro (la popolazione fra i 60 e 64 anni) e coloro che vi stanno per entrare (la popolazione tra i 15 e i 19 anni) e fornisce un'indicazione sulla capacità che ha una popolazione di sostituire la generazione di lavoratori anziani con nuove forze lavorative. L'indice di dipendenza strutturale è il rapporto percentuale tra la popolazione in età non attiva (con età inferiore o pari a 14 anni e con 65 anni e più) e la popolazione in età attiva (fra i 15 e i 64 anni) e misura la capacità di supporto di una popolazione. Il tasso di attività è calcolato da ISTAT come rapporto fra la forza lavoro fra i 15 e i 64 anni e la popolazione con 15 e più anni²⁴ e misura l'offerta di lavoro. Il tasso di occupazione è calcolato come rapporto percentuale fra occupati e popolazione e misura l'incidenza della popolazione con un lavoro sulla popolazione in generale. Il tasso di disoccupazione, infine, è dato dal rapporto percentuale fra le persone in cerca di lavoro e la forza lavoro e misura l'eccesso di forza lavoro.

Per quanto riguarda le abitazioni, anche le informazioni relative al titolo di godimento della casa, alle dimensioni e alla condizione dell'immobile, alla tipologia di impianti disponibili possono essere utili per valutare le disuguaglianze socio-economiche.

²⁴ Eurostat utilizza una definizione differente, limitando il denominatore alla popolazione fra 15 e 64 anni.

4.3 Le Anagrafi comunali

Il terzo tipo di dati utilizzati in questo studio sono i dati di tipo demografico e socio-economico relativi ai singoli pazienti, ossia Comune di residenza, indirizzo, data di nascita, sezione di Censimento 2001 di residenza, genere, stato civile, titolo di studio e condizione professionale. Queste informazioni non sono raccolte centralmente ma localmente dalle singole Anagrafi; è quindi stato necessario farne richiesta a tutte le Anagrafi comunali della regione, contattando tutti i 92 comuni umbri.

L'Anagrafe della popolazione residente è, infatti, una fonte fondamentale per ottenere informazioni di tipo individuale su una popolazione estesa ma circoscritta come quella di un comune. Essa registra nominativamente le caratteristiche demografiche e socio-economiche di tutti gli abitanti residenti in un comune, nonché le successive variazioni che si verificano nella popolazione stessa.

Questa attività di registrazione e aggiornamento delle informazioni individuali è particolarmente importante non solo per il territorio specifico su cui opera, ma anche perché costituisce la base su cui vengono organizzati numerosi servizi pubblici (elettorale, scolastico, tributario, assistenziale, ecc.).

Rispetto alla modalità di acquisizione delle informazioni, per ogni nato registrato nel comune di nascita viene istituito un *record* in cui sono inserite alcune informazioni sull'individuo e in cui saranno di seguito registrate tutte le variazioni nello stato civile che interverranno nella vita di un individuo, informazioni relative a titolo di studio, stato occupazionale, eventuali disabilità, ecc.

Le modalità di archiviazione dei dati variano da comune a comune, a seconda che esse siano conservate in formato elettronico, cartaceo o entrambe (la maggioranza dei casi). Le procedure di sicurezza devono comunque essere uniformemente adeguate a quanto previsto dalla normativa sulla *privacy*²⁵ e dalle modalità attuative definite sia a livello nazionale che locale.

In più, le informazioni raccolte sono potenzialmente collegabili fra loro dalle caratteristiche individuali di nome, cognome, data di nascita, indirizzo di residenza e codice fiscale che costituiscono le chiavi di unione, costruendo *record* di informazioni unici per ogni residente, anche se questa possibilità di *data linkage* è sfruttata solo da

²⁵ Decreto legislativo del 30 Giugno 2003, n. 196.

pochi comuni (in realtà non è possibile quantificare con precisione quanti comuni sfruttino questa opportunità).

Considerando, infine la qualità dei dati raccolti dalle Anagrafi, essi tendono a presentare in maniera ricorrente alcuni errori. I più comuni in questo tipo di dati sono fondamentalmente due.

1) **Errore di imputazione.** Non sempre gli operatori che ricevono/raccolgono le informazioni dagli individui le riportano o le inseriscono nel *database* in maniera omogenea. Sono errori tipici gli spazi a volte inseriti e a volte omessi in cognomi composti da più parole (De Ponti, DePonti), o la diversa imputazione dello stesso luogo di residenza (Via S. Giovanni, Via San Giovanni), o l'errore di scrittura in strutture alfanumeriche complesse come il codice fiscale. Questi errori sono totalmente imprevedibili, legati ad un'incompleta conoscenza condivisa delle regole di imputazione delle informazioni fra i dipendenti comunali che gestiscono tali dati. Sono critici in fase di *linkage* fra *database*, perché portano alla perdita completa delle informazioni relative ad un individuo. Non vi è, presso i Comuni, una sistematicità nei controlli di qualità e nelle correzioni di questi errori; in genere sono realizzati quando sono stimolati da richieste da altri enti (ad esempio INPS, Uffici elettorali, ASL, lo stesso Censimento).

I controlli sono realizzati tramite confronto con altre fonti amministrative, come l'anagrafe fiscale o l'anagrafe sanitaria, e/o tramite contatto diretto con l'individuo interessato.

2) **Errore di aggiornamento.** Le informazioni conservate nei *database* delle anagrafi non sono sempre le più recenti a disposizione. Magari le informazioni sono già state raccolte, ma giacciono inevase sulle scrivanie degli operatori. Di questa situazione ci si rende conto, in particolare, quando si valutano informazioni come, ad esempio, lo stato in vita di un individuo, o altre informazioni che possono cambiare in maniera repentina. Anche in questo caso, il controllo di qualità o la correzione dell'errore non sono attività sistematiche, ma vengono poste in essere quando l'errore si rende palese. Questo accade quando istituzioni o enti fanno richiesta di tali informazioni e ci si trova in presenza di difformità generate dall'incrocio di diverse fonti informative. Come visto nel paragrafo precedente, i diversi Censimenti sono un esempio di momento di controllo e verifica dell'aggiornamento delle informazioni individuali conservate dalle anagrafi comunali.

4.3.1 I dati delle Anagrafi comunali umbre

Tutte le Anagrafi dei 92 comuni umbri hanno partecipato alla raccolta dati, secondo una modalità in grado di garantire il rispetto della normativa sulla *privacy*. Questa problematica riguardava l'eventuale autorizzazione per interrogare i *database* comunali per ottenere le informazioni individuali. Tale autorizzazione, in realtà, non era necessaria in questo caso, perché il Registro è da sempre parte della Rete Oncologica regionale, di cui fanno parte anche i comuni tramite le ASL che vi operano, ed aveva già avuto le necessarie autorizzazioni dai comuni per ricevere le informazioni dai municipi quando aveva iniziato ad operare nel 1994. Questa raccolta dati si configurava, quindi, come procedura di aggiornamento del *database* del Registro, e come tale non necessitava di ulteriori autorizzazioni.

La raccolta dati si è svolta fra il 2013 e il 2015 e sono state considerate le informazioni disponibili al 31/12/2012.

A questo punto è stato necessario definire con il personale del Registro Tumori una procedura standard a cui ogni comune potesse aderire per organizzare e trasmettere le informazioni richieste.

Si sono, inizialmente, individuati all'interno del *database* generale del Registro Tumori gli individui su cui condurre lo studio, suddividendoli per comune di residenza. Sono stati considerati solo gli individui che al 31/12/2012, data di fine *follow-up* per questo studio, fossero residenti in un comune umbro.

In questa fase ad ogni individuo è stato anche associato un codice numerico univoco, diverso dal codice fiscale, in grado di identificarlo in maniera anonima; è stato un passaggio fondamentale perché ha permesso all'autore di questo studio, non autorizzato a trattare i dati nominativi, di gestire successivamente una parte delle procedure di *record linkage* fra i dati di Registro e i dati dalle Anagrafi.

Per ogni comune è, quindi, stata costruita una scheda contenente gli identificativi di ogni individuo (codice fiscale, nome, cognome, data di nascita ed eventuale data di morte) e le informazioni richieste. La tabella seguente (Tab. 4.3) riporta l'elenco dei comuni e il numero di casi per patologia che sono stati richiesti ad ogni Anagrafe.

Tab. 4.3 – Distribuzione dei casi per comune.

| Comuni | Mammella Donne | Colon- Retto Entrambi i generi | Comuni | Mammella Donne | Colon- Retto Entrambi i generi | Comuni | Mammella Donne | Colon- Retto Entrambi i generi |
|--------------------------|-------------------|---|-------------------------------|-------------------|---|-----------------------------|-------------------|---|
| ACQUASPARTA | 44 | 56 | FOLIGNO | 553 | 766 | PARRANO | 9 | 12 |
| ALLERONA | 17 | 30 | FOSSATO DI VICO | 19 | 40 | PASSIGNANO SUL TRASIMENO | 37 | 84 |
| ALVIANO | 13 | 12 | FRATTA TODINA | 11 | 30 | PENNA IN TEVERINA | 12 | 7 |
| AMELIA | 98 | 159 | GIANO DELL'UMBRIA | 29 | 37 | PERUGIA | 1.575 | 2.364 |
| ARRONE | 28 | 42 | GIOVE | 17 | 24 | PIEGARO | 28 | 65 |
| ASSISI | 243 | 354 | GUALDO CATTANEO | 56 | 75 | PIETRALUNGA | 26 | 27 |
| ATTIGLIANO | 19 | 24 | GUALDO TADINO | 169 | 236 | POGGIODOMO | 2 | 3 |
| AVIGLIANO UMBRO | 18 | 32 | GUARDEA | 17 | 26 | POLINO | 1 | 8 |
| BASCHI | 19 | 48 | GUBBIO | 283 | 375 | PORANO | 17 | 28 |
| BASTIA UMBRA | 175 | 251 | LISCIANO NICCONE | 4 | 7 | PRECI | 6 | 8 |
| BETTONA | 30 | 56 | LUGNANO IN TEVERINA | 11 | 17 | SAN GEMINI | 45 | 56 |
| BEVAGNA | 58 | 57 | MAGIONE | 114 | 204 | SAN GIUSTINO | 91 | 129 |
| CALVI DELL'UMBRIA | 11 | 23 | MARSCIANO | 136 | 225 | SAN VENANZO | 17 | 22 |
| CAMPELLO SUL CLITUNNO | 36 | 35 | MASSA MARTANA | 35 | 58 | SANT'ANATOLIA DI NARCO | 1 | 9 |
| CANNARA | 44 | 42 | MONTE CASTELLO DI VIBIO | 13 | 27 | SCHEGGIA E PASCELUPO | 19 | 24 |
| CASCIA | 23 | 30 | MONTE SANTA MARIA TIBERINA | 3 | 16 | SCHEGGINO | 3 | 10 |
| CASTEL GIORGIO | 19 | 45 | MONTECASTRILLI | 35 | 50 | SELLANO | 8 | 15 |
| CASTEL RITALDI | 27 | 37 | MONTECCHIO | 7 | 25 | SIGILLO | 25 | 31 |
| CASTEL VISCARDO | 27 | 64 | MONTEFALCO | 50 | 74 | SPELLO | 78 | 114 |
| CASTIGLIONE DEL LAGO | 125 | 225 | MONTEFRANCO | 7 | 17 | SPOLETO | 509 | 560 |
| CERRETO DI SPOLETO | 10 | 15 | MONTEGABBIONE | 16 | 29 | STRONCONE | 44 | 60 |
| CITERNA | 25 | 42 | MONTELEONE D'ORVIETO | 14 | 35 | TERNI | 1.307 | 1.490 |
| CITTA' DELLA PIEVE | 59 | 97 | MONTELEONE DI SPOLETO | 4 | 10 | TODI | 147 | 268 |
| CITTA' DI CASTELLO | 339 | 555 | MONTONE | 11 | 17 | TORGIANO | 42 | 79 |
| COLLAZZONE | 21 | 38 | NARNI | 217 | 273 | TREVI | 66 | 95 |
| CORCIANO | 172 | 227 | NOCERA UMBRA | 43 | 67 | TUORO SUL TRASIMENO | 33 | 69 |
| COSTACCIARO | 9 | 11 | NORCIA | 39 | 62 | UMBERTIDE | 143 | 230 |
| DERUTA | 72 | 93 | ORVIETO | 177 | 367 | VALFABBRICA | 29 | 51 |
| FABRO | 26 | 56 | OTRICOLI | 14 | 17 | VALLO DI NERA | 4 | 4 |
| FERENTILLO | 14 | 26 | PACIANO | 5 | 12 | VALTOPINA | 7 | 19 |
| FICULLE | 14 | 38 | PANICALE | 42 | 108 | | | |

Le operazioni iniziali di contatto si sono svolte telefonicamente fra Novembre e Dicembre 2013 con la mediazione dell'allora Direttore di Registro, Prof. Francesco La Rosa. Il contatto telefonico ha permesso di verificare la disponibilità dei comuni a collaborare. Poiché l'attività è stata presentata come inquadrata nelle attività di Registro

a cui i comuni già partecipano, non vi sono state particolari difficoltà ad ottenere la collaborazione di tutti i comuni. E' stata richiesta, da parte dei comuni stessi, una certa flessibilità nei tempi di consegna e non è stata data garanzia che tutte le informazioni richieste sarebbero state disponibili.

E' poi stata preparata una lettera ufficiale di presentazione dell'iniziativa per ogni singolo comune, a firma del Direttore del Registro e in cui era identificato l'autore di questo studio come referente finale²⁶. La lettera, a cui è stata allegata la tabella-elenco con i pazienti corrispondenti, è stata trasmessa via fax a Gennaio 2014.

Al fax è seguita anche una *e-mail*, con allegati sia la lettera che gli elenchi, per facilitare eventuali risposte in formato elettronico. Quest'ultima modalità è stata utilizzata dalla grande maggioranza delle Anagrafi; solo alcuni piccoli comuni con un numero estremamente ridotto di casi ha utilizzato la possibilità di rispondere via fax.

Il mese di Marzo ha visto un primo richiamo telefonico di verifica a tutti i comuni, anche se nel periodo Febbraio-Aprile erano state inviate tutte le informazioni disponibili relative alla conferma dei dati anagrafici e allo stato civile. Queste erano infatti le informazioni immediatamente a disposizione delle Anagrafi; per le altre era necessario un lavoro di confronto fra le diverse banche dati all'interno dei comuni, le cui modalità operative non sono però state specificate dalle anagrafi.

Il periodo da Maggio a Luglio ha visto un rallentamento nel flusso informativo; è stato quindi necessario un nuovo richiamo telefonico a tutti i comuni e un incontro in Umbria con i tecnici del Registro per riattivare la raccolta informativa. La raccolta dati si è conclusa ai primi di Ottobre.

Tutti i comuni sono stati in grado di fornire le informazioni richieste, riportate nella tabella 4.4.

²⁶ Consultabile in Appendice (Doc. A.1).

Tab. 4.4 – Le variabili demografiche e socio-economiche individuali.

| Variabile | Categorie |
|--------------------------|--|
| Comune di Residenza | |
| Indirizzo | |
| Data di nascita | |
| Stato civile | Celibi/nubili |
| | Coniugati/Conviventi |
| | Separati/divorziati |
| | Vedovi |
| Titolo di studio | Licenza elementare o analfabeta |
| | Licenza media inferiore |
| | Licenza media superiore |
| | Laurea |
| Condizione professionale | Legislatore, dirigente o imprenditore |
| | Professione intellettuale, scientifica o di elevata specializzazione |
| | Professione qualificata nelle attività commerciali o nei servizi |
| | Professione tecnica |
| | Impiegato |
| | Artigiano, operaio specializzato o agricoltore |
| | Professione non qualificata |
| | Pensionato |
| | Disoccupato |
| | Casalinga |
| Studente | |

Le informazioni relative a Stato civile, Titolo di studio e Condizione professionale sono le variabili fondamentali per valutare le differenze socio-economiche individuali e compararle con le differenze socio-economiche espresse dalle variabili che descrivono la sezione di Censimento di residenza.

I comuni, come richiesto, hanno inviato i dati in formato Excel o via fax; questo ha permesso al personale del Registro di riportare agevolmente i dati nel proprio *database*, senza particolari problemi di lettura delle informazioni fornite.

Il processo di controllo di qualità del dato è stato eseguito dal personale tecnico del Registro, con il supporto dell'autore, secondo due diverse modalità:

- per le informazioni richieste come aggiornamento delle variabili già presenti (Comune di residenza, indirizzo, data di nascita, sezione di Censimento 2001) sono stati verificati tutti i casi con una procedura automatizzata interna al Registro, che permette di sostituire l'informazione più recente a quella più datata, laddove vi fossero differenze. Questa modalità di controllo fa parte delle procedure abituali di aggiornamento dati del Registro;
- per le informazioni relative alle nuove variabili, la loro correttezza è stata verificata tramite controllo telefonico su un campione casuale di individui, il cui numero di telefono (sia fisso che cellulare) è contenuto nel *database* del Registro. Questo controllo si è reso necessario, in quanto nel momento in cui queste informazioni diverranno parte delle informazioni utilizzate in maniera routinaria, esse dovranno essere il più corrette possibili. Il controllo, effettuato su circa 200 individui non ha fatto rilevare errori significativi.

A questo punto, con tutte le informazioni richieste per lo sviluppo dello studio a disposizione, il lavoro si è concentrato sulla costruzione dei due *dataset*.

4.4 L'unione dei *dataset* (*data linkage*)

Per ognuna delle due sedi tumorali considerate è stato necessario creare un unico *dataset* che includesse le variabili provenienti dai *database* del Registro Tumori, dalle Anagrafi comunali e dal Censimento. Per la creazione di ogni *dataset* è stata utilizzata una procedura di *record linkage*.

4.4.1 Definizione di *record linkage*

Il *record linkage*, procedura che viene citata per la prima volta in letteratura da Dunn (1946), consiste in una tecnica algoritmica il cui scopo è l'unione di più file di diversa

natura (*dataset* amministrativi, *survey*, ecc.), che raccolgono informazioni sulle stesse unità di analisi (Belin et al., 1995, ripreso da Scanu, 2003).

Lo scopo primario di una procedura di *record linkage* è di aumentare il numero di informazioni disponibili per ogni unità di analisi, permettendo così di rispondere a domande di ricerca complesse.

Tale procedura è però possibile solo se sono raccolte le informazioni relative ad uno o più elementi identificativi di ogni individuo (ad esempio, nome e cognome, data di nascita, codice fiscale).

Ovviamente, la sua diffusione è stata direttamente proporzionale al crescere della quantità delle informazioni individuali presenti nei *database* provenienti, appunto, da fonti sempre più numerose e differenziate (da enti pubblici, aziende private, agenzie di ricerca, ecc.), che raccolgono informazioni sugli stessi individui.

Lo sviluppo delle tecnologie informatiche necessarie alla realizzazione del *data linkage* ha contribuito in maniera fondamentale alla sua diffusione (Machado, 2004) e ha permesso la sua applicazione anche a *dataset* di natura fra loro differente. Se, infatti, il *record linkage* almeno inizialmente si proponeva di sviluppare una tecnica per il *matching* di informazioni presenti in *database* principalmente amministrativi (demografici, ospedalieri, ecc.), attualmente si sta diffondendo la pratica di *linkage* fra dati amministrativi e dati provenienti da inchieste campionarie.

In Italia, l'interesse per questi metodi si è sviluppato a partire dalla metà degli anni '80, inizialmente in ambito demografico (Pinelli, 1984). In particolare in ISTAT questi metodi sono stati usati per i dati delle forze lavoro (ad esempio, Giusti et al., 1991) oltre che per la costruzione dell'Archivio Statistico delle Imprese Attive (ASIA) (Garofalo, 1998). E' aperta anche la discussione sulla fattibilità del *linkage* fra dati amministrativi e inchieste campionarie (ad esempio, Sala, 2013), con questioni sia di tipo metodologico che legate alla sensibilità dei dati e al consenso del *linkage* fra differenti fonti (ad esempio, Marchi, 2003).

Si è soliti distinguere fra due modelli principali di *record linkage*: deterministico e probabilistico. Questi modelli differiscono principalmente per le procedure di implementazione.

1) **Il *record linkage* deterministico.** Il *record linkage* di tipo deterministico collega le informazioni relative alla stessa unità di analisi presente in differenti *dataset* tramite un (o più) codice identificativo univoco associato ad ogni unità di analisi. Esempi di tali identificativi possono essere il codice fiscale in Italia o il National Insurance Number nel Regno Unito (NINO, Jenkins et al., 2008), il nome, il cognome, la data di nascita. La presenza di questo tipo di identificativi nei *dataset* che si vogliono unire è quindi condizione necessaria per realizzare un *record linkage* deterministico.

La tecnica di *linkage* deterministica presenta però una serie di criticità rilevanti, legate all'incertezza potenziale che esiste quando si opera il collegamento fra i *dataset* (Scheuren, 1999).

Tali criticità riguardano, fondamentalmente, la qualità dei dati utilizzati per il *linkage*, che a volte non permettono la corretta identificazione degli individui, e le questioni attinenti alla richiesta di autorizzazione ad unire i dati.

In merito alla prima criticità, è possibile la presenza di errori che influenzano la qualità delle variabili chiave e che possono pregiudicare la possibilità di aggancio dei *record* e/o rendere meno accurato il risultato (Scanu, 2003):

- errori di trascrizione e registrazione: immissione errate dei codici, variazioni di codice come ad esempio l'uso di diverse versioni del nome;
- errori legati alla mancanza di una o più variabili chiave per qualche unità di analisi.

Questi errori possono produrre:

- falsi non abbinamenti: alcuni *record* delle due basi di dati fanno riferimento alla stessa unità ma il collegamento non è in grado di individuarli in quanto almeno una variabile chiave è affetta da qualche errore o è mancante;
- falsi abbinamenti: alcuni *record* possono essere abbinati anche se in realtà fanno riferimento a unità diverse. Le variabili chiave coincidono per via degli errori sopra descritti.

E' comunque possibile ridurre il manifestarsi di tali errori tramite controlli incrociati da effettuarsi sulle unità di analisi nelle banche dati originarie, prima delle operazioni di *linkage* (Scanu, 2003).

In merito alla seconda criticità, va ricordato come in molte circostanze la natura nominativa dei dati implica che il *record linkage* deterministico sia possibile solo in presenza del consenso esplicito da parte degli individui presenti nel *database*, in accordo con la normativa sulla *privacy*.

Quest'ultimo problema normativo è, in realtà, presente in ogni nazione dove esiste una legge che tutela l'utilizzo di dati sensibili e la possibilità unirli attraverso diversi *database*.

Il modo di superare legalmente questo possibile ostacolo è differente per ogni realtà nazionale, ma esistono alcune strategie operative condivise:

- ai soggetti interessati, le cui informazioni confluiscono nel *database*, viene chiesto il consenso preventivo ad operare sui dati, nel momento in cui forniscono l'informazione. L'utilizzo diretto dei dati sensibili deve però avvenire esclusivamente da parte del personale espressamente autorizzato al trattamento di tali dati; qualsiasi utilizzo esterno deve avvenire in forma esclusivamente anonima e non riconducibile ai singoli individui;
- ai soggetti interessati viene chiesto il consenso ad operare sui loro dati nel momento in cui si rende necessaria l'operazione di *linkage* per uno scopo specifico (ad esempio, uno studio scientifico).

Un interessante esempio di *record linkage* deterministico è il lavoro di Jenkins et al. (2008) che ha valutato la fattibilità nel Regno Unito dell'unione fra i *record* individuali di un'inchiesta campionaria e il *dataset* governativo sui benefici fiscali. In questo studio sono state utilizzate cinque chiavi di *linkage*: codice di Previdenza Sociale Nazionale fornito dai rispondenti all'inchiesta, genere, nome e cognome, indirizzo e data di nascita.

Lo studio ha mostrato come il *record linkage* fra dati da inchiesta campionaria e dati amministrativi è fattibile e può produrre buoni risultati in termini di numero di abbinamenti e loro accuratezza.

2) **Il *record linkage* probabilistico.** Il *record linkage* probabilistico è una tecnica di abbinamento che collega le informazioni presenti in differenti *dataset*, utilizzando il calcolo probabilistico per determinare se le informazioni presenti appartengano o meno alla stessa unità di analisi.

La probabilità di abbinamento viene calcolata su un insieme di variabili identificative rilevanti (ad esempio, nome, sesso, data di nascita, ecc.); il risultato di tale calcolo produce una serie di pesi probabilistici associati ad ogni variabile. Tali pesi sono quindi utilizzati per determinare se una coppia di record si riferisce in maniera probabilistica alla stessa unità statistica (Li et al., 2006).

Questa tecnica è utilizzata quando non sono presenti valori identificativi univoci che permettano il *record linkage* deterministico, perché in questa situazione teoricamente massimizza la probabilità di abbinamento corretto e può permettere anche di ridurre eventuali errori di codifica nelle variabili utilizzate per il *linkage* (Newcombe et al., 1959; Newcombe, 1988). Può tuttavia portare a distorsioni su alcuni *link* potenziali e inoltre richiede dettagliate conoscenze a priori su alcune importanti misure relative a specifici valori identificativi, per esempio la distribuzione di frequenza delle variabili da utilizzare per il *linkage* (Van Den Brandt et al., 1990).

I metodi per il *record linkage* probabilistico sono stati sviluppati a partire dal lavoro di Newcombe et al. (1959), mirato a risolvere il problema di collegare le informazioni fra archivi ospedalieri differenti. Questi autori introdussero per primi l'utilizzo di pesi probabilistici, basati sulle frequenze con cui si presentavano le stesse informazioni all'interno delle variabili considerate per il *record linkage* (ad esempio, nome e cognome). Tali pesi si ottenevano sommando, ad esempio, i pesi probabilistici di nome, cognome ed età; valori alti nei pesi delle combinazioni presenti erano considerati indicazioni di *matching*.

Fellegi e Sunter (1969) sono partiti dal lavoro di Newcombe per introdurre le basi matematiche e statistiche del *record linkage* probabilistico usate correntemente tutt'oggi. Il loro approccio può essere sintetizzato nei seguenti punti (Machado, 2004).

a) La procedura di abbinamento conduce a stimare, per ognuna delle N possibili coppie di *record*, il valore ignoto di una variabile indicatrice e , a tal fine, è possibile utilizzare i valori assunti nei *record* da alcune variabili di confronto presenti in entrambi i *dataset*. In particolare, Newcombe (1988) osserva che tali variabili devono avere la capacità di

discriminare al meglio gli individui presenti nei due archivi, in caso di discordanza, o di concordanza o, nella migliore delle ipotesi, in entrambi i casi. La variabile “Sesso”, ad esempio, fornisce poche informazioni se è concordante, mentre fornisce una forte indicazione negativa sull'abbinamento se si osserva una discordanza.

b) Appare pertanto chiaro che, oltre alla scelta delle variabili di confronto, assume estrema importanza la definizione di concordanza che viene assegnata ad ogni possibile confronto fra le variabili. Al fine di sfruttare al meglio le informazioni provenienti dal confronto, sarebbe necessario tenere in considerazione tutte le possibili combinazioni di modalità che possono ottenersi quando si confronta la stessa variabile presente nei due archivi. E' evidente che ciò è tanto più difficile quanto maggiore è il numero di modalità. Sono pertanto necessarie delle modifiche nella definizione di concordanza che permette di aggregare quei risultati che forniscono informazioni simili sull'abbinamento. La dimensione di tale processo di aggregazione dipende essenzialmente dalla parsimonia richiesta al modello e dalle numerosità campionarie di cui si dispone.

c) Fra i metodi più utilizzati, il confronto può dare semplicemente un risultato dicotomico, con valori pari a 1 in caso di concordanza e a 0 in caso di discordanza fra le variabili. Per avere un maggior dettaglio, possono essere previsti diversi livelli di concordanza (ad esempio per l'età, tenendo conto della differenza in anni), o diverse capacità discriminanti per specifici valori delle variabili (ad esempio nel caso di cognomi più o meno comuni, per cui la concordanza di cognomi diffusi fornisce minori informazioni).

d) Una volta scelte le variabili di confronto e le definizioni di concordanza, l'informazione ottenibile per ogni coppia può essere riassunta in un vettore, che ha come singolo elemento il risultato del confronto fra le diverse variabili.

Il vantaggio di tale strategia è di ridurre, spesso drasticamente, il numero di confronti ammissibili; di contro, tale procedura può condurre ad un aumento di falsi negativi se non è elevata l'affidabilità delle variabili di confronto prescelte.

e) Definiti i vettori di confronto, rimane da stabilire come questi possano essere utilizzati per la decisione sulla classificazione delle coppie. Una possibilità è l'assegnazione ad ogni vettore di un peso probabilistico, sul valore del quale si basa il processo decisionale per l'abbinamento delle coppie.

Nella formulazione di Fellegi e Sunter (1969) i pesi vengono stimati in modo probabilistico, attraverso il rapporto fra le due verosimiglianze del vettore di confronto, rispettivamente nel caso di coppie relative allo stesso individuo e coppie abbinare casualmente.

Essi dimostrano che l'abbinamento è ottimale quando minimizza la regione di indecisione, ed ha come conseguenza, ad esempio, la possibilità di fissare a priori i livelli di errore desiderati, sia per quanto riguarda i falsi positivi che i falsi negativi, rendendo minimo il numero di coppie da abbinare in altra maniera (ad esempio, manualmente).

f) Al fine di valutare l'efficienza di un qualunque metodo di abbinamento di *record*, è cruciale disporre di stime del numero di errati abbinamenti e mancati abbinamenti che conseguono alla sua applicazione. Tali stime, fra l'altro, consentono di affrontare razionalmente il problema della scelta della soglia che consente di decidere quali coppie abbinare e quali no. La soglia deve essere determinata in un'ottica di minimizzazione degli errori, che devono essere stimati con precisione.

Un metodo spesso utilizzato per una prima approssimazione della soglia migliore, o per definire quali siano le coppie da verificare manualmente, consiste in un'analisi grafica della distribuzione dei pesi sull'intero campione. In questo modo, la regione di incertezza rimane nella zona in cui le code delle due distribuzioni condizionate si intersecano e gli errori dipendono dal punto esatto in cui si posiziona la soglia, con una conferma della relazione inversa fra le proporzioni di falsi positivi e falsi negativi (Garofalo et al., 2001).

Negli sviluppi successivi, altri autori hanno anche osservato che, quando vi è una probabilità elevata che una specifica unità di analisi sia presente una sola volta nell'altro *dataset* considerato per il *linkage*, allora la probabilità di un *matching* corretto cresce in maniera esponenziale (McLeod et al., 1998). In questo modo si è cercato di risolvere le problematiche legate all'eventuale presenza di informazioni di bassa qualità e non sempre omogenee nei *dataset* su cui effettuare il *linkage*.

4.4.2 La costruzione dei dataset finali

Come già riportato in precedenza, l'obiettivo di questa fase dello studio era di arrivare alla costruzione di due *dataset* finali, uno per i malati di tumore del colon-retto e uno

per le malate di tumore alla mammella, contenenti le informazioni provenienti dal Registro Tumori, dal Censimento fornito dall'Ufficio Statistico della Regione e dalle Anagrafi comunali.

La tabella 4.5 mostra una sintesi del lavoro di preparazione dei dati e di *record linkage*.

Tab. 4.5 – Le variabili identificative e l'unione dei casi.

| Fonte dati | Variabili identificative | Dati uniti ai dati del registro | % casi uniti al 1° passo | % casi uniti al 2° passo |
|-------------------|---|---------------------------------|---|-------------------------------------|
| Registro Tumori | Codice fiscale; data di nascita; indirizzo di residenza e sezione di Censimento di residenza di ogni individuo; codice numerico univoco e anonimo identificativo del database del Registro. | - | - | - |
| Censimento | Sezione di Censimento. | Sì | 100% dei casi. | - |
| Anagrafi comunali | Codice fiscale; data di nascita; indirizzo di residenza e sezione di censimento di residenza di ogni individuo. | Si | 98,7% dei casi (mammella) 97,2% dei casi (colon-retto) | 100% dei casi per entrambi i tumori |

La procedura di *record linkage* utilizzata è di tipo deterministico e ha permesso l'unione dei dati secondo diverse fasi, sviluppate fra Novembre e Dicembre 2014.

1) Il primo passo è stato unire i dati del Registro Tumori con i dati del Censimento; la chiave identificativa utilizzata è stata la sezione di Censimento, che ha permesso di collegare ad ogni individuo tutte le informazioni da Censimento relative alla propria sezione. Essendo stato completato in precedenza il controllo di qualità sulla variabile identificativa da parte del personale tecnico del Registro, è stato possibile abbinare tutti i casi alle informazioni da Censimento, ossia 8.317 casi per i pazienti di tumore alla mammella e 12.087 casi per i pazienti di tumore al colon-retto. Questa parte di procedura di *linkage* è stata sviluppata dall'autore di questo studio.

2) Il secondo passo ha riguardato l'abbinamento delle informazioni ricevute dalle Anagrafi comunali con i dati di Registro che erano già stati sottoposti al *record linkage* con i dati di Censimento. Questa fase è stata gestita dal personale del Registro, perché ha implicato l'uso di variabili identificative contenenti informazioni sensibili ai sensi

della normativa sulla *privacy*. I dati sono infatti stati collegati utilizzando, per il *merging*, contemporaneamente il codice fiscale, l'indirizzo di residenza, la data di nascita e la sezione di Censimento. Anche in questo caso le operazioni preventive di controllo della qualità dei dati relativi alle variabili identificative ha permesso di ridurre al minimo il numero di errori, ossia di casi per cui non è stato possibile effettuare l'abbinamento fra i due *dataset*. Più precisamente, sono stati collegati 8.209 casi per i casi con tumore alla mammella e 11.749 casi per i casi con tumore al colon-retto (per una perdita di casi pari a 1.3% e 2.8% rispettivamente). Si è cioè ottenuto un risultato decisamente buono sotto tutti i punti di vista, ma che è stato ulteriormente migliorato nella terza fase.

3) Il terzo passo è stato rivolto a identificare gli errori che hanno impedito il *linkage* completo dei casi e cercare di risolverli. Per fare questo, il personale del Registro ha operato un controllo caso per caso sugli individui non collegati, ricontrollando i dati relativi alle variabili identificative pervenuti dalle Anagrafi rispetto ai dati posseduti dal Registro, affiancando Nome e Cognome alle variabili utilizzate per il *linkage*. E' emerso che, malgrado i controlli di qualità eseguiti in precedenza, erano ancora presenti errori legati all'imputazione non omogenea dell'indirizzo di residenza e, in alcuni casi, ad un'imputazione errata del codice fiscale. Ad esempio, in un caso lo stesso individuo riportava come indirizzo VIA SANTA MARIA DELLE ROSE, 4 nelle informazioni da Registro e VIA S. MARIA DELLE ROSE nelle informazioni trasmesse dalle Anagrafi; oppure, un altro caso presentava due codici fiscali appena leggermente differenti: PRNMGH35S50G478P e PRMMGH35S50G478P (quest'ultimo errato).

Gli errori sono stati quindi corretti ed è stato quindi possibile inserire anche questi ultimi casi nei due *dataset* finali, ottenendo un abbinamento del 100% (e, quindi, di nuovo 8.317 casi per i pazienti di tumore alla mammella e 12.087 casi per i pazienti di tumore al colon-retto).

4) Il passo finale ha visto l'anonimizzazione dei dati, ottenuta associando ad ogni caso il codice numerico univoco predisposto dal Registro ed eliminando dai *dataset* tutte le informazioni sensibili che potessero far risalire all'identità degli individui coinvolti nello studio (nome, cognome, codice fiscale, indirizzo di residenza e numero di telefono). A questo punto il *dataset* è stato trasmesso all'autore di questo studio per poter procedere nelle analisi.

Il lavoro di controllo di qualità dei dati e di *record linkage* è stato sviluppato tramite *software* sviluppati internamente al Registro Tumori e tramite Microsoft Access 2010. I due *dataset* sono stati tradotti e trasmessi in formato IBM SPSS 19.0 e Stata 12.0.

Il risultato ottenuto nelle procedure di *record linkage* è ottimo, ma è frutto della particolare situazione fortunata in cui ci si è trovati ad operare.

Come detto, infatti, sia le fonti consultate che le procedure di controllo e unione dei dati fanno parte di un flusso informativo regionale consolidato.

E' anche fondamentale ricordare che, per questo motivo, non è stato necessario richiedere consensi ulteriori per il *record linkage*, in quanto la procedura si è configurata come aggiornamento del database del Registro Tumori all'interno dell'attività prevista dalle convenzioni che regolano il flusso informativo fra Comuni, Regione e membri della Rete Oncologica.

Queste peculiarità, positive per lo studio qui descritto, potrebbero quindi rendere problematico applicare le procedure di successo sopra descritte ad altre realtà territoriali.

Capitolo 5

I metodi di analisi

L'obiettivo principale di questo lavoro, come è stato più volte ripetuto, è verificare l'esistenza di un'associazione tra le caratteristiche della diagnosi e della presa in carico dei pazienti, delle loro caratteristiche socio-economiche, delle caratteristiche della sezione di Censimento di residenza e la sopravvivenza ai due tumori considerati.

Per fare questo si è cercato di applicare un modello lineare per il calcolo della sopravvivenza, che permettesse di utilizzare tutte le covariate di interesse, restituendone i coefficienti in termini di “*hazard ratio*”, ossia in termini di rapporto di coefficienti di rischio relativo (Dickman et al., 2004). Se, ad esempio, il rischio relativo associato alla covariata “Genere” è di 1,5 per gli uomini rispetto alle donne, vuol dire che il rischio di morte per gli uomini è 50% più alto che nelle donne.

Fra i vari modelli di sopravvivenza disponibili, basati su differenti procedure di calcolo e caratteristiche dei dati utilizzati, il modello più adeguato per questo studio è risultato essere quello della sopravvivenza relativa (Dickman et al., 2004; Lambert et al., 2005).

Questo capitolo, quindi, si propone di:

- definire la variabile dipendente utilizzata in questo studio, ossia la sopravvivenza relativa, specificando le ragioni che hanno portato alla scelta di tale variabile;
- presentare le caratteristiche e le procedure dei modelli lineari a parametri flessibili per il calcolo della sopravvivenza;
- descrivere il modello lineare multilivello ad effetti misti implementato in questo studio, che ha permesso di valutare gli effetti delle covariate descritte nel capitolo precedente,

Verranno anche indicate le modalità con cui, dopo l'analisi del modello generale, sono state analizzate le relazioni fra le variabili di tipo diagnostico, clinico e socio-economico che sono risultate avere un effetto sulla sopravvivenza.

5.1 La sopravvivenza relativa

La sopravvivenza relativa è la variabile generalmente scelta per stimare la sopravvivenza dei paziente neoplastici, quando si utilizzano i dati raccolti dai Registri Tumori in studi di popolazione (come è stato fatto in questa tesi) (Dickman et al., 2004 e 2006), anche se la sua utilità non è limitata a studiare la sopravvivenza ai tumori (Nelson et al., 2007).

Essa è, in sintesi, calcolata come il rapporto tra la sopravvivenza osservata nei pazienti (considerando tutti i decessi a prescindere dalla causa di morte) e la sopravvivenza attesa all'interno di una popolazione di cui tali pazienti fanno parte (considerando anche qui i decessi per tutte le cause di morte). In questo modo viene anche misurato l'eccesso di rischio di morte che si registra nella popolazione affetta dalla patologia neoplastica, ossia il reciproco della sopravvivenza relativa stessa.

Lo scopo degli studi di sopravvivenza basati sulla popolazione è di descrivere la sopravvivenza del paziente in gruppi demografici opportunamente definiti (ad esempio, per genere, età, stadio della malattia, ecc.), in modo tale che i risultati siano rappresentativi per tutta la popolazione di cui i soggetti in studio fanno parte.

Di conseguenza, gli studi di popolazione implicano un maggior numero di pazienti che sono seguiti per un periodo di tempo lungo. Questo può però causare problemi di affidabilità con la qualità dell'informazione relativa ad alcuni risultati clinici, in particolare l'eventuale causa di morte.

La descrizione di tale evento è affidata al medico che la registra e che, come è stato verificato da numerosi studi (ad esempio, Harteloh et al., 2010; Moussa et al., 1990; Nielsen et al., 1991), non sempre registra il tumore di cui soffriva il deceduto. Quindi, l'uso di metodi di calcolo della sopravvivenza basati sulla causa specifica di morte può essere problematico, in quanto le informazioni sulla causa della morte possono essere non affidabili, in particolare per quanto riguarda i pazienti affetti da neoplasie²⁷ (Fall et al., 2008; Percy et al., 1981).

²⁷ È importante ricordare che i tumori considerati in questo studio sono fra quelli che, più frequentemente, incorrono in questo tipo di problematica negli studi di popolazione, anche a causa della sopravvivenza medio-lunga (e, quindi, della maggiore durata del periodo osservato) che li caratterizza (Percy et al., 1981 e 1990).

Può accadere, ad esempio, che quando una metastasi al fegato del cancro del colon provoca il decesso del paziente, la causa di morte venga erroneamente registrato come cancro del fegato (Dickman et al., 2004). In un'analisi di sopravvivenza dei pazienti con diagnosi di tumore del colon basata sulle cause specifiche di morte, questa errata segnalazione sarà considerata come "morte per altre cause". Il tempo di sopravvivenza del caso sarà quindi considerato come "censurato" al momento della morte, ossia non sarà attribuito al tumore al colon che ha causato la metastasi, ingenerando quindi un errore nel calcolo del tempo di sopravvivenza stesso.

In più, anche se l'informazione sulla causa di morte neoplastica è accurata, può essere difficile determinare se una morte possa o meno essere imputata al cancro che si sta studiando (Percy et al., 1981 e 1990).

Ad esempio, non è evidente come debba essere qualificato un decesso dovuto agli effetti secondari delle cure. In un'analisi per cause specifiche, non è infatti possibile classificare una morte solo parzialmente come provocata dal cancro; le uniche due alternative sono o interamente dovuta al cancro di interesse o interamente dovuta ad altre cause (Dickman et al., 2004).

Per queste ragioni, è comune utilizzare la sopravvivenza relativa in studi di popolazione. Essa è generalmente stimata grazie alle tavole di mortalità della popolazione (Ederer et al., 1961), che sono messe a disposizione in maniera ufficiale dagli uffici statistici nazionali e locali (ad esempio, ISTAT, 2015), aggiornate al dato di mortalità verificato più recente²⁸.

Le tavole di mortalità sono, in genere, stratificate per causa di morte, sesso, età, anno solare e, dove l'informazione è considerata interessante a livello di salute pubblica, etnia (ad esempio, lo studio sulla sopravvivenza per tumore alla mammella sviluppato da McKenzie et al., 2011). Negli ultimi anni si stanno iniziando a stratificare le tavole di mortalità anche per status socio-economico, con tutte le problematiche che tale procedura comporta (Cancer Research UK Cancer Survival Group, 2009)²⁹.

Anche se queste tavole includono i decessi dovuti ai tumori specifici, in pratica questi non incidono sulle stime di sopravvivenza (Ederer et al., 1961; Estève et al., 1994), perché la mortalità per una causa specifica costituisce generalmente solo una piccola

²⁸ In Italia le Tavole di Mortalità ISTAT sono attualmente disponibili aggiornate al 2012.

²⁹ In Italia, le informazioni sull'etnia e sullo status socio-economico dei deceduti non sono informazioni disponibili per la stratificazione delle tavole di mortalità.

frazione della mortalità totale e correggere per questa mortalità ha un effetto trascurabile sulle stime di sopravvivenza, anche quando sono calcolate per tumori comuni, come i tumori del colon-retto o della mammella qui considerati.

I maggiori vantaggi della sopravvivenza relativa consistono quindi:

- nel fatto che l'informazione sulla causa di morte specifica non è richiesta;
- nel fornire una misura della mortalità in eccesso per i pazienti con diagnosi di tumore, indipendentemente dal fatto che l'eccesso di mortalità sia direttamente o indirettamente imputabile al cancro.

Vista la semplificazione che queste caratteristiche comportano, la sopravvivenza relativa è diventata l'indicatore ottimale per valutare questo esito sanitario negli studi oncologici di salute pubblica internazionali, nazionali e locali che utilizzano i dati dei Registri Tumori.

Esempi recenti e particolarmente rilevanti sono, ad esempio, lo studio Eurocare 5 che mette a confronto la sopravvivenza a 46 tumori in 29 nazioni (De Angelis et al., 2014), oppure lo studio sulla sopravvivenza in Italia, analizzata prendendo in considerazione tutti i casi di tumore diagnosticati fra il 1990 e il 2007 inseriti nei database dei Registri Tumori (AIRTUM Working Group, 2011).

5.1.1 Il modello di base della sopravvivenza relativa

Calcolare la sopravvivenza relativa vuol dire calcolare l'eccesso di rischio di morte legato alla diagnosi di cancro: sopravvivenza e rischio di morte sono infatti legati fra loro da una relazione logaritmica (Ederer et al., 1961; Dickman et al., 2013). Il modello di sopravvivenza relativa è infatti un modello additivo dei rischi di morte e, nella sua forma di base, può essere scritto come segue:

$$r(t; \mathbf{z}) = S(t; \mathbf{z}) / S^*(t; \mathbf{z}) \quad (1)$$

dove:

$S(t; \mathbf{z})$ = sopravvivenza osservata cumulata nei pazienti considerati al tempo t date le covariate \mathbf{z} ;

$S^*(t; \mathbf{z})$ = sopravvivenza attesa nella popolazione considerata, di cui fanno parte i pazienti, al tempo t date le covariate \mathbf{z} ;

t = tempo dalla diagnosi alla fine del periodo di osservazione (fine del follow-up);

\mathbf{z} = vettore delle covariate.

Data la relazione logaritmica che lega sopravvivenza e rischi di morte, al fine di calcolare l'eccesso di rischio di morte legato alla diagnosi di cancro la (1) può essere scritta in questo modo:

$$v(t; \mathbf{z}) = \lambda(t; \mathbf{z}) - \lambda^*(t; \mathbf{z}) \quad (2)$$

dove:

$v(t; \mathbf{z})$ = eccesso di rischio legato alla diagnosi di cancro,

$\lambda(t; \mathbf{z})$ = rischio osservato nei pazienti considerati;

$\lambda^*(t; \mathbf{z})$ = rischio atteso nella popolazione considerata, di cui fanno parte i pazienti, al tempo t date le covariate \mathbf{z} ;

La forma (2) dell'equazione è la base del modello di calcolo che è stato utilizzato in questa tesi per analizzare il ruolo delle covariate nei confronti del eccesso di rischio di morte legato alla diagnosi di tumore.

Vi sono però alcuni aspetti peculiari del presente studio da prendere in considerazione, che vanno oltre le caratteristiche del modello di base.

Innanzitutto, l'equazione (2) assume che gli eccessi di rischio fra i gruppi individuati dai diversi strati identificati dalle covariate nel vettore \mathbf{z} siano proporzionali nel tempo e che, di conseguenza, tali covariate non siano influenzate dal tempo di osservazione e non interagiscano tra di loro (Ederer et al., 1961; Estève et al., 1994).

In realtà, tale condizione non si verifica frequentemente e diverse situazioni possono manifestarsi.

Ad esempio, l'eccesso di rischio di morte può presentare variazioni legate all'età e all'anno di fine del *follow-up*, nonché risultare influenzato da come lo stadio alla diagnosi varia al variare dell'anno di fine *follow-up*, cosa che succede molto frequentemente quando si utilizzano dati di registro su intervalli di osservazione piuttosto estesi e, comunque, tendenzialmente superiori ai 3 anni (Dickman et al., 2004 e 2013).

Non considerare questo aspetto significa inserire un elemento di errore importante nella stima della sopravvivenza che può portare a sovrastime o sottostime significative e ad un'errata interpretazione del ruolo delle covariate, come è stato ad esempio dimostrato dallo studio di Bolard et al. (2001) su individui affetti da tumore al colon del Registro Tumori della Côte-d'Or.

Viste le caratteristiche dei dati utilizzati, si è quindi accettato l'ipotesi di eccessi di rischio non proporzionali, legati all'età e allo stadio alla diagnosi rispetto alla durata del periodo osservato (si ricorda che l'incidenza è stata osservata fra il 2001 e il 2010, gli esiti della malattia fino al 31/12/2012).

Un secondo aspetto peculiare, peraltro comune a tutti gli studi che analizzano se e come le covariate di varia natura influenzano la sopravvivenza relativa, è dato dal ruolo di tali variabili nel modello di calcolo.

Il vettore delle covariate \mathbf{z} , infatti, contiene solamente quelle utilizzate nella stratificazione delle tavole di mortalità, ossia sesso, età, stadio alla diagnosi, anno di fine *follow-up*. Le altre variabili diagnostiche e terapeutiche e le variabili socio-economiche vanno considerate aggiungendo un ulteriore membro all'equazione:

$$v(t; \mathbf{z}) = \lambda(t; \mathbf{z}) - \lambda^*(t; \mathbf{z}) + \exp(\mathbf{x}\beta) \quad (3)$$

dove il vettore \mathbf{x} rappresenta l'insieme delle covariate di cui sopra, mentre β rappresenta l'insieme dei loro coefficienti (Dickman, 2013).

Per sviluppare un'analisi in grado di valutare adeguatamente il ruolo delle covariate nel definire l'eccesso di rischio (e, quindi, la sopravvivenza relativa), ci si è rivolti ai modelli lineari di calcolo della sopravvivenza a parametri flessibili che permettono, allo stesso tempo, di gestire gli effetti di non proporzionalità³⁰.

Per raggiungere tale obiettivo, fra i diversi modelli a parametri flessibili si è quindi scelto il modello di sopravvivenza multilivello a parametri flessibili ed effetti misti (Crowther et al., 2014).

³⁰ Si ricorda, infatti, che l'obiettivo principale di questo studio non è calcolare la sopravvivenza relativa ai tumori presi in esame, bensì trovare se esiste una associazione statisticamente significativa fra la sopravvivenza e le variabili di tipo diagnostico-terapeutico e socio-economico presentate nel capitolo precedente.

5.2 Il modello di sopravvivenza relativa a parametri flessibili

I modelli di sopravvivenza a parametri flessibili sono modelli lineari che utilizzano funzioni a *spline* cubici ristretti per modellare il rischio cumulativo di base, la probabilità cumulativa di sopravvivenza di base, o altre distribuzioni di base in modelli di analisi di sopravvivenza.

Gli *spline* sono funzioni matematiche flessibili, definite da polinomi applicati a sotto-intervalli che compongono l'intervallo di osservazione; i punti in cui i polinomi si uniscono sono definiti "nodi". Tali polinomi sono condizionati da vincoli che garantiscono il "lisciamento" della curva stimata (ad esempio, Lambert et al., 2009), riducendo quindi il problema di picchi e irregolarità legati a situazioni contingenti (Lambert et al., 2005).

Le funzioni *spline* usate più comunemente sono le *spline* cubiche, che possono essere incorporate in qualsiasi modello di regressione che preveda almeno un predittore lineare.

In origine, questi modelli consentivano solo l'utilizzo dei rischi proporzionali, ma sono stati successivamente estesi anche ai rischi non proporzionali, incorporando gli effetti dipendenti dal tempo sulle covariate (Lambert et al., 2009). Per questa ragione, nel presente studio è stato necessario utilizzare un modello appartenente a questa famiglia, che offre la possibilità di modellare gli effetti dipendenti dal tempo e calcolare gli effetti delle covariate, tenendo conto della collinearità tra variabili.

Più precisamente, quando si intende utilizzare un modello che stimi gli effetti non proporzionali, ci si rivolge ad un particolare gruppo di funzioni *spline* cubiche: le cosiddette funzioni *spline* cubiche limitate o con restrizione (Lambert et al., 2009). Tale restrizione consiste nel forzare la funzione modellata ad essere lineare prima del primo nodo e dopo l'ultimo nodo. Da un punto di vista matematico, ciò si esprime dicendo che la funzione modellata tramite *spline* cubici limitati a K nodi è implementata creando $K-1$ variabili derivate:

$$s(x) = \gamma_0 + \gamma_1 z_1 + \gamma_2 z_2 + \dots + \gamma_{K-1} z_{K-1} \quad (4)$$

dove:

$$z_1 = x;$$

$$z_j = (x - k_j)^3 - \varphi_j(x - k_1)^3 - (1 - \varphi_j)(x - k_K)^3;$$

$$j = 2, \dots, K - 1;$$

$$\varphi_j = (k_K - k_j)/(k_K - k_1)$$

L'equazione (4) è di fondamentale importanza, perché permette di passare a modelli parametrici di stima dei rischi non proporzionali. Grazie all'equazione (4), infatti, la funzione di base per i rischi non proporzionali può essere scritta nella seguente maniera³¹ (equazione di Lambert e Royston):

$$\ln \{H_i(t | \mathbf{x}_i)\} = s\{\ln(t) | \gamma, \mathbf{k}_0\} + \sum_{j=1}^D s\{\ln(t) | \delta_k, \mathbf{k}_j\} x_{ij} + \mathbf{x}_i \beta \quad (5)$$

Nell'equazione (5):

- $s\{\ln(t) | \gamma, \mathbf{k}_0\}$ rappresenta gli (eventuali) effetti delle covariate stratificanti al tempo di inizio dell'osservazione;
- $\sum_{j=1}^D s\{\ln(t) | \delta_k, \mathbf{k}_j\} x_{ij}$ rappresenta gli effetti non proporzionali delle covariate stratificanti;
- $\mathbf{x}_i \beta$ rappresenta gli effetti di tutte le altre covariate.

In termini di sopravvivenza relativa, la (5) diventa:

$$R(t) = \exp[-\exp(\ln \{H_i(t | \mathbf{x}_i)\})] \quad (6)$$

che è l'equazione che permette di calcolare la sopravvivenza relativa a partire dalla funzione di calcolo dei rischi di morte non proporzionali.

Questo approccio a parametri flessibili di calcolo della sopravvivenza, basato sul calcolo dei rischi cumulativi di morte, presenta il vantaggio di una maggiore stabilità del modello, che vuol dire una maggiore attendibilità nel calcolo della sopravvivenza e una migliore lettura del ruolo delle covariate grazie ad una migliore stima dei coefficienti (in termini di *hazard ratio*), come è stato dimostrato ad esempio dallo studio di Remontet al. (2007) che ha comparato diversi modelli di calcolo della sopravvivenza relativa applicati ad un'ampia casistica di pazienti di tumore catalogati nei database dei Registri Tumori francesi.

³¹ Per la dimostrazione puntuale, qui non riportata per comodità di lettura, vedi ad esempio Lambert et al. (2005 e 2009); Nelson et al. (2007).

Il risultato finale di questo studio francese ha permesso di evidenziare come il modello ottenuto dall'applicazione delle equazioni (5) e (6) abbia fornito le stime di sopravvivenza più attendibili e, contemporaneamente, di modellizzare accuratamente gli effetti delle diverse covariate utilizzate nel modello.

Rispetto agli obiettivi di questa tesi, però, l'equazione (5) presenta un limite: permette di valutare solo gli effetti fissi delle covariate e non gli effetti casuali, legati a condizioni particolari come, nel nostro caso, l'appartenenza dei pazienti alla sezione di Censimento di residenza che potrebbe creare un effetto di *clustering* nei casi in studio (Jin et al., 2003).

Questo effetto di *clustering* è di tipo ecologico ed è un confondente che deve essere tenuto sotto controllo se si vuole valutare più accuratamente l'effetto delle covariate individuali sulla salute degli individui (Goldstein, 1995).

Altri studi, che hanno analizzato come le disuguaglianze socio-economiche influiscono sullo stato di salute degli individui, hanno affrontato tale problematica rivolgendosi ai modelli lineari multivello per gestire l'effetto dell'area di residenza.

In Italia si può ad esempio ricordare il lavoro di Biolcati Rinaldi et al. (2013), che ha indagato lo stato di salute percepito dal campione di intervistati appartenenti a tre edizioni dell'Indagine ISTAT Multiscopo³², residenti in Lombardia, a confronto con il totale del campione italiano. In questo lavoro, l'effetto di *clustering* dovuto al comune di residenza è analizzato tramite un modello di regressione lineare multilivello, implementato nel programma MLwiN. Tale modello non permette però di stimare gli effetti multilivello quando la variabile dipendente è la sopravvivenza.

Per tenere conto di questa condizione, nella presente ricerca si è deciso di applicare un modello lineare multilivello a parametri flessibili ed effetti misti per il calcolo della sopravvivenza, sviluppato recentemente per risolvere questo tipo di problema da Crowther et al. (2014).

³² Indagine ISTAT Multiscopo sulle famiglie "Aspetti della vita quotidiana".

5.2.1 Il modello multilivello a parametri flessibili ed effetti misti per il calcolo della sopravvivenza relativa

La presenza di effetti casuali, generati da una situazione di *clustering* dei dati, è abbastanza comune quando si affrontano studi su base di popolazione, sia nelle scienze sociali che nelle scienze mediche. Nel caso specifico, in questo studio si è in presenza di dati che sono clusterizzati per sezione di Censimento di residenza dei pazienti, poiché le stesse sezioni di Censimento forniscono le informazioni relative alle caratteristiche socio-economiche di tali aree, attribuite ad ogni singolo paziente (Crowther et al., 2014; Jin et al., 2003).

Si è quindi reso necessario applicare un modello di analisi che tenesse in considerazione gli effetti casuali della distribuzione spaziale degli individui considerati. Il modello di Crowther et al. (2014) è stato sviluppato proprio per risolvere questo tipo di questione. Esso, infatti, utilizza il metodo di quadratura di Gauss-Hermite (in merito, vedi anche Liu et al., 2008) per permettere all'equazione di Lambert-Royston di calcolare sia gli effetti fissi delle covariate, sia gli effetti casuali. In questo modo, il modello additivo di calcolo dei rischi, espresso dalla (5), si trasforma in questa maniera:

$$\ln \{H_i(t | \mathbf{x}_i)\} = s\{\ln(t) | \gamma, \mathbf{k}_0\} + \sum_{j=1}^D s\{\ln(t) | \delta_k, \mathbf{k}_j\} x_{ij} + \mathbf{x}_i\beta + \mathbf{z}_i\mathbf{b}_i \quad (7)$$

dove:

$\mathbf{z}_i\mathbf{b}_i$ = espressione degli effetti casuali.

L'integrazione di tale espressione avviene grazie ad una specifica formulazione del calcolo della verosimiglianza, ossia dell'indicatore che permette di verificare la bontà di adattamento di un modello di stima, applicata alla distribuzione in *cluster* dei casi.

Crowther et al. riportano la seguente formulazione per il calcolo della verosimiglianza applicato all'*i*-esimo *cluster*:

$$L_i = \int_{-\infty}^{\infty} \left[\prod_{j=1}^{n_i} p(T_{ij}, d_{ij} | \mathbf{b}_i, \theta) \right] p(\mathbf{b}_i | \theta) d\mathbf{b}_i \quad (8)$$

Fonte: Crowther et al., 2014.

dove θ è il vettore complessivo delle covariate, incluso il termine casuale.

All'equazione (8) viene applicata la quadratura di Gauss-Hermite, ossia una formula di quadratura di massimo grado di precisione in grado di approssimare l'integrale (Liu et al., 2008; Rabe-Hesketh et al., 2002), permettendone l'implementazione nel modello.

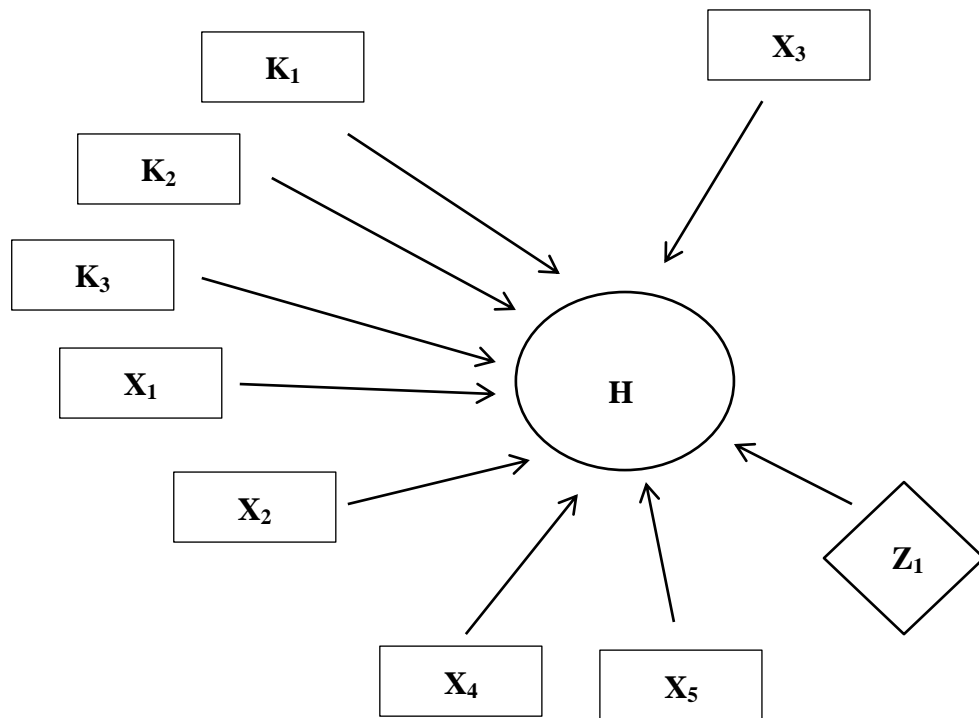
Il *software* necessario per applicare questo modello è attualmente disponibile per Stata come pacchetto da installare sotto il nome di STMIXED (Stata Corp., 2014), sviluppato dagli autori del lavoro sopracitato e verificato dal Dipartimento di Epidemiologia Medica e Biostatistica del Karolinska Institutet (Stoccolma) e dal Dipartimento di Scienze della Salute dell'Università di Leicester.

Il lavoro di questa tesi ha utilizzato la versione più recente di tale software (STMIXED v1.1.0) e ha contribuito a perfezionarlo e a valutarne la capacità di utilizzo su *dataset* di grosse dimensioni, come quelli utilizzati per studi di popolazione su dati provenienti da Registri Tumori.

5.3 Il modello analizzato

Il modello multilivello a parametri flessibili ed effetti misti è stato declinato in questo studio secondo la struttura presentata nella figura 5.1, dove sono indicati gli effetti principali fissi e casuali delle covariate sul calcolo dei rischi di morte.

Fig. 5.1 – Il modello analizzato: gli effetti fissi.



Più precisamente, i singoli simboli rappresentano:

le variabili indipendenti, di cui si studiano gli effetti fissi:

K₁ = età alla diagnosi; variabile stratificante con effetto non proporzionale;

K₂ = stadio della malattia; variabile stratificante con effetto non proporzionale;

K₃ = sesso; variabile stratificante con effetto proporzionale³³;

X₁ = gruppo delle variabili diagnostiche, fra cui il grado di malignità;

X₂ = gruppo delle variabili relative alle terapie applicate;

X₃ = presenza di comorbidità alla diagnosi;

X₄ = gruppo delle variabili socio-economiche individuali;

X₅ = gruppo delle variabili socio-economiche relative alla sezione di censimento di residenza;

la variabile indipendente, responsabile degli effetti casuali:

Z₁ = sezione di Censimento di residenza;

³³ Per il tumore della mammella, avendo considerato solamente il genere femminile, la variabile non è stata utilizzata.

la variabile dipendente:

H = rischio di morte legato al tumore.

Oltre agli effetti principali, per la parte relativa agli effetti fissi si sono anche calcolati gli effetti di interazione fra le covariate sulla variabile dipendente. Più precisamente, basandosi sulla letteratura in materia che ha già trattato in questo modo il ruolo delle covariate rispetto alla sopravvivenza (ad esempio, Chang et al., 2005; oppure Polednak et al., 2002), sono state analizzate le seguenti interazioni a due livelli:

K₁ # K₂ K₁ # X₁ K₁ # X₂ K₁ # X₃ K₁ # X₄
K₂ # X₂ K₂ # X₄
X₁ # X₂ X₁ # X₄ X₁ # X₅
X₂ # X₃ X₂ # X₄ X₂ # X₅
X₃ # X₄
X₄ # X₅

E, solo per i tumori del colon-retto:

K₁ # K₃ K₃ # X₁ K₃ # X₂ K₃ # X₃ K₃ # X₄

Non è stato possibile analizzare gli effetti delle interazioni a livelli superiori, a causa dell'eccessiva richiesta di memoria necessaria all'elaboratore per poter eseguire tutti i calcoli necessari.

Non sono state considerate le interazioni a due o più vie con la variabile responsabile degli effetti casuali, perché al momento tale funzione non è disponibile in STMIXED o in altri software per l'analisi della sopravvivenza relativa (almeno al momento in cui sono state eseguite le analisi per questa tesi).

Per quanto riguarda la collinearità fra le variabili dipendenti, essa è stata controllata automaticamente all'interno dell'analisi sviluppata da STMIXED, che esegue un controllo di tolleranza in cui vengono accettate nel modello solo le variabili che superano tale controllo (l'intervallo di confidenza per tale verifica è posto automaticamente al 95%).

L'implementazione del modello è stata eseguita tramite un processo additivo dei gruppi di variabili in diversi passi, secondo un criterio proposto da Wang et al. (2015). Tale processo è presentato nella tabella 5.1. Più precisamente, sono stati valutati 24 modelli differenti per ogni sede di tumore, coprendo tutte le possibili permutazioni fra i diversi gruppi di variabili.

Per poter scegliere il modello con il miglior adattamento si è utilizzato il Criterio Informativo di Akaike (AIC; Akaike, 1974 e 1985; Sugiura, 1978). Tale Criterio permette di confrontare diversi modelli statistici riferiti ad una stessa variabile dipendente e fornisce una misura della qualità della stima di ogni modello confrontato, prendendo in considerazione sia la bontà dell'adattamento, sia la complessità del modello, espressa dal numero di covariate. Il valore AIC più basso identifica il modello migliore.

Negli studi che prendono in considerazione più modelli di calcolo della sopravvivenza, l'AIC è stato utilizzato sia per scegliere il miglior modello fra modelli calcolati con lo stesso metodo ma con variabili e condizioni differenti³⁴ (ad esempio, Crowther et al., 2014), sia per valutare modelli basati su metodi differenti (ad esempio, Remontet et al., 2007).

³⁴ Come nel caso di questo studio.

Tab. 5.1 – Schema del processo di inserimento dei gruppi di variabili.

| Età | Stadio | Variabili diagnostiche | Comorbidità | Variabili terapeutiche | Effetti fissi Caratteristiche socio-economiche individuali | Caratteristiche socio-economiche Sezione di Censimento | Effetti casuali Sezione di Censimento di residenza |
|-----|--------|------------------------|-------------|------------------------|---|--|---|
| X | X | | | | | | X |
| X | X | | | | X | | X |
| X | X | | | | | X | X |
| X | X | X | | | | | X |
| X | X | X | X | | | | X |
| X | X | X | | | X | | X |
| X | X | X | | | | X | X |
| X | X | X | X | | X | | X |
| X | X | X | X | | | X | X |
| X | X | | | | X | X | X |
| X | X | X | | | X | X | X |
| X | X | X | X | | X | X | X |
| X | X | | | X | | | X |
| X | X | | | X | X | | X |
| X | X | | | X | | X | X |
| X | X | X | | X | | | X |
| X | X | X | X | X | | | X |
| X | X | X | | X | X | | X |
| X | X | X | | X | | X | X |
| X | X | X | X | X | X | | X |
| X | X | X | X | X | | X | X |
| X | X | X | | X | X | X | X |
| X | X | X | X | X | X | X | X |

5.3.1 Interventi tecnici sulle variabili utilizzate nel modello

Le analisi sono state eseguite presso il Dipartimento di Epidemiologia Medica e Biostatistica, sotto la supervisione del prof. Paul Dickman e con il supporto del gruppo di Epidemiologia Oncologica del Dipartimento (in particolare il prof. Pär Sparen).

Questo supporto ha permesso di effettuare scelte di analisi che hanno tenuto conto non solo degli aspetti statistici e metodologici, ma anche di aspetti di tipo medico e biologico, in modo da tale da avere una maggiore attendibilità nei risultati finali.

Una prima scelta importante è stata fatta rispetto alle sedi di tumore considerate: le analisi per il colon-retto sono state separate nelle due differenti sedi per il colon e il retto, poiché le due sedi specifiche si differenziano, da un punto di vista clinico per parecchi aspetti che riguardano in particolare lo stadio alla diagnosi, il grado di malignità e le terapie utilizzate. Anche se in moltissimi studi queste due sedi sono analizzate assieme (come mostrato nel capitolo 2 di questa tesi), quando si affronta uno studio che cerca di analizzare in maniera approfondita le determinanti della sopravvivenza, scendendo al livello individuale degli approcci di cura e delle caratteristiche socio-economiche, si preferisce lavorare separatamente sulle due sedi (Hong et al., 2012).

Da un punto di vista del Criterio Internazionale Standard per la Classificazione delle Malattie (Ministero della Salute, 2001), le sedi di tumore sono state così raggruppate (Tab. 5.2):

Tab. 5.2 – Sedi di tumore per codici ICD-10-

| Sede del tumore | Codice ICD-10 |
|-----------------|------------------|
| Mammella | C50 |
| Colon | Da C18.0 a C18.9 |
| Retto | C19; C20; C21 |

Una seconda riflessione importante ha riguardato le caratteristiche del comportamento dei tumori rispetto a come sono state classificate nel *dataset* messo a disposizione dal Registro Tumori.

Più specificatamente, per tutte le sedi considerate erano presenti tumori classificati come maligni, altri come benigni e altri dal comportamento incerto. Viste le finalità

dello studio, si è deciso di eliminare dall'analisi i casi relativi a tali tumori, per non incorrere in problemi di sopravvivenza troppo alta, perché comprendente anche casi non a rischio di morte per il tumore.

Anche rispetto alle variabili di stadio della malattia alla diagnosi e di grado di malignità sono state fatte alcune scelte importanti, orientate ad eliminare potenziali distorsioni nel calcolo della sopravvivenza. Sono, infatti, stati eliminati dall'analisi i casi in cui lo stadio del tumore era indicato come "Non noto" nel *datataset* originale.

Gli effetti di queste esclusioni, in termini di numero assoluto di casi esclusi dall'analisi per sede di tumore, saranno mostrati nel prossimo capitolo, in cui saranno presentati i risultati di tutte le analisi effettuate.

Un lavoro di ricodifica è stato, invece, fatto per quanto riguarda l'età degli individui considerati e la presenza di comorbidità alla diagnosi.

Per quanto riguarda l'età, la variabile è stata riclassificata in tre gruppi: 0-49 anni, 50-69 anni, 70 e più anni. Tali gruppi hanno, infatti, un significato ben preciso per la diagnosi precoce dei tumori.

Il primo gruppo, 0-49 anni è composto da individui considerati giovani, in cui il tumore è tendenzialmente più raro e per i quali non è previsto l'inserimento nei programmi di screening organizzato.

Il secondo gruppo, 50-69 anni, contiene coloro per cui è prevista la partecipazione ai programmi di screening organizzato, attivi in Umbria per entrambi i tumori.

Il terzo gruppo, 70 e più anni, è composto dalla parte anziana dei pazienti, non compresi nei programmi di screening organizzato e per i quali l'incidenza di questi tumori è generalmente più alta rispetto al resto della popolazione, come ampiamente osservato nei capitoli precedenti.

Rispetto alle comorbidità, è stata ricostruita una variabile dicotomica che registra la presenza/assenza di comorbidità al momento della diagnosi.

5.4 Le analisi post-modello

Lo scopo di tutta l'analisi precedente è di individuare se, dopo il controllo tramite tutte le variabili indipendenti di tipo medico, le variabili che descrivono le caratteristiche

socio-economiche individuali e relative alla sezione di Censimento esprimono ancora un effetto statisticamente significativo.

Questa situazione è stata verificata ed è noto che questi effetti sulla sopravvivenza al cancro sono soprattutto indiretti, perché condizionano aspetti di tipo clinico, come la diagnosi precoce della malattia, la capacità del malato di seguire in maniera ottimale le cure o l'aver a disposizione opportunità diagnostiche o terapeutiche adeguate (in merito, si rimanda al capitolo 2 di questa tesi, ma anche ad Adler et al., 1999; Galobardes et al., 2007).

Sono quindi state effettuate una serie di analisi di regressione logistica in cui le variabili socio-economiche, che nei modelli di sopravvivenza avevano coefficienti statisticamente significativi, sono state messe a confronto come indipendenti rispetto alle singole variabili di tipo diagnostico e terapeutico, considerate le variabili dipendenti. Di tali regressioni sono stati valutati gli effetti principali delle covariate sulla variabile indipendente, calcolandone gli odds ratio. L'intervallo di confidenza è stato posto al 95%.

In questo modo si è potuto individuare su quali aspetti del percorso di diagnosi e cura, le variabili socio-economiche producevano un effetto statisticamente significativo.

Tutte le analisi sono state effettuate con software statistico Stata 13.0; come già ricordato in precedenza, per l'analisi dei modelli di sopravvivenza relativa multilivello a parametri flessibili è stato utilizzato il modulo aggiuntivo STMIXED, disponibile per il download presso l'archivio della Stata Corporation nella versione 1.1.0, utilizzata in questa tesi.

Capitolo 6

I Risultati

I risultati ottenuti dal percorso di analisi sono presentati separatamente per le tre sedi di tumore. Per ognuna di esse sono riportati:

- i risultati della selezione dei casi utilizzati nei modelli a parametri flessibili ed effetti misti;
- le caratteristiche di base delle variabili utilizzate nei modelli;
- i risultati dei modelli;
- i risultati delle analisi di regressione logistica.

6.1 Il tumore alla mammella

Partendo da un totale di 8.317 pazienti, solo 5.379 sono state inserite nell'analisi di sopravvivenza. Gli altri 2.938 casi sono stati esclusi in quanto affetti da tumori benigni, o colpiti da tumore maligno per cui lo stadio alla diagnosi non è disponibile, o perché non è stato possibile ricostruire tutte e tre le informazioni socio-economiche individuali richieste alle Anagrafi comunali.

La tabella 6.1 mostra i risultati di tale selezione.

Tab. 6.1 – Casi esclusi e motivazioni.

| Ragioni dell'esclusione | Casi esclusi | % |
|---|---------------------|-------------|
| Casi con tumore a comportamento benigno | 1.267 | 15,2 |
| Casi con stadio non definito o non disponibile | 1.472 | 17,7 |
| Casi per cui è mancante almeno una informazione fra stato civile, titolo di studio, professione | 199 | 2,4 |
| <i>Totale casi esclusi</i> | <i>2.938</i> | <i>35,3</i> |
| Casi mantenuti nell'analisi | 5.379 | 64,7 |
| Totale casi | 8.317 | 100,0 |

Come si nota, il 15% circa dei casi esclusi sono pazienti che presentano tumori a comportamento di tipo benigno, ossia non a rischio di morte per tale causa oncologica. Tali casi sono comunque di interesse per altre analisi di tipo epidemiologico e socio-

economico e fanno regolarmente parte del database di ogni Registro Tumori, in cui devono comunque essere riportati in quanto patologia che deve essere curata.

Le informazioni delle pazienti con tumore benigno non vengono però utilizzate quando viene studiata la sopravvivenza, perché ovviamente possono incrementarla in maniera fuorviante rispetto alla patologia neoplastica maligna e fornire informazioni fortemente contraddittorie quando si studiano le associazioni fra sopravvivenza e covariate. Pertanto esse sono state escluse dal presente studio.

Le pazienti in questa condizione sono comunque state inserite nella raccolta dati, per ricostruire un *dataset* di informazioni completo, i cui casi corrispondono a quelli messi originariamente a disposizione dal Registro Tumori.

In questo modo anche i casi benigni hanno partecipato alle operazioni di *record linkage* e verifica del dato, contribuendo a dare una maggiore robustezza in termini di numerosità dei *record* alla valutazione di validità e attendibilità di tali operazioni, di cui si è discusso nel capitolo 4 di questa tesi.

Un 18% circa di casi esclusi, invece, è composto da pazienti affette da tumore maligno della mammella per i quali, però, tutta la documentazione fornita dalle diverse fonti diagnostiche non ha permesso di avere una classificazione precisa dello stadio, più precisamente ha portato all'informazione "Stadio indeterminato".

Anche in questo caso, per i Registri Tumori tali individui sono molto rilevanti per altri studi che non siano quelli di sopravvivenza, dove il loro inserimento potrebbe provocare risultati non correttamente interpretabili³⁵. Si è scelto di inserirli comunque nel percorso di raccolta dati per le stesse ragioni relative ai casi con tumore benigno³⁶.

Va osservato che la percentuale di tumori maligni dallo stadio indefinito è in linea con i criteri di alta qualità del dato richiesti dall'Associazione Registri (Ferretti et al., 2007).

Infine, circa il 2% dei casi sono stati esclusi perché, nel processo di raccolta delle informazioni da parte delle anagrafi comunali, non è stato possibile recuperare tutte e tre le informazioni relative a stato civile, titolo di studio, condizione professionale. Poiché l'analisi di sopravvivenza richiede che nelle covariate utilizzate non vi siano dati mancanti, i casi in tale condizione sono stati esclusi.

³⁵ Si ricorda che lo stadio alla diagnosi è uno dei principali determinanti della sopravvivenza. Uno stadio classificato come indeterminato può, quindi, essere un qualsiasi stadio e, di conseguenza, un andamento di sopravvivenza non può essere letto correttamente se il modello di calcolo comprende anche pazienti così classificati.

³⁶ Queste considerazioni e le precedenti sono valide anche per i casi affetti da neoplasia al colon-retto.

6.1.1 Le caratteristiche di base

La tabella 6.2 presenta le caratteristiche demografiche e socio-economiche individuali delle persone che sono state utilizzate nell'analisi finale.

Tab. 6.2 – Le caratteristiche demografiche e socio-economiche individuali delle pazienti.

| Caratteristiche individuali | | N | % |
|-----------------------------|--|--------------|------|
| Età all'incidenza | 0-49 | 1.150 | 21,4 |
| | 50-69 | 2.607 | 48,5 |
| | 70+ | 1.622 | 30,2 |
| Stato civile | Celibe/nubile | 514 | 9,6 |
| | Coniugato/a o convivente | 3.873 | 72,0 |
| | Separato/a o divorziato/a | 197 | 3,7 |
| | Vedovo/a | 795 | 14,8 |
| Titolo di studio | Licenza elementare o analfabeta | 2.562 | 47,6 |
| | Licenza media inferiore | 930 | 17,3 |
| | Licenza media superiore | 1.219 | 22,7 |
| | Laurea | 668 | 12,4 |
| Condizione professionale | Legislatore, dirigente o imprenditore | 13 | 0,2 |
| | Professione intellettuale, scientifica o di elevata specializzazione | 54 | 1,0 |
| | Professione qualificata nelle attività commerciali o nei servizi | 695 | 12,9 |
| | Professione tecnica | 460 | 8,6 |
| | Impiegato | 179 | 3,3 |
| | Artigiano, operaio specializzato o agricoltore | 220 | 4,1 |
| | Professione non qualificata | 669 | 12,4 |
| | Pensionato | 303 | 5,6 |
| | Disoccupato | 1.088 | 20,2 |
| | Casalinga | 1.698 | 31,6 |
| Totale | | 5.379 | |

Rispetto alle età, i casi analizzati sono per il 48% circa nella fascia compresa fra i 50 e i 69 anni, ossia nella fascia di età interessata dalle campagne per la prevenzione secondaria e la diagnosi precoce, mentre il 30% circa è formato da anziane (70 e più anni) e il 21% circa da giovani (0-49 anni). Rispetto ai dati italiani per lo stesso periodo (AIRTUM, 2014), si osserva una minor percentuale di anziane (34% circa in Italia) e

una percentuale sostanzialmente identica di giovani (22% circa in Italia), ma una percentuale più elevata di donne malate nel gruppo di mezza età (44% circa in Italia).

Le pazienti sono per lo più coniugate (o conviventi), seguite dalla percentuale di vedove, nubili e separate/divorziate. Rispetto al dato dell'intera popolazione umbra (dati ISTAT, 2001), si nota una distribuzione simile per quanto riguarda le vedove (14% circa in tutta la regione) e le separate/divorziate (2% in tutta la regione), ma una percentuale decisamente più alta di coniugate/conviventi (49% in tutta la regione) e più bassa di nubili (35% in tutta la regione).

Rispetto al titolo di studio, si osserva una presenza elevata di titoli elementari, seguiti dai diplomi di scuola superiore, licenze medie inferiori e lauree. Le percentuali dei titoli inferiori sono più alte dei valori regionali e nazionali, mentre le lauree sono simili ai livelli regionali e nazionali. Più basse, quindi, le percentuali dei diplomi di media superiore (Vercelli et al., 2013).

Le donne considerate nell'analisi sono per lo più casalinghe e disoccupate (complessivamente il 52% circa) o impiegate in professioni non qualificate (12% circa) e pensionate (6% circa).

La tabella 6.3 mostra le distribuzioni combinate età-stato civile ed età-titolo di studio; la tabella 6.4 mostra la distribuzione combinata titolo di studio-condizione professionale; la loro connessione è stata verificata tramite test del χ^2 (significatività statistica del test posta a $p < 0.05$).

Tab. 6.3 – Stato civile e titolo di studio per età alla diagnosi.

| | | 0-49 | | 50-69 | | 70+ | | Sig. χ^2 |
|------------------|---------------------------------|------|------|-------|------|-------|------|---------------|
| | | N | % | N | % | N | % | |
| Stato civile | Nubile | 239 | 20,8 | 173 | 6,6 | 102 | 6,3 | p<0,05 |
| | Coniugata o convivente | 825 | 71,7 | 2.128 | 81,6 | 920 | 56,7 | |
| | Separata o divorziata | 69 | 6,0 | 86 | 3,3 | 42 | 2,6 | |
| | Vedova | 17 | 1,5 | 220 | 8,4 | 558 | 34,4 | |
| Titolo di studio | Licenza elementare o analfabeta | 168 | 14,6 | 1.136 | 43,6 | 1.258 | 77,6 | p<0,05 |
| | Licenza media inferiore | 300 | 26,1 | 551 | 21,1 | 79 | 4,9 | |
| | Licenza media superiore | 443 | 38,5 | 612 | 23,5 | 164 | 10,1 | |
| | Laurea | 239 | 20,8 | 308 | 11,8 | 121 | 7,5 | |

Tab. 6.4 – Condizione professionale per titolo di studio.

| | | Licenza elementare o analfabeta | | Licenza media inferiore | | Licenza media superiore | | Laurea | | Sig. χ^2 |
|--------------------------|--|---------------------------------|------|-------------------------|------|-------------------------|------|--------|------|---------------|
| | | N | % | N | % | N | % | N | % | |
| Condizione professionale | Legislatore, dirigente o imprenditore | 0 | 0,0 | 0 | 0,0 | 12 | 1,0 | 1 | 0,1 | p<0,05 |
| | Professione intellettuale, scientifica o di elevata specializzazione | 4 | 0,2 | 1 | 0,1 | 17 | 1,4 | 32 | 4,8 | |
| | Professione qualificata nelle attività commerciali o nei servizi | 182 | 7,1 | 9 | 1,0 | 35 | 2,9 | 469 | 70,2 | |
| | Professione tecnica | 166 | 6,5 | 159 | 17,1 | 102 | 8,4 | 33 | 4,9 | |
| | Impiegato | 10 | 0,4 | 16 | 1,7 | 141 | 11,6 | 12 | 1,8 | |
| | Artigiano, operaio specializzato o agricoltore | 151 | 5,9 | 22 | 2,4 | 46 | 3,8 | 1 | 0,1 | |
| | Professione non qualificata | 420 | 16,4 | 61 | 6,6 | 181 | 14,8 | 7 | 1,0 | |
| | Pensionato | 198 | 7,7 | 21 | 2,3 | 80 | 6,6 | 4 | 0,6 | |
| | Disoccupato | 386 | 15,1 | 519 | 55,8 | 147 | 12,1 | 36 | 5,4 | |
| Casalinga | 1.045 | 40,8 | 122 | 13,1 | 458 | 37,6 | 73 | 10,9 | | |

Stato civile e titolo di studio si distribuiscono secondo un andamento connesso ai gruppi di età: nubili e separate/divorziate decrescono al crescere dell'età, mentre le vedove ovviamente aumentano; il livello del titolo di studio decresce al crescere dell'età.

Per quanto riguarda il rapporto fra titolo di studio e condizione professionale, si nota anche qui un andamento ben noto, in cui il livello professionale cresce al crescere del livello di istruzione.

Per descrivere le caratteristiche della malattia, della diagnosi e della cura, sono state presentate nella tabella 6.5 le variabili che, in letteratura, sono più frequentemente utilizzate e che sono state inserite nell'analisi di sopravvivenza. La stessa tabella mostra anche la loro distribuzione nei tre gruppi di età³⁷.

Le pazienti considerate nello studio presentano in genere uno stadio piuttosto precoce della malattia (stadio I o II, complessivamente 83% circa), ossia situazioni di malattia in area anatomica contenuta, con interessamento nullo o scarso dei linfonodi e una prognosi più favorevole. Decisamente più bassa la presenza di neoplasie in stadio avanzato (III) o metastatiche (IV), che hanno una prognosi sfavorevole o pessima. Mentre lo stadio IV segue un andamento crescente al crescere dell'età, lo stadio III

³⁷ Come abbiamo ricordato ripetutamente nei capitoli precedenti, è stata ampiamente verificata in letteratura la relazione fra l'età del malato, le modalità della diagnosi, le scelte terapeutiche.

presenta una presenza in percentuale rilevante anche nel gruppo più giovane. La fascia d'età centrale, coinvolta nello screening regionale, presenta la percentuale più alta di stadi a prognosi favorevole.

Per quanto riguarda il grado di malignità, non vi è una relazione statisticamente significativa con l'età delle pazienti. Si nota una maggiore presenza di malattia con gradi a malignità bassa o intermedia (1 e 2, ossia neoplasie cellule ben differenziate), mentre sono meno numerosi i casi a malignità medio-alta o alta (3 e 4, ossia situazioni a scarsa differenziazione cellulare), che fanno prevedere una prognosi peggiore.

Tab. 6.5 – Caratteristiche cliniche delle pazienti, per età.

| | | Totale | | 0-49 | | 50-69 | | 70+ | | Sig. χ^2 |
|--------------------------------|-----------------------|--------|------|-------|-------|-------|------|-------|------|---------------|
| | | N | % | N | % | N | % | N | % | |
| Stadio alla diagnosi | I | 2.576 | 47,9 | 513 | 44,6 | 1.406 | 53,9 | 657 | 40,5 | p<0,05 |
| | II | 1.907 | 35,5 | 434 | 37,7 | 838 | 32,1 | 635 | 39,1 | |
| | III | 608 | 11,3 | 148 | 12,9 | 249 | 9,6 | 211 | 13,0 | |
| | IV | 288 | 5,4 | 55 | 4,8 | 114 | 4,4 | 119 | 7,3 | |
| Grado alla diagnosi | Grado 1 | 704 | 13,0 | 151 | 13,1 | 379 | 14,5 | 174 | 10,7 | n.s |
| | Grado 2 | 2.854 | 53,1 | 568 | 49,4 | 1.393 | 53,4 | 893 | 55,1 | |
| | Grado 3 | 1.505 | 28,0 | 374 | 32,5 | 706 | 27,1 | 425 | 26,2 | |
| | Grado 4 | 316 | 5,9 | 57 | 5,0 | 129 | 4,9 | 130 | 8,0 | |
| Modalità diagnosi mammografia | No | 3.114 | 57,9 | 627 | 54,5 | 1.537 | 59,0 | 950 | 58,6 | p<0,05 |
| | Diagnosi: mammografia | 2.265 | 42,1 | 523 | 45,5 | 1.070 | 41,0 | 672 | 41,4 | |
| Modalità diagnosi ecografia | No | 2.798 | 52,0 | 522 | 45,4 | 1.388 | 53,2 | 888 | 54,7 | p<0,05 |
| | Diagnosi: ecografia | 2.581 | 48,0 | 628 | 54,6 | 1.219 | 46,8 | 734 | 45,3 | |
| Modalità diagnosi citologia | No | 3.565 | 66,3 | 706 | 61,4 | 1.729 | 66,3 | 1.130 | 69,7 | p<0,05 |
| | Diagnosi: citologia | 1.814 | 33,7 | 444 | 38,6 | 878 | 33,7 | 492 | 30,3 | |
| Modalità diagnosi biopsia | No | 2.750 | 51,1 | 568 | 49,4 | 1.266 | 48,6 | 916 | 56,5 | p<0,05 |
| | Diagnosi: biopsia | 2.629 | 48,9 | 582 | 50,6 | 1.341 | 51,4 | 706 | 43,5 | |
| Modalità diagnosi solo clinica | No | 5.319 | 98,9 | 1.133 | 98,5 | 2.584 | 99,1 | 1.602 | 98,8 | n.s. |
| | Diagnosi: clinica | 60 | 1,1 | 17 | 1,5 | 23 | 0,9 | 20 | 1,2 | |
| Modalità diagnosi RMN | No | 5.217 | 97,0 | 1.093 | 95,0 | 2.537 | 97,3 | 1.587 | 97,8 | p<0,05 |
| | Diagnosi: RNM | 162 | 3,0 | 57 | 5,0 | 70 | 2,7 | 35 | 2,2 | |
| Modalità diagnosi anamnestica | No | 5.353 | 99,5 | 1.145 | 99,6 | 2.594 | 99,5 | 1.614 | 99,5 | n.s. |
| | Diagnosi: anamnestica | 26 | 0,5 | 5 | 0,4 | 13 | 0,5 | 8 | 0,5 | |
| Modalità diagnosi screening | No | 5.061 | 94,1 | 1.150 | 100,0 | 2.296 | 88,1 | 1.615 | 99,6 | p<0,05 |
| | Diagnosi: screening | 318 | 5,9 | 0 | 0,0 | 311 | 11,9 | 7 | 0,4 | |

Nota: n.s. = assenza di significatività statistica.

(Continua)

(Continua)

| | | Totale | | 0-49 | | 50-69 | | 70+ | | Sig. χ^2 |
|------------------------------|----------------------------------|--------|------|-------|------|-------|------|-------|------|---------------|
| | | N | % | N | % | N | % | N | % | |
| Recettori estrogeni segno | Recettori estrogeni: negativo | 1.146 | 21,3 | 302 | 26,3 | 568 | 21,8 | 276 | 17,0 | n.s. |
| | Recettori estrogeni: positivo | 4233 | 78,7 | 848 | 73,7 | 2.039 | 78,2 | 1.346 | 83,0 | |
| Recettori progesterone segno | Recettori progesterone: negativo | 1950 | 36,3 | 376 | 32,7 | 1.011 | 38,8 | 563 | 34,7 | n.s. |
| | Recettori progesterone: positivo | 3429 | 63,8 | 774 | 67,3 | 1.596 | 61,2 | 1.059 | 65,3 | |
| Comorbidità alla diagnosi | No comorbidità | 3451 | 64,2 | 988 | 85,9 | 1.673 | 64,2 | 790 | 48,7 | p<0,05 |
| | Comorbidità | 1928 | 35,8 | 162 | 14,1 | 934 | 35,8 | 832 | 51,3 | |
| Chemioterapia | No chemioterapia | 2746 | 51,1 | 298 | 25,9 | 1.162 | 44,6 | 1.286 | 79,3 | p<0,05 |
| | Chemioterapia | 2633 | 49,0 | 852 | 74,1 | 1.445 | 55,4 | 336 | 20,7 | |
| Ormonoterapia | No ormonoterapia | 1817 | 33,8 | 338 | 29,4 | 801 | 30,7 | 678 | 41,8 | p<0,05 |
| | Ormonoterapia | 3562 | 66,2 | 812 | 70,6 | 1.806 | 69,3 | 944 | 58,2 | |
| Radioterapia | No radioterapia | 1975 | 36,7 | 364 | 31,7 | 722 | 27,7 | 889 | 54,8 | p<0,05 |
| | Radioterapia | 3404 | 63,3 | 786 | 68,3 | 1.885 | 72,3 | 733 | 45,2 | |
| Chirurgia | No chirurgia | 169 | 3,1 | 26 | 2,3 | 66 | 2,5 | 77 | 4,7 | p<0,05 |
| | Chirurgia | 5210 | 96,9 | 1.124 | 97,7 | 2.541 | 97,5 | 1.545 | 95,3 | |

Nota: n.s. = assenza di significatività statistica.

Le modalità di diagnosi per immagini e biopsia, che aumentano le probabilità di una identificazione precoce, sono maggiormente utilizzate nelle due fasce di età più giovani (si notano percentuali decrescenti al crescere dell'età, statisticamente significative). Le pazienti hanno potuto accedere a più di una modalità di diagnosi, combinando soprattutto mammografia o ecografia con biopsia.

In tabella è anche mostrata la distribuzione dei recettori estrogenici e progesteronici negativi e positivi. Questi ultimi, migliorando la capacità di ricezione delle terapie ormonali mirate, possono incrementare le probabilità di sopravvivenza (si veda, ad esempio, Bal et al., 2015). Non sono legate alle età e, nelle pazienti in esame, sono in maggioranza positivi entrambi.

E' invece fortemente legata all'età la presenza di comorbidità, ossia di altre patologie, alla diagnosi; circa il 36% dei casi in esame mostra tale situazione e, di essi, più del 51% sono ultrasessantenni.

Per quanto riguarda le terapie, la chirurgia è, come da protocolli di cura, utilizzata in più del 95% dei casi, ma si nota una debole riduzione della sua applicazione al crescere dell'età, statisticamente significativa. La seconda terapia più utilizzata è l'ormonoterapia, seguita da radioterapia e chemioterapia. Anche esse vedono una

riduzione dell'applicazione nel gruppo delle ultrasessantenni, in particolare per quanto riguarda la chemioterapia.

Un numero residuale di donne non è stato trattato in alcun modo per la patologia, circa l'1% (62 individui). Sono per lo più anziane e in stadio avanzato e sperimentano una prognosi non favorevole, ma per questo tumore costituiscono un gruppo residuale non significativo (dati riportati in appendice, variabile utilizzata nei modelli di sopravvivenza).

La tabella 6.6, infine, presenta i valori medi percentuali delle caratteristiche (utilizzate come covariate nel modello di sopravvivenza) delle sezioni di Censimento in cui sono residenti le pazienti comprese nell'analisi, rispetto alla loro distribuzione regionale.

Tab. 6.6 – Caratteristiche delle sezioni di Censimento delle pazienti a confronto con quelle regionali.

| | Casi in analisi | | | Regione |
|--|--------------------------|--------------------------------|--------------------------|-------------------------|
| | | Intervallo di confidenza (95%) | | |
| | Valore medio aree | Estremo inferiore | Estremo superiore | Valore regionale |
| Cittadinanza italiana (%) | 96,4 | 96,2 | 96,6 | 96,6 |
| Cittadinanza straniera (%) | 3,2 | 3,1 | 3,3 | 3,4 |
| Numero persone per famiglia | 2,6 | 2,6 | 2,6 | 2,6 |
| Dipendenti o in altra posizione subordinata (%) | 70,1 | 69,8 | 70,5 | 68,8 |
| Imprenditori (%) | 3,5 | 3,4 | 3,6 | 3,3 |
| Liberi professionisti (%) | 5,2 | 5,1 | 5,4 | 3,4 |
| Lavoratori in proprio (%) | 16,5 | 16,3 | 16,7 | 20,0 |
| Soci di cooperative (%) | 2,0 | 2,0 | 2,1 | 1,9 |
| Coadiuvanti familiari (%) | 1,6 | 1,5 | 1,7 | 2,0 |
| % Occupati | 43,6 | 43,3 | 43,8 | 41,5 |
| % In cerca di prima occupazione | 1,6 | 1,5 | 1,6 | 1,5 |
| % Disoccupati (in cerca di nuova occupazione) | 2,8 | 2,8 | 2,9 | 2,8 |
| % In attesa di iniziare un lavoro già trovato | 0,3 | 0,3 | 0,3 | 0,3 |
| % Studenti | 7,6 | 7,5 | 7,7 | 6,7 |
| % Casalinghe | 14,1 | 13,9 | 14,2 | 15,4 |
| % Ritirati dal lavoro | 24,8 | 24,6 | 25,0 | 26,4 |
| % In servizio di leva o in servizio civile sostitutivo | 0,2 | 0,2 | 0,2 | 0,2 |
| % Inabile al lavoro | 1,4 | 1,4 | 1,5 | 1,7 |
| % In altra condizione | 3,2 | 3,1 | 3,3 | 3,6 |

(Continua)

(Continua)

| | Casi in analisi | | | Regione |
|---|--------------------------------|-------------------|-------------------|------------------|
| | Intervallo di confidenza (95%) | | | |
| | Valore medio aree | Estremo inferiore | Estremo superiore | Valore regionale |
| Celibe / Nubile (%) | 35,7 | 35,6 | 35,9 | 35,1 |
| Coniugato/a (%) | 51,2 | 51,0 | 51,5 | 52,7 |
| Separato/a di fatto (%) | 0,4 | 0,3 | 0,4 | 0,3 |
| Separato/a legalmente (%) | 1,4 | 1,4 | 1,5 | 1,2 |
| Divorziato/a (%) | 1,3 | 1,2 | 1,3 | 0,8 |
| Vedovo/a (%) | 9,6 | 9,4 | 9,7 | 9,9 |
| Non sa leggere o scrivere (%) | 1,1 | 1,1 | 1,2 | 1,6 |
| Sa leggere e scrivere (%) | 7,0 | 6,9 | 7,2 | 9,8 |
| Licenza di scuola elementare (%) | 24,1 | 23,9 | 24,3 | 27,7 |
| Licenza di scuola media inferiore (%) | 25,9 | 25,7 | 26,1 | 27,7 |
| Diploma di scuola secondaria superiore (%) | 31,5 | 31,3 | 31,8 | 27,6 |
| Diploma non universitario post maturità (%) | 0,5 | 0,5 | 0,5 | 0,4 |
| Diploma universitario (%) | 0,9 | 0,9 | 1,0 | 0,6 |
| Laurea (%) | 8,6 | 8,4 | 8,8 | 4,7 |
| Indice di ricambio | 160,9 | 157,4 | 164,4 | 144,5 |
| Indice di vecchiaia | 227,5 | 220,0 | 232,9 | 220,8 |
| Indice di dipendenza strutturale | 56,4 | 55,6 | 57,3 | 59,2 |
| Tasso di attività | 65,4 | 65,2 | 65,7 | 64,2 |
| Tasso di occupazione | 59,0 | 58,7 | 59,3 | 57,8 |
| Tasso di disoccupazione | 9,2 | 9,1 | 9,4 | 9,4 |
| % alloggi di proprietà | 67,1 | 66,6 | 67,5 | 59,1 |
| % alloggi in affitto | 13,5 | 13,2 | 13,8 | 8,4 |
| % alloggi con altro titolo di godimento | 8,5 | 8,0 | 8,7 | 8,2 |
| % alloggi senza alcuna specifica | 10,0 | 9,7 | 10,3 | 24,2 |

Le differenze significative mostrano che, tendenzialmente, le persone considerate nell'analisi sono residenti in un insieme di sezioni di Censimento più ricco del resto della regione, in quanto si registrano migliori percentuali nell'area dell'occupazione, maggiore presenza di titoli di studio elevati, occupazioni a reddito potenzialmente più alto, migliori indicatori di sintesi (ad esempio, indice di ricambio più elevato e indice di dipendenza strutturale più basso).

6.1.2 I risultati del modello di sopravvivenza a parametri flessibili ed effetti misti

La tabella 6.7 mostra il risultato del confronto fra i diversi modelli di sopravvivenza implementati.

Tab. 6.7 – Modelli di sopravvivenza a parametri flessibili ed effetti misti implementati: caratteristiche e confronto.

| Età | Stadio | Variabili diagnostiche | Comorbidità | Variabili terapeutiche | Effetti fissi Caratteristiche socio-economiche individuali | Caratteristiche socio-economiche sezione di Censimento | Effetti casuali Sezione di Censimento di residenza | N | AIC |
|-----|--------|------------------------|-------------|------------------------|---|--|---|-------|---------|
| X | X | | | | | | X | 5.379 | 7.060,9 |
| X | X | | | | X | | X | 5.379 | 7.061,3 |
| X | X | | | | | X | X | 5.379 | 7.076,7 |
| X | X | X | | | | | X | 5.379 | 6.946,2 |
| X | X | X | X | | | | X | 5.379 | 6.946,1 |
| X | X | X | | | X | | X | 5.379 | 6.954,4 |
| X | X | X | | | | X | X | 5.379 | 6.967,6 |
| X | X | X | X | | X | | X | 5.379 | 6.954,1 |
| X | X | X | X | | | X | X | 5.379 | 6.967,5 |
| X | X | | | | X | X | X | 5.379 | 7.082,2 |
| X | X | X | | | X | X | X | 5.379 | 6.973,8 |
| X | X | X | X | | X | X | X | 5.379 | 6.973,4 |
| X | X | | | X | | | X | 5.379 | 6.948,3 |
| X | X | | | X | X | | X | 5.379 | 6.956,0 |
| X | X | | | X | | X | X | 5.379 | 6.973,2 |
| X | X | X | | X | | | X | 5.379 | 6.899,0 |
| X | X | X | X | X | | | X | 5.379 | 6.897,7 |
| X | X | X | | X | X | | X | 5.379 | 6.906,7 |
| X | X | X | | X | | X | X | 5.379 | 6.925,8 |
| X | X | X | X | X | X | | X | 5.379 | 6.905,0 |
| X | X | X | X | X | | X | X | 5.379 | 6.925,0 |
| X | X | | | X | X | X | X | 5.379 | 6.979,1 |
| X | X | X | | X | X | X | X | 5.379 | 6.931,4 |
| X | X | X | X | X | X | X | X | 5.379 | 6.890,1 |

In essa sono evidenziati i gruppi di variabili utilizzati per ogni modello (a partire dal modello base che comprende età e stadio per quanto riguarda gli effetti fissi e la sezione di Censimento di residenza per quanto riguarda gli effetti casuali), il numero di casi utilizzati in ogni modello e il valore del Criterio Informativo di Akaike (AIC).

Il modello che presenta l'AIC più basso è quello in cui tutte le aree sono rappresentate. Se si osserva l'andamento dei valori dell'AIC, si può già fare qualche considerazione iniziale su quali siano le aree che, a partire dal modello di base, incidono in maniera più rilevante sul calcolo della sopravvivenza.

Si può, infatti, facilmente notare che le riduzioni più marcate si ottengono quando nel modello sono inserite le variabili relative alla diagnosi, ai trattamenti terapeutici e alle caratteristiche socio-economiche individuali, mentre effetti di riduzione minori si registrano quando si inseriscono le comorbidità e le caratteristiche dell'area di residenza.

Questi effetti sembrano confermati dai risultati ottenuti con il modello di sopravvivenza illustrato nella tabella 6.8, che corrisponde a quello identificato dall'AIC come il più adeguato.

Per comodità di lettura, nella tabella sono state riportate solo le covariate che hanno una relazione statisticamente significativa con il rischio di morte, escludendo tutte le variabili non statisticamente significative e quelle escluse dal test di collinearità. Similarmente, si sono riportate solo le interazioni a due vie statisticamente significative e non escluse dal medesimo test.

Tab. 6.8 – Il modello di sopravvivenza relativa a parametri flessibili ed effetti misti.

| | | | <i>Intervallo di confidenza (95%)</i> | | | |
|------------------------------|---|----------------------------------|---|-----------------------------------|-----------------|--------|
| | | <i>HR</i> | <i>Sig.</i> | <i>Est. Inf.</i> | <i>Est. Sup</i> | |
| Effetti fissi | Stadio | Stadio I | 1 | <i>(categoria di riferimento)</i> | | |
| | | Stadio II | 6,06 | 0,00 | 2,97 | 12.36 |
| | | Stadio III | 20,61 | 0,00 | 10,09 | 42.12 |
| | | Stadio IV | 83,91 | 0,00 | 39,15 | 179.83 |
| | Diagnosi | Grado 1 | 1 | <i>(categoria di riferimento)</i> | | |
| | | Grado 2 | 1,63 | 0,05 | 1,00 | 2.71 |
| | | Grado 3 | 2,89 | 0,00 | 1,73 | 4.82 |
| | | Grado 4 | 1,81 | 0,03 | 1,05 | 3.13 |
| | | Recettore estrogeni negativo | 1 | <i>(categoria di riferimento)</i> | | |
| | | Recettore estrogeni positivo | 0,68 | 0,01 1 | 0,50 | 0.92 |
| | | Assenza di comorbidità | 1 | <i>(categoria di riferimento)</i> | | |
| | | Presenza di comorbidità | 1,26 | 0,05 | 1,00 | 1.58 |
| | Terapia | No ormonoterapia | 1 | <i>(categoria di riferimento)</i> | | |
| | | Si ormonoterapia | 0,65 | 0,00 | 0,50 | 0.84 |
| | | No radioterapia | 1 | <i>(categoria di riferimento)</i> | | |
| | | Si radioterapia | 0,74 | 0,01 | 0,59 | 0.93 |
| | | No terapia chirurgica | 1 | <i>(categoria di riferimento)</i> | | |
| | | Si terapia chirurgica | 0,39 | 0,00 | 0,26 | 0.57 |
| | SE individuali | Coniugata/convivente | 1 | <i>(categoria di riferimento)</i> | | |
| | | Nubile | 1,27 | 0,05 | 1,01 | 1.76 |
| | | Separata/divorziata | 1,41 | 0,26 | 0,77 | 2.56 |
| | | Vedova | 1,36 | 0,03 | 1,04 | 1.91 |
| | SE d'area | Indice di dipendenza strutturale | 1,01 | 0,05 | 1,00 | 1.01 |
| | Effetti fissi di interazione a due vie | 70+ # Stadio IV | 1.26 | 0,01 | 1,03 | 1,47 |
| | | 70+ # Grado 3 | 1.26 | 0,00 | 1,12 | 1,39 |
| | | Stadio IV # Ormonoterapia | 0.63 | 0,04 | 0,31 | 0,95 |
| | | Grado 3 # Chemioterapia | 0.51 | 0,05 | 0,27 | 0,75 |
| Grado 4 # Chirurgia | | 0.43 | 0,03 | 0,25 | 0,63 | |
| Grado 2 # Licenza elementare | | 2.43 | 0,01 | 1,98 | 2,94 | |
| Grado 3 # Licenza elementare | | 2.70 | 0,00 | 2,20 | 3,21 | |
| Effetti casuali | Sezione di Censimento | 0.57 | | 0,35 | 1,01 | |

Nota: HR = hazard ratio (eccesso di rischio di morte); SE = variabili socio-economiche.

Il modello presenta un valore di verosimiglianza (*log-likelihood*), ossia di bontà del modello, pari a -3.396,8, statisticamente significativo, raggiunto utilizzando una

quadratura di Gauss-Hermite a 12 nodi. Il modello e i suoi risultati sono pertanto da considerarsi attendibili.

Osservando gli effetti principali, si nota subito che lo stadio è il principale predittore del rischio di morte, con eccessi di rischio che si incrementano al crescere dello stadio ossia al peggiorare della malattia alla diagnosi, confermando quindi quanto noto in letteratura e commentato in precedenza.

Per quanto riguarda le variabili legate alla diagnosi, il grado di malignità che indica una peggiore sopravvivenza sembra essere il grado 3, con un eccesso di rischio quasi tre volte più alto rispetto al grado 1. Alti anche i rischi per i gradi 2 (63% di eccesso di rischio) e 4 (81%).

Anche la presenza di comorbidità alla diagnosi gioca un ruolo peggiorativo sulle probabilità di sopravvivenza (26%), mentre sembra essere confermato l'effetto protettivo dei recettori estrogenici positivi (riduzione del rischio di morte del 32%).

Fra le covariate che descrivono le terapie spicca l'assenza della chemioterapia, che è stata esclusa dal modello perché non statisticamente significativa, mentre chirurgia, ormonoterapia e radioterapia sembrano confermare la loro efficacia.

L'ipotesi dello studio è, però, che le caratteristiche socio-economiche individuali e relative all'area di residenza abbiano una relazione con la sopravvivenza, malgrado l'inserimento nel modello delle variabili diagnostiche e terapeutiche di cui, invece, è noto l'effetto diretto sulla sopravvivenza.

Per quanto riguarda il tumore alla mammella nelle donne, questa ipotesi sembra verificata. A livello individuale, infatti, lo stato civile delle pazienti presenta un effetto piuttosto rilevante: le nubili, infatti, sperimentano un eccesso di rischio di morte del 27% rispetto alle coniugate, mentre le vedove un eccesso pari al 36%, due valori quindi niente affatto trascurabili che verranno esplorati come effetti sulla diagnosi e la cura (quindi effetti indiretti rispetto alla sopravvivenza) nel paragrafo successivo.

Anche la seconda ipotesi di una differenza fra gli effetti socio-economici individuali e relativi alla sezione di Censimento di residenza sembra verificata. Emerge, infatti, come statisticamente significativo il ruolo dell'indice di dipendenza strutturale, ossia dell'indice ISTAT che misura il rapporto tra popolazione non attiva e attiva: al crescere dell'indice aumenta dell'1% circa l'eccesso del rischio di morte. L'effetto è, evidentemente, debole ma non può essere ignorato.

Fra le variabili di base, emerge l'assenza di significatività statistica per l'età alla diagnosi, che però emerge negli effetti fissi legati all'interazione a due vie.

Infatti, l'effetto combinato di età anziana (70 e più anni) con stadio 4 e grado 3, rispettivamente, portano in entrambi i casi ad un eccesso di rischio di morte del 26%.

L'effetto combinato di ormonoterapia e stadio IV, chirurgia e grado 4 e chemioterapia e grado 3, invece, migliorano le probabilità di sopravvivenza, riducendo gli eccessi di rischio di morte.

Nelle interazioni entra in gioco anche una variabile socio-economica individuale: il titolo di studio che, a livello di licenza elementare (o analfabeta) si combina con i gradi 2 e 3, incrementando significativamente gli eccessi di rischio di morte.

La sezione di censimento di residenza, ossia la variabile di *clustering* responsabile per gli effetti casuali, non ha un ruolo statisticamente significativo nel modello.

6.1.3 L'analisi degli effetti dello stato civile sulla diagnosi e sulle cure

Come ricordato in precedenza, l'effetto delle variabili socio-economiche sulla sopravvivenza si esprime soprattutto in maniera indiretta, in genere condizionando la tempestività della diagnosi e l'adesione alle cure. Nello specifico del tumore alla mammella, solo lo stato civile presenta un ruolo rilevante, mentre il titolo di studio si limita ad un'interazione con il grado di malignità, caratteristica innata del tumore, non legata alle caratteristiche socio-economiche dell'individuo³⁸.

In questo paragrafo esamineremo, quindi il ruolo dello stato civile nell'influenzare la tempestività della diagnosi (ossia lo stadio) e le cure a cui le pazienti hanno avuto accesso.

La tabella 6.9 mostra come si distribuiscono lo stadio e le terapie rispetto allo stato civile, valutando la connessione fra le variabili tramite significatività statistica del test del χ^2 ($p < 0,05$).

³⁸ In Appendice sono riportati i risultati delle regressioni logistiche fra stadio alla diagnosi, terapie applicate e titolo di studio, da cui si evince l'assenza di relazioni statisticamente significative, ad eccezione delle pazienti con licenza elementare, che hanno meno probabilità di essere trattate con chemioterapia, ormonoterapia o radioterapia (Tab. A.3 e Tab. A.4).

Tab. 6.9 – Stato civile, stadio alla diagnosi e terapie seguite.

| | | Stato civile | | | | | | | | Sig. χ^2 |
|----------------------|------------------|---------------|------|--------------------------|------|---------------------------|------|----------|------|---------------|
| | | Celibe/nubile | | Coniugato/a o convivente | | Separato/a o divorziato/a | | Vedovo/a | | |
| | | N | % | N | % | N | % | N | % | |
| Stadio alla diagnosi | I | 228 | 44,4 | 1.925 | 49,7 | 96 | 48,7 | 327 | 41,1 | p<0,05 |
| | II | 172 | 33,5 | 1.357 | 35,0 | 62 | 31,5 | 316 | 39,7 | |
| | III | 78 | 15,2 | 407 | 10,5 | 31 | 15,7 | 92 | 11,6 | |
| | IV | 36 | 7,0 | 184 | 4,8 | 8 | 4,1 | 60 | 7,5 | |
| Ormonoterapia | No ormonoterapia | 171 | 33,3 | 1.263 | 32,6 | 71 | 36,0 | 312 | 39,2 | p<0,05 |
| | Ormonoterapia | 343 | 66,7 | 2.610 | 67,4 | 126 | 64,0 | 483 | 60,8 | |
| Radioterapia | No radioterapia | 196 | 38,1 | 1.310 | 33,8 | 75 | 38,1 | 394 | 49,6 | p<0,05 |
| | Radioterapia | 318 | 61,9 | 2.563 | 66,2 | 122 | 61,9 | 401 | 50,4 | |
| Chirurgia | No chirurgia | 23 | 4,5 | 108 | 2,8 | 3 | 1,5 | 35 | 4,4 | p<0,05 |
| | Chirurgia | 491 | 95,5 | 3.765 | 97,2 | 194 | 98,5 | 760 | 95,6 | |
| Chemioterapia | No chemioterapia | 222 | 43,2 | 1.877 | 48,5 | 83 | 42,1 | 564 | 70,9 | p<0,05 |
| | Chemioterapia | 292 | 56,8 | 1.996 | 51,5 | 114 | 57,9 | 231 | 29,1 | |

Si può notare come lo stadio IV sia più elevato nelle nubili e nelle vedove, evidenziando quindi come lo stato civile non sia solo *proxy* dell'età delle pazienti, ma abbia anche a che fare con altri aspetti relativi al modo in cui viene affrontata la malattia, di cui si parlerà nel prossimo capitolo. Questa affermazione sembra confermata da come si distribuiscono gli altri stadi; lo stadio I è meno presente fra le nubili, così come lo stadio II è più presente nelle vedove e lo stadio III fra nubili e separate/divorziate.

Rispetto alle terapie, si osserva come, tendenzialmente, nubili e vedove vedano una minore adesione ai protocolli di cura, in particolare per quanto riguarda la radioterapia e, per le vedove, la chemioterapia.

Una lettura più dettagliata del ruolo dello stato civile su tempestività alla diagnosi e adesione ai protocolli di cura è data dai risultati dei modelli di regressione logistica presentati nelle tabelle 6.10 e 6.11, in cui è utilizzato anche l'indice di dipendenza strutturale, unica variabile socio-economica d'area che ha mostrato un effetto statisticamente significativo sulla sopravvivenza.

Tab. 6.10 – Stadio alla diagnosi e variabili socio-economiche: modelli di regressione logistica.

| Stadio alla diagnosi | | OR | Sig. |
|----------------------|----------------------------------|------|----------------------------|
| Stadio I | Coniugato/convivente | 1 | (categoria di riferimento) |
| | Celibe/nubile | 0,81 | p<0,05 |
| | Separato/divorziato | 0,96 | n.s. |
| | Vedovo | 0,71 | p<0,05 |
| | Indice di dipendenza strutturale | 1,00 | n.s. |
| | Costante | 0,99 | n.s. |
| Stadio II | Coniugato/convivente | 1 | (categoria di riferimento) |
| | Celibe/nubile | 0,93 | n.s. |
| | Separato/divorziato | 0,85 | n.s. |
| | Vedovo | 1,22 | p<0,05 |
| | Indice di dipendenza strutturale | 1,00 | n.s. |
| | Costante | 0,52 | p<0,05 |
| Stadio III | Coniugato/convivente | 1 | (categoria di riferimento) |
| | Celibe/nubile | 1,53 | p<0,05 |
| | Separato/divorziato | 1,59 | p<0,05 |
| | Vedovo | 1,12 | n.s. |
| | Indice di dipendenza strutturale | 1,00 | n.s. |
| | Costante | 0,13 | p<0,05 |
| Stadio IV | Coniugato/convivente | 1 | (categoria di riferimento) |
| | Celibe/nubile | 1,51 | p<0,05 |
| | Separato/divorziato | 0,85 | n.s. |
| | Vedovo | 1,64 | p<0,05 |
| | Indice di dipendenza strutturale | 1,00 | n.s. |
| | Costante | 0,05 | p<0,05 |

Nota: OR = Odds ratio.

Nubili e vedove hanno meno probabilità di avere uno stadio I alla diagnosi, mentre la probabilità di avere stadi più alti è particolarmente alta in questi due gruppi. In particolare, le nubili presentano alte probabilità (superiori al 50%) di presentarsi alla diagnosi con stadi III e IV, mentre le vedove hanno il 64% di probabilità di avere uno stadio IV. Non vi sono, invece, effetti sullo stadio legati all'indice di dipendenza strutturale.

Tab. 6.11 – Terapie e variabili socio-economiche: modelli di regressione logistica.

| Terapie | | OR | Sig. |
|---------------|----------------------------------|-------|----------------------------|
| Chemioterapia | Coniugato/convivente | 1 | (categoria di riferimento) |
| | Celibe/nubile | 1,24 | p<0,05 |
| | Separato/divorziato | 1,29 | n.s. |
| | Vedovo | 0,39 | p<0,05 |
| | Indice di dipendenza strutturale | 1,00 | n.s. |
| | Costante | 1,16 | p<0,05 |
| Ormonoterapia | Coniugato/convivente | 1 | (categoria di riferimento) |
| | Celibe/nubile | 0,97 | n.s. |
| | Separato/divorziato | 0,86 | n.s. |
| | Vedovo | 0,75 | p<0,05 |
| | Indice di dipendenza strutturale | 1,00 | n.s. |
| | Costante | 2,19 | p<0,05 |
| Radioterapia | Coniugato/convivente | 1 | (categoria di riferimento) |
| | Celibe/nubile | 0,83 | p<0,05 |
| | Separato/divorziato | 0,83 | n.s. |
| | Vedovo | 0,52 | p<0,05 |
| | Indice di dipendenza strutturale | 1,00 | n.s. |
| | Costante | 2,13 | p<0,05 |
| Chirurgia | Coniugato/convivente | 1 | (categoria di riferimento) |
| | Celibe/nubile | 0,61 | p<0,05 |
| | Separato/divorziato | 1,85 | n.s. |
| | Vedovo | 0,62 | p<0,05 |
| | Indice di dipendenza strutturale | 1,00 | n.s. |
| | Costante | 34,25 | p<0,05 |

Nota: OR = Odds ratio.

Rispetto alle terapie si osserva che le vedove hanno probabilità sensibilmente ridotte di ricevere le diverse terapie previste dai protocolli di cura, mentre le nubili hanno probabilità ridotte di essere trattate chirurgicamente o con radioterapia, mentre è più probabile che vengano avviate al trattamento chemioterapico (così come le separate/divorziate). Anche per le terapie non si evidenziano effetti legati all'indice di dipendenza strutturale. Vi è invece un ruolo rilevante della costante in tutte le modalità terapeutiche, in particolare nell'applicazione della chirurgia, che probabilmente esprime altri aspetti di tipo clinico.

6.2 Il tumore al colon

Come spiegato nel capitolo 5, le neoplasie del colon-retto sono state considerate separatamente per le sedi relative al colon e al retto (in merito, oltre a quanto già citato in precedenza, vd. anche Tamas et al., 2015).

Il processo di esclusione, basato ovviamente sugli stessi criteri che hanno riguardato le malate di tumore alla mammella, ha coinvolto i 12.087 pazienti affetti da tali neoplasie. I risultati dell'esclusione sono riportati nella tabella 6.12.

Tab. 6.12 – Casi esclusi e motivazioni.

| Ragioni dell'esclusione | Casi esclusi | % |
|---|---------------------|-------------|
| Casi con tumore a comportamento benigno | 3.451 | 28,6 |
| Casi con stadio non definito o non disponibile | 629 | 5,2 |
| Casi per cui è mancante almeno una informazione fra stato civile, titolo di studio, professione | 217 | 1,8 |
| <i>Totale casi esclusi</i> | <i>4.297</i> | <i>35,6</i> |
| Casi mantenuti nell'analisi | 7.790 | 64,4 |
| <i>Casi con tumore al colon</i> | <i>5.276</i> | <i>67,7</i> |
| <i>Casi con tumore al retto</i> | <i>2.514</i> | <i>32,3</i> |
| Totale casi | 12.087 | 100,0 |

Il 29% circa dei casi è stato escluso, perché composto da individui a cui sono stati diagnosticati tumori benigni; solo il 5% circa è stato escluso perché caratterizzato da stadio non definito, mentre il 2% circa dei casi è stato escluso in quanto non è stato possibile recuperare dalle Anagrafi comunali tutte le informazioni relative a stato civile, titolo di studio, condizione professionale.

Sono quindi rimasti a disposizione per l'analisi 7.790 casi, di cui 5.276 persone affette da tumore del colon e 2.514 affette da tumore del retto.

6.2.1 Le caratteristiche di base

La tabella 6.13 mostra le caratteristiche socio-economiche individuali delle persone affette dal tumore al colon, considerate nello studio.

Tab. 6.13 – Le caratteristiche demografiche e socio-economiche individuali dei pazienti.

| Caratteristiche individuali | | N | % |
|------------------------------------|--|--------------|----------|
| Genere | F | 2.363 | 44,8 |
| | M | 2.913 | 55,2 |
| Età all'incidenza | 0-49 | 161 | 3,1 |
| | 50-69 | 1.547 | 29,3 |
| | 70+ | 3.568 | 67,6 |
| Stato civile | Celibe/nubile | 580 | 11,0 |
| | Coniugato/a o convivente | 3.566 | 67,6 |
| | Separato/a o divorziato/a | 155 | 2,9 |
| | Vedovo/a | 975 | 18,5 |
| Titolo di studio | Licenza elementare o analfabeta | 3.345 | 63,4 |
| | Licenza media inferiore | 1.068 | 20,2 |
| | Licenza media superiore | 699 | 13,2 |
| | Laurea | 164 | 3,1 |
| Condizione professionale | Professione intellettuale, scientifica o di elevata specializzazione | 2 | 0,0 |
| | Professione qualificata nelle attività commerciali o nei servizi | 139 | 2,6 |
| | Impiegato | 473 | 9,0 |
| | Artigiano, operaio specializzato o agricoltore | 949 | 18,0 |
| | Pensionato | 3.343 | 63,4 |
| | Disoccupato | 3 | 0,1 |
| | Casalinga | 367 | 7,0 |
| Totale | | 5.276 | |

Fra i casi analizzati, vi sono un po' più uomini che donne, con una proporzione che rispecchia perfettamente quella registrata su tutta la popolazione italiana affetta da tale tumore (dati ITACAN, AIRTUM, 2014).

Rispetto alle età, i pazienti sono in maggioranza appartenenti alla fascia anziana (il 68% ha 70 e più anni), mentre meno del 30% ha avuto diagnosi di tumore nella fascia di età interessata dallo screening organizzato. Molto bassa la percentuale di giovani affetti da tale tumore (meno del 3%).

Rispetto ai dati italiani per lo stesso periodo (dati ITACAN, AIRTUM, 2014), si osservano proporzioni di neoplasia al colon molto simili.

Gli individui considerati sono per lo più coniugati (o conviventi), poi vedovi, celibi/nubili e, infine, separati/divorziati in bassa percentuale. Rispetto al dato

dell'intera popolazione umbra (dati ISTAT, 2001), si nota una distribuzione simile per quanto riguarda i vedovi (14% circa in tutta la regione) e i separati/divorziati (2% in tutta la regione), ma una percentuale decisamente più alta di coniugati/conviventi (49% in tutta la regione) e più bassa di celibi/nubili (35% in tutta la regione).

Rispetto al titolo di studio, si osserva una presenza elevata dei titoli elementari, seguiti dalle licenze medie inferiori, dai diplomi di scuola superiore e, infine dalle, lauree. Le percentuali dei titoli inferiori sono più alte dei valori regionali e nazionali, mentre diplomi di scuola secondaria superiore e lauree sono molto meno presenti rispetto a regione e nazione (Vercelli et al., 2013).

La condizione professionale dominante è l'essere pensionati, coerentemente con la distribuzione delle età; rilevante pure la presenza di artigiani, operai specializzati e agricoltori, meno presenti le altre occupazioni.

Le caratteristiche sembrano quindi essere connesse alle età dei malati; la tabella 6.14 analizza la distribuzione di genere, stato civile e titolo di studio per i tre gruppi di età.

Tab. 6.14 – Genere, stato civile e titolo di studio per età alla diagnosi.

| | | 0-49 | | 50-69 | | 70+ | | Sig. χ^2 |
|------------------|---------------------------------|------|------|-------|------|-------|------|---------------|
| | | N | % | N | % | N | % | |
| Genere | F | 85 | 52,8 | 632 | 40,9 | 1.646 | 46,1 | p<0,05 |
| | M | 76 | 47,2 | 915 | 59,1 | 1.922 | 53,9 | |
| Stato civile | Celibe/nubile | 54 | 33,5 | 192 | 12,4 | 334 | 9,4 | p<0,05 |
| | Coniugato/a o convivente | 102 | 63,4 | 1.207 | 78,0 | 2.257 | 63,3 | |
| | Separato/a o divorziato/a | 4 | 2,5 | 72 | 4,7 | 79 | 2,2 | |
| | Vedovo/a | 1 | 0,6 | 76 | 4,9 | 898 | 25,2 | |
| Titolo di studio | Licenza elementare o analfabeta | 1 | 0,6 | 442 | 28,6 | 2.902 | 81,3 | p<0,05 |
| | Licenza media inferiore | 61 | 37,9 | 447 | 28,9 | 560 | 15,7 | |
| | Licenza media superiore | 65 | 40,4 | 541 | 35,0 | 93 | 2,6 | |
| | Laurea | 34 | 21,1 | 117 | 7,6 | 13 | 0,4 | |

E' immediato notare (nonché verificato dalla significatività statistica del χ^2) come stato civile e titolo di studio si distribuiscano in maniera coerente con l'età (ad esempio, il livello dei titoli di studio decresce al crescere di essa). La differenza di genere mostra come, pur essendo in presenza di numeri piccoli, le donne incorrano in tale tumore quando sono più giovani rispetto agli uomini, e viceversa. Questa differenza verrà brevemente discussa nel prossimo capitolo.

La tabella 6.15 analizza il rapporto fra condizione professionale e titolo di studio.

Tab. 6.15 – Condizione professionale per titolo di studio.

| | | Licenza elementare o analfabeta | | Licenza media inferiore | | Licenza media superiore | | Laurea | | Sig. χ^2 |
|--------------------------|--|---------------------------------|------|-------------------------|------|-------------------------|------|--------|------|---------------|
| | | N | % | N | % | N | % | N | % | |
| Condizione professionale | Professione intellettuale, scientifica o di elevata specializzazione | 0 | 0,0 | 0 | 0,0 | 0 | 0,0 | 2 | 1,2 | p<0,05 |
| | Professione qualificata nelle attività commerciali o nei servizi | 3 | 0,1 | 3 | 0,3 | 40 | 5,7 | 93 | 56,7 | |
| | Impiegato | 188 | 5,6 | 36 | 3,4 | 226 | 32,3 | 23 | 14,0 | |
| | Artigiano, operaio specializzato o agricoltore | 249 | 7,4 | 364 | 34,1 | 336 | 48,1 | 0 | 0,0 | |
| | Pensionato | 2.589 | 77,4 | 660 | 61,8 | 94 | 13,4 | 0 | 0,0 | |
| | Disoccupato | 2 | 0,1 | 0 | 0,0 | 1 | 0,1 | 0 | 0,0 | |
| | Casalinga | 314 | 9,4 | 5 | 0,5 | 2 | 0,3 | 46 | 28,0 | |

La forte presenza di bassi titoli di studio condiziona la distribuzione delle condizioni professionali: chi è ancora in attività è, per lo più, artigiano, operaio specializzato o agricoltore e, in parte, impiegato, mentre le professione più qualificate, ovviamente, sono proprie dei titoli di studio più elevati.

Le caratteristiche della malattia, della diagnosi e della cura sono presentate nella tabella 6.16 tramite le variabili che, in letteratura, sono più frequentemente utilizzate e che sono state inserite nell'analisi di sopravvivenza. La stessa tabella mostra anche la loro distribuzione nei tre gruppi di età.

I pazienti considerati nello studio presentano uno stadio piuttosto avanzato della malattia (stadio III o IV, complessivamente 61% circa), ossia situazioni di malattia in stadio avanzato (III) o metastatica (IV), che hanno una prognosi sfavorevole o pessima. Mentre lo stadio IV non segue un andamento regolare al crescere dell'età, lo stadio III presenta un incremento significativo. Lo stadio I presenta un picco rilevante di presenze nella fascia d'età centrale interessata dalle attività di prevenzione secondaria, mentre lo stadio II è abbastanza stabile attraverso i tre gruppi.

Per quanto riguarda il grado di malignità, non vi è una relazione statisticamente significativa con l'età dei pazienti. Si nota una maggiore presenza di malattia con gradi a malignità bassa o intermedia (1 e 2, ossia neoplasie cellule ben differenziate), mentre

sono meno numerosi i casi a malignità medio-alta o alta (3 e 4, ossia situazioni a scarsa differenziazione cellulare), che fanno prevedere una prognosi peggiore.

Tab. 6.16 – Caratteristiche cliniche dei pazienti, per età.

| | | Totale | | 0-49 | | 50-69 | | 70+ | | Sig. χ^2 |
|--|----------------------------------|--------|-------|------|-------|-------|-------|-------|-------|---------------|
| | | N | % | N | % | N | % | N | % | |
| Stadio alla diagnosi | I | 832 | 15,8 | 24 | 14,9 | 339 | 21,9 | 469 | 13,1 | p<0,05 |
| | II | 1.231 | 23,3 | 37 | 23,0 | 343 | 22,2 | 851 | 23,9 | |
| | III | 1.359 | 25,8 | 31 | 19,3 | 352 | 22,8 | 976 | 27,4 | |
| | IV | 1.854 | 35,1 | 69 | 42,9 | 513 | 33,2 | 1.272 | 35,7 | |
| Grado alla diagnosi | Grado 1 | 944 | 17,9 | 27 | 16,8 | 286 | 18,5 | 631 | 17,7 | n.s |
| | Grado 2 | 2.399 | 45,5 | 66 | 41,0 | 777 | 50,2 | 1.556 | 43,6 | |
| | Grado 3 | 1.119 | 21,2 | 46 | 28,6 | 341 | 22,0 | 732 | 20,5 | |
| | Grado 4 | 814 | 15,4 | 22 | 13,7 | 143 | 9,2 | 649 | 18,2 | |
| Modalità diagnosi solo clinica | No | 4.444 | 99,7 | 149 | 99,3 | 1.414 | 99,6 | 2.881 | 99,8 | n.s |
| | Diagnosi: solo clinica | 13 | 0,3 | 1 | 0,7 | 5 | 0,4 | 7 | 0,2 | |
| Modalità diagnosi RX (clisma opaco) | No | 4.414 | 99,0 | 149 | 99,3 | 1.408 | 99,2 | 2.857 | 98,9 | n.s. |
| | Diagnosi: RX (clisma opaco) | 43 | 1,0 | 1 | 0,7 | 11 | 0,8 | 31 | 1,1 | |
| Modalità diagnosi ecografia transrettale | No | 4.456 | 100,0 | 150 | 100,0 | 1.419 | 100,0 | 2.887 | 100,0 | n.s |
| | Diagnosi: ecografia transrettale | 1 | 0,0 | 0 | 0,0 | 0 | 0,0 | 1 | 0,0 | |
| Modalità diagnosi endoscopia | No | 1.798 | 40,3 | 65 | 43,3 | 527 | 37,1 | 1.206 | 41,8 | p<0,05 |
| | Diagnosi: endoscopia | 2.659 | 59,7 | 85 | 56,7 | 892 | 62,9 | 1.682 | 58,2 | |
| Modalità diagnosi RMN | No | 4.446 | 99,8 | 149 | 99,3 | 1.416 | 99,8 | 2.881 | 99,8 | n.s. |
| | Diagnosi: RMN | 11 | 0,2 | 1 | 0,7 | 3 | 0,2 | 7 | 0,2 | |
| Modalità diagnosi Tc spirale | No | 4.341 | 97,4 | 145 | 96,7 | 1.381 | 97,3 | 2.815 | 97,5 | n.s |
| | Diagnosi: Tc spirale | 116 | 2,6 | 5 | 3,3 | 38 | 2,7 | 73 | 2,5 | |
| Modalità diagnosi SOF | No | 4.454 | 99,9 | 149 | 99,3 | 1.419 | 100,0 | 2.886 | 99,9 | p<0,05 |
| | Diagnosi: SOF | 3 | 0,1 | 1 | 0,7 | 0 | 0,0 | 2 | 0,1 | |
| Modalità diagnosi screening | No | 4.425 | 99,3 | 150 | 100,0 | 1.397 | 98,4 | 2.878 | 99,7 | p<0,05 |
| | Diagnosi: screening | 32 | 0,7 | 0 | 0,0 | 22 | 1,6 | 10 | 0,3 | |
| Comorbidità alla diagnosi | No comorbidità | 4.296 | 81,4 | 149 | 92,5 | 1.292 | 83,5 | 2.855 | 80,0 | p<0,05 |
| | Comorbidità | 980 | 18,6 | 12 | 7,5 | 255 | 16,5 | 713 | 20,0 | |
| | Trattato | 3.998 | 75,8 | 143 | 88,8 | 1.366 | 88,3 | 2.489 | 69,8 | |

Nota: n.s. = assenza di significatività statistica.

(Continua)

(Continua)

| | | Totale | | 0-49 | | 50-69 | | 70+ | | Sig. χ^2 |
|--------------------------------|------------------|--------|------|------|------|-------|------|-------|------|---------------|
| | | N | % | N | % | N | % | N | % | |
| Chemioterapia | No chemioterapia | 4.332 | 82,1 | 87 | 54,0 | 1.028 | 66,5 | 3.217 | 90,2 | p<0,05 |
| | Chemioterapia | 944 | 17,9 | 74 | 46,0 | 519 | 33,5 | 351 | 9,8 | |
| Radioterapia | No radioterapia | 5.239 | 99,3 | 159 | 98,8 | 1.526 | 98,6 | 3.554 | 99,6 | p<0,05 |
| | Radioterapia | 37 | 0,7 | 2 | 1,2 | 21 | 1,4 | 14 | 0,4 | |
| Immunoterapia | No immunoterapia | 5.220 | 98,9 | 154 | 95,7 | 1.511 | 97,7 | 3.555 | 99,6 | p<0,05 |
| | Immunoterapia | 56 | 1,1 | 7 | 4,3 | 36 | 2,3 | 13 | 0,4 | |
| Chirurgia | No chirurgia | 1.327 | 25,2 | 24 | 14,9 | 201 | 13,0 | 1.102 | 30,9 | p<0,05 |
| | Chirurgia | 3.949 | 74,8 | 137 | 85,1 | 1.346 | 87,0 | 2.466 | 69,1 | |
| Pazienti trattati/non trattati | Non trattato | 1.278 | 24,2 | 18 | 11,2 | 181 | 11,7 | 1.079 | 30,2 | p<0,05 |

Nota: n.s. = assenza di significatività statistica.

La modalità di diagnosi più utilizzata è di tipo endoscopico (colonscopia o sigmoidoscopia) che riguarda quasi il 60% dei pazienti, con un incremento (63% circa) nella fascia d'età interessata dalle attività di screening organizzato. Meno utilizzate le altre modalità di diagnosi.

E' legata all'età la presenza di comorbidità: il 19% dei casi in esame mostra presenza di altre patologie al momento della diagnosi, con percentuali crescenti al crescere dell'età.

Per quanto riguarda le terapie, la chirurgia è, come da protocolli, la cura più utilizzata (in circa il 75% dei casi), ma si nota una riduzione della sua applicazione al crescere dell'età. La chemioterapia è il secondo percorso di cura, spesso assieme alla chirurgia, ma la sua applicazione decresce fortemente al crescere dell'età. Immunoterapia e radioterapia sono residuali nel percorso di cura, in particolare la radioterapia che, nel tumore al colon, non è prevista nei protocolli di cura usati più frequentemente.

Per questa neoplasia, infine, i pazienti non trattati ammontano al 24% e sono per lo più ultrasessantenni che combinano età e stadio avanzati (dato tabulato riportato in appendice). Questi sono ovviamente pazienti per cui la prognosi è tendenzialmente negativa (il 77% sopravvive meno di tre anni alla diagnosi).

La tabella 6.17, infine, presenta i valori medi percentuali delle caratteristiche (utilizzate come covariate nel modello di sopravvivenza) delle sezioni di Censimento in cui sono residenti gli individui compresi nell'analisi, rispetto alla loro distribuzione regionale.

Tab. 6.17 – Caratteristiche delle sezioni di Censimento dei pazienti a confronto con quelle regionali.

| | Casi in analisi | | | Regione |
|--|--------------------------|--------------------------------|--------------------------|-------------------------|
| | | Intervallo di confidenza (95%) | | |
| | Valore medio aree | <i>Estremo inferiore</i> | <i>Estremo superiore</i> | Valore regionale |
| Cittadinanza italiana (%) | 96,4 | 96,2 | 96,6 | 96,6 |
| Cittadinanza straniera (%) | 3,3 | 3,2 | 3,4 | 3,4 |
| Numero persone per famiglia | 2,6 | 2,6 | 2,6 | 2,6 |
| Dipendenti o in altra posizione subordinata (%) | 70,1 | 69,8 | 70,4 | 68,8 |
| Imprenditori (%) | 3,5 | 3,4 | 3,6 | 3,3 |
| Liberi professionisti (%) | 5,1 | 4,9 | 5,3 | 3,4 |
| Lavoratori in proprio (%) | 16,6 | 16,4 | 16,9 | 20,0 |
| Soci di cooperative (%) | 2,0 | 1,9 | 2,1 | 1,9 |
| Coadiuvanti familiari (%) | 1,7 | 1,6 | 1,7 | 2,0 |
| % Occupati | 43,0 | 42,7 | 43,2 | 41,5 |
| % In cerca di prima occupazione | 1,5 | 1,5 | 1,6 | 1,5 |
| % Disoccupati (in cerca di nuova occupazione) | 2,8 | 2,7 | 2,8 | 2,8 |
| % In attesa di iniziare un lavoro già trovato | 0,3 | 0,3 | 0,3 | 0,3 |
| % Studenti | 7,5 | 7,4 | 7,5 | 6,7 |
| % Casalinghe | 14,3 | 14,1 | 14,4 | 15,4 |
| % Ritirati dal lavoro | 25,5 | 25,3 | 25,7 | 26,4 |
| % In servizio di leva o in servizio civile sostitutivo | 0,2 | 0,2 | 0,2 | 0,2 |
| % Inabile al lavoro | 1,5 | 1,4 | 1,5 | 1,7 |
| % In altra condizione | 3,2 | 3,1 | 3,4 | 3,6 |
| Celibe / Nubile (%) | 35,6 | 35,4 | 35,8 | 35,1 |
| Coniugato/a (%) | 51,3 | 51,1 | 51,5 | 52,7 |
| Separato/a di fatto (%) | 0,4 | 0,4 | 0,4 | 0,3 |
| Separato/a legalmente (%) | 1,4 | 1,4 | 1,4 | 1,2 |
| Divorziato/a (%) | 1,3 | 1,2 | 1,3 | 0,8 |
| Vedovo/a (%) | 9,8 | 9,7 | 9,9 | 9,9 |
| Non sa leggere o scrivere (%) | 1,1 | 1,1 | 1,2 | 1,6 |
| Sa leggere e scrivere (%) | 7,4 | 7,3 | 7,5 | 9,8 |
| Licenza di scuola elementare (%) | 24,4 | 24,2 | 24,7 | 27,7 |
| Licenza di scuola media inferiore (%) | 25,8 | 25,6 | 26,0 | 27,7 |
| Diploma di scuola secondaria superiore (%) | 31,3 | 31,0 | 31,5 | 27,6 |
| Diploma non universitario post maturità (%) | 0,5 | 0,5 | 0,5 | 0,4 |
| Diploma universitario (%) | 0,9 | 0,8 | 0,9 | 0,6 |
| Laurea (%) | 8,3 | 8,1 | 8,5 | 4,7 |

(Continua)

(Continua)

| | Casi in analisi | | | Regione |
|---|--------------------------------|--------------------------|--------------------------|-------------------------|
| | Intervallo di confidenza (95%) | | | |
| | Valore medio aree | <i>Estremo inferiore</i> | <i>Estremo superiore</i> | Valore regionale |
| Indice di ricambio | 161,2 | 157,6 | 164,7 | 144,5 |
| Indice di vecchiaia | 238,0 | 231,9 | 244,2 | 220,8 |
| Indice di dipendenza strutturale | 57,7 | 57,0 | 58,4 | 59,2 |
| Tasso di attività | 65,1 | 64,9 | 65,4 | 64,2 |
| Tasso di occupazione | 58,8 | 58,6 | 59,1 | 57,8 |
| Tasso di disoccupazione | 9,1 | 9,0 | 9,3 | 9,4 |
| % alloggi di proprietà | 66,7 | 66,3 | 67,1 | 59,1 |
| % alloggi in affitto | 13,2 | 12,9 | 13,5 | 8,4 |
| % alloggi con altro titolo di godimento | 8,4 | 8,2 | 8,6 | 8,2 |
| % alloggi senza alcuna specifica | 10,9 | 10,6 | 11,2 | 24,2 |

Anche per questa patologia le differenze mostrano che, tendenzialmente, le persone considerate nell'analisi sono residenti in un insieme di sezioni di Censimento più ricco del resto della regione, in quanto si registrano migliori percentuali nell'area dell'occupazione, maggiore presenza di titoli di studio elevati, occupazioni a reddito potenzialmente più alto, migliori indicatori di sintesi (ad esempio, indice di ricambio più elevato e indice di dipendenza strutturale più basso), a parte l'indice di vecchiaia che, invece, risulta più elevato in maniera significativa.

6.2.2 I risultati del modello di sopravvivenza a parametri flessibili ed effetti misti

La tabella 6.18 riporta il risultato del confronto fra i diversi modelli di sopravvivenza implementati.

In essa sono evidenziati i gruppi di variabili utilizzati per ogni modello (a partire dal modello base che comprende età e stadio per quanto riguarda gli effetti fissi e la sezione di Censimento di residenza per quanto riguarda gli effetti casuali), il numero di casi utilizzati in ogni modello e il valore dell'AIC.

Tab. 6.18 – Modelli di sopravvivenza a parametri flessibili ed effetti misti implementati: caratteristiche e confronto.

| Età | Stadio | Variabili diagnostiche | Comorbidità | Variabili terapeutiche | Effetti fissi Caratteristiche socio-economiche individuali | Caratteristiche socio-economiche sezione di Censimento | Effetti casuali Sezione di Censimento di residenza | N | AIC |
|-----|--------|------------------------|-------------|------------------------|---|--|---|-------|----------|
| X | X | | | | | | X | 5.276 | 11.296,6 |
| X | X | | | | X | | X | 5.276 | 11.293,9 |
| X | X | | | | | X | X | 5.276 | 11.337,7 |
| X | X | X | | | | | X | 5.276 | 11.115,4 |
| X | X | X | X | | | | X | 5.276 | 11.113,7 |
| X | X | X | | | X | | X | 5.276 | 11.119,1 |
| X | X | X | | | | X | X | 5.276 | 11.151,6 |
| X | X | X | X | | X | | X | 5.276 | 11.117,5 |
| X | X | X | X | | | X | X | 5.276 | 11.149,9 |
| X | X | | | | X | X | X | 5.276 | 11.332,9 |
| X | X | X | | | X | X | X | 5.276 | 11.154,3 |
| X | X | X | X | | X | X | X | 5.276 | 11.152,6 |
| X | X | | | X | | | X | 5.276 | 11.262,8 |
| X | X | | | X | X | | X | 5.276 | 11.260,1 |
| X | X | | | X | | X | X | 5.276 | 11.303,9 |
| X | X | X | | X | | | X | 5.276 | 11.101,4 |
| X | X | X | X | X | | | X | 5.276 | 11.101,6 |
| X | X | X | | X | X | | X | 5.276 | 11.103,9 |
| X | X | X | | X | | X | X | 5.276 | 11.097,8 |
| X | X | X | X | X | X | | X | 5.276 | 11.096,3 |
| X | X | X | X | X | | X | X | 5.276 | 11.098,2 |
| X | X | | | X | X | X | X | 5.276 | 11.299,8 |
| X | X | X | | X | X | X | X | 5.276 | 11.099,4 |
| X | X | X | X | X | X | X | X | 5.276 | 11.095,4 |

Anche per il tumore al colon, l'AIC identifica come modello da preferire quello in cui tutte le aree sono rappresentate. Le riduzioni più marcate del valore dell'AIC si ottengono quando nel modello sono inserite le variabili relative alla diagnosi, ai trattamenti terapeutici e alle caratteristiche socio-economiche individuali, mentre effetti di riduzione minori si registrano quando si inseriscono le comorbidità e le caratteristiche dell'area di residenza.

Questi effetti sembrano confermati dai risultati ottenuti con il modello di sopravvivenza illustrato nella tabella 6.19, che corrisponde a quello identificato dall'AIC come il più adeguato.

Solo le covariate che hanno una relazione statisticamente significativa con il rischio di morte sono state riportate in tale tabella; per rendere più agevole la lettura, sono state escluse tutte le variabili e le interazioni a due vie non statisticamente significative e quelle che non hanno superato il test di collinearità.

Il modello presenta un valore di verosimiglianza pari a $-5.492,7$, statisticamente significativo, raggiunto utilizzando una quadratura di Gauss-Hermite a 16 nodi. Il modello è quindi da considerarsi attendibile.

Osservando gli effetti principali, si nota subito che lo stadio è anche qui il principale predittore del rischio di morte, con eccessi di rischio fortemente crescenti quando esso più avanzato alla diagnosi, confermando quindi quanto noto in letteratura e commentato in precedenza.

Tab. 6.19 – Il modello di sopravvivenza relativa a parametri flessibili ed effetti misti.

| | | | | Intervallo di confidenza (95%) | | |
|---|---|-----------------------|-------------|---|-----------------|--------|
| | | HR | Sig. | Est. Inf. | Est. Sup | |
| Effetti fissi | Stadio | Stadio I | 1 | <i>(categoria di riferimento)</i> | | |
| | | Stadio II | 3,85 | 0,02 | 1,20 | 12,36 |
| | | Stadio III | 12,97 | 0,00 | 4,19 | 40,16 |
| | | Stadio IV | 94,81 | 0,00 | 30,90 | 290,88 |
| | Diagnosi | Grado 1 | 1 | <i>(categoria di riferimento)</i> | | |
| | | Grado 2 | 1,73 | 0,00 | 1,35 | 2,22 |
| | | Grado 3 | 3,18 | 0,00 | 2,45 | 4,13 |
| | | Grado 4 | 3,80 | 0,00 | 2,92 | 4,95 |
| | Terapia | No chemioterapia | 1 | <i>(categoria di riferimento)</i> | | |
| | | Sì chemioterapia | 0,61 | 0,00 | 0,53 | 0,72 |
| | | No terapia chirurgica | 1 | <i>(categoria di riferimento)</i> | | |
| | | Sì terapia chirurgica | 0,52 | 0,00 | 0,30 | 0,77 |
| | SE individuali | Coniugata/convivente | 1 | <i>(categoria di riferimento)</i> | | |
| | | Celibe/Nubile | 1,27 | 0,13 | 0,95 | 1,55 |
| | | Separato/divorziato | 0,98 | 0,88 | 0,70 | 1,36 |
| | | Vedovo | 1,20 | 0,03 | 1,02 | 1,41 |
| | SE d'area | Indice di vecchiaia | 1,04 | 0,04 | 1,00 | 1,09 |
| Effetti fissi di interazione a due vie | 70+ # Stadio II | 6,13 | 0,01 | 2,16 | 10,54 | |
| | 50-69 # Diploma media superiore | 0,35 | 0,00 | 0,18 | 0,67 | |
| | Stadio III # Chemioterapia | 0,45 | 0,04 | 0,12 | 0,71 | |
| | Stadio II # Licenza elementare o analfabeta | 5,32 | 0,00 | 3,21 | 7,98 | |
| | Stadio II # Licenza media inferiore | 4,77 | 0,01 | 2,35 | 7,67 | |
| Effetti casuali | Sezione di Censimento | 0,91 | 0,43 | 0,35 | 1,56 | |

Nota: HR = hazard ratio(eccesso di rischio di morte); SE = variabili socio-economiche.

Per quanto riguarda le variabili legate alla diagnosi, solo il grado di malignità è mantenuto nel modello: gli eccessi di rischio di morte si incrementano tanto più il tumore è maligno, ossia la differenziazione cellulare con i tessuti sani è minore.

Dalle covariate che descrivono le terapie restano escluse l'immunoterapia e la radioterapia che, come visto in precedenza, riguardano però solo una parte minima dei pazienti. Restano, invece in gioco chirurgia e chemioterapia che mostrano un effetto di forte riduzione degli eccessi di rischio (circa il 48% per la prima e circa il 39% per la seconda).

Fra le variabili socio-economiche individuali, solo lo stato civile presenta un effetto statisticamente significativo: i pazienti vedovi fanno registrare un eccesso di rischio pari al 20%, i cui effetti sulla diagnosi e la cura (quindi effetti indiretti rispetto alla sopravvivenza) verranno analizzati nel paragrafo successivo.

Gli effetti socio-economici relativi alla sezione di Censimento di residenza mostrano un ruolo statisticamente significativo dell'indice di vecchiaia, ossia dell'indice ISTAT che misura il rapporto tra popolazione con 65 e più anni e popolazione con 14 o meno anni. Al crescere del valore dell'indice aumenta del 4% circa l'eccesso del rischio di morte, generando un effetto debole, ma statisticamente significativo.

L'età, esclusa dagli effetti principali, entra in gioco negli effetti legati alle interazioni a due vie: l'effetto combinato di età anziana (70 e più anni) con stadio II porta ad un eccesso di rischio di morte di oltre il 600%.

Sono rilevanti anche gli effetti combinati fra basso titolo di studio (licenza elementare o di media inferiore) e stadio II: anche qui, a fronte di uno stadio che potrebbe portare ad una buona prognosi, l'effetto combinato con i titoli di studio bassi porta ad eccessi di rischio molto alti.

Questo effetto rende interessante analizzare, nel paragrafo successivo, non solo gli effetti di stato civile e indice di vecchiaia, ma anche del titolo di studio su diagnosi e cure.

Un effetto protettivo è invece dato dalla combinazione fra stadio III e applicazione della chemioterapia, che riduce sensibilmente l'eccesso di rischio di morte anche in presenza di uno stadio avanzato.

Infine, la sezione di Censimento di residenza, ossia la variabile di *clustering* responsabile per gli effetti casuali, non ha un ruolo statisticamente significativo nel modello.

6.2.3 L'analisi degli effetti dello stato civile e del titolo di studio sulla diagnosi e sulle cure

Nel tumore al colon, lo stato civile presenta un ruolo rilevante fra le covariate che producono gli effetti fissi, mentre il titolo di studio quando è basso agisce in maniera importante nell'interazione con lo stadio II aumentando gli eccessi di rischio di morte,

in contrasto con il fatto che tale stadio, tendenzialmente, permette una buona prognosi di sopravvivenza.

In questo paragrafo vengono esaminati, quindi, il ruolo sia dello stato civile che del titolo di studio nell'influenzare la tempestività della diagnosi (ossia lo stadio) e le cure a cui i pazienti hanno avuto accesso.

Le tabelle 6.20 e 6.21 mostrano come si distribuiscono lo stadio alla diagnosi e le terapie rispetto alle due variabili socio-economiche, valutandone la connessione tramite significatività statistica del test del χ^2 ($p < 0,05$).

Tab. 6.20 – Stato civile, stadio alla diagnosi e terapie seguite.

| | | Stato civile | | | | | | | | Sig. χ^2 |
|--------------------------------|------------------|---------------|------|--------------------------|------|---------------------------|-------|----------|------|---------------|
| | | Celibe/nubile | | Coniugato/a o convivente | | Separato/a o divorziato/a | | Vedovo/a | | |
| | | N | % | N | % | N | % | N | % | |
| Stadio alla diagnosi | I | 99 | 17,1 | 611 | 17,1 | 15 | 9,7 | 107 | 11,0 | p<0,05 |
| | II | 129 | 22,2 | 842 | 23,6 | 42 | 27,1 | 218 | 22,4 | |
| | III | 143 | 24,7 | 941 | 26,4 | 49 | 31,6 | 226 | 23,2 | |
| | IV | 209 | 36,0 | 1.172 | 32,9 | 49 | 31,6 | 424 | 43,5 | |
| Chemioterapia | No chemioterapia | 457 | 78,8 | 2.857 | 80,1 | 122 | 78,7 | 896 | 91,9 | p<0,05 |
| | Chemioterapia | 123 | 21,2 | 709 | 19,9 | 33 | 21,3 | 79 | 8,1 | |
| Immunoterapia | No immunoterapia | 576 | 99,3 | 3.518 | 98,7 | 154 | 99,4 | 972 | 99,7 | p<0,05 |
| | Immunoterapia | 4 | 0,7 | 48 | 1,3 | 1 | 0,6 | 3 | 0,3 | |
| Radioterapia | No radioterapia | 578 | 99,7 | 3.534 | 99,1 | 155 | 100,0 | 972 | 99,7 | n.s. |
| | Radioterapia | 2 | 0,3 | 32 | 0,9 | 0 | 0,0 | 3 | 0,3 | |
| Chirurgia | No chirurgia | 139 | 24,0 | 772 | 21,6 | 43 | 27,7 | 373 | 38,3 | p<0,05 |
| | Chirurgia | 441 | 76,0 | 2.794 | 78,4 | 112 | 72,3 | 602 | 61,7 | |
| Pazienti trattati/non trattati | Non trattato | 130 | 22,4 | 740 | 20,8 | 40 | 25,8 | 368 | 37,7 | p<0,05 |
| | Trattato | 450 | 77,6 | 2.826 | 79,2 | 115 | 74,2 | 607 | 62,3 | |

Si può notare come lo stadio IV sia più frequente nei celibi/nubili e nei vedovi, mentre gli stadi II e III sono distribuiti piuttosto equamente fra i vari stati civili; lo stadio I è invece meno frequente nei vedovi e nei separati/divorziati.

Rispetto alle terapie si osserva come, tendenzialmente, i vedovi vedano una minore adesione ai protocolli di cura, radioterapia a parte. Questo sotto-trattamento dei vedovi si evidenzia anche nella variabile dicotomica fra pazienti trattati e non trattati, con più di un terzo di essi che ricade nel secondo gruppo.

Tab. 6.21 – Titolo di studio, stadio alla diagnosi e terapie seguite.

| | | Stato civile | | | | | | | | Sig. χ^2 |
|--------------------------------|------------------|--------------------|------|-------------------------|------|-------------------------|------|--------|------|---------------|
| | | Licenza elementare | | Licenza media inferiore | | Diploma media superiore | | Laurea | | |
| | | N | % | N | % | N | % | N | % | |
| Stadio alla diagnosi | I | 471 | 14.1 | 178 | 16.7 | 149 | 21.3 | 34 | 20.7 | p<0,05 |
| | II | 815 | 24.4 | 239 | 22.4 | 147 | 21.0 | 30 | 18.3 | |
| | III | 869 | 26.0 | 285 | 26.7 | 165 | 23.6 | 40 | 24.4 | |
| | IV | 1.190 | 35.6 | 366 | 34.3 | 238 | 34.0 | 60 | 36.6 | |
| Chemioterapia | No chemioterapia | 2.937 | 87.8 | 819 | 76.7 | 476 | 68.1 | 100 | 61.0 | p<0,05 |
| | Chemioterapia | 408 | 12.2 | 249 | 23.3 | 223 | 31.9 | 64 | 39.0 | |
| Immunoterapia | No immunoterapia | 3.323 | 99.3 | 1.057 | 99.0 | 686 | 98.1 | 154 | 93.9 | p<0,05 |
| | Immunoterapia | 22 | 0.7 | 11 | 1.0 | 13 | 1.9 | 10 | 6.1 | |
| Radioterapia | No radioterapia | 3.325 | 99.4 | 1.060 | 99.3 | 692 | 99.0 | 162 | 98.8 | n.s. |
| | Radioterapia | 20 | 0.6 | 8 | 0.7 | 7 | 1.0 | 2 | 1.2 | |
| Chirurgia | No chirurgia | 953 | 28.5 | 250 | 23.4 | 106 | 15.2 | 18 | 11.0 | p<0,05 |
| | Chirurgia | 2.392 | 71.5 | 818 | 76.6 | 593 | 84.8 | 146 | 89.0 | |
| Pazienti trattati/non trattati | Non trattato | 931 | 27,8 | 235 | 22,0 | 96 | 13,7 | 16 | 9,8 | p<0,05 |
| | Trattato | 2.414 | 72,2 | 833 | 78,0 | 603 | 86,3 | 148 | 90,2 | |

Considerando la relazione fra stadio e titolo di studio, si può osservare che, mentre la presenza di malattie in stadio I alla diagnosi cresce al crescere del titolo di studio, gli stadi III e IV sono sostanzialmente stabili attraverso i vari titoli, la presenza di stadi II decresce al crescere del livello di istruzione.

Per quanto riguarda i trattamenti, le applicazioni di chirurgia e chemioterapia (e anche immunoterapia) crescono al crescere del titolo di studio, mentre l'applicazione di radioterapia è sostanzialmente stabile.

Sul totale della differenza fra pazienti trattati e non trattati, emerge anche qui come la percentuale di applicazione dei curati aumenti all'aumentare del titolo di studio.

Una lettura più dettagliata del ruolo di tali covariate su tempestività alla diagnosi e adesione ai protocolli di cura è data dai risultati dei modelli di regressione logistica presentati nelle tabelle 6.22 e 6.23, in cui è utilizzato anche l'indice di vecchiaia, unica variabile socio-economica d'area che ha mostrato un effetto statisticamente significativo sulla sopravvivenza.

Rispetto allo stadio, si osserva come vedovi e separati/divorziati abbiano meno probabilità di essere diagnosticati ad uno stadio I, così come le persone in possesso della licenza elementare. L'essere vedovi aumenta drasticamente le probabilità di essere diagnosticati in maniera tardiva (63% di probabilità in più di avere un tumore in stadio IV), mentre è meno chiaro l'effetto di "protezione" che tale stato civile ha rispetto allo stadio III. Il titolo di studio rientra in gioco nello stadio II, dove si registra un 52% di probabilità di essere diagnosticati in tale situazione se si ha un titolo elementare,

Per quanto riguarda le cure seguite³⁹, si osserva una generale riduzione delle probabilità di essere adeguatamente trattati se si è celibi/nubili o vedovi, in possesso di titoli di studio bassi o medio bassi.

Costoro, infatti, hanno probabilità più basse di ricevere chemioterapia, immunoterapia o chirurgia. Poiché, però, non sembra invece esserci una relazione fra l'essere trattati o meno e le caratteristiche socio-economiche (solo chi è in possesso di licenza media inferiore vede un 78% di probabilità di non essere trattato), ciò significa che ai pazienti con tali stati civili e titoli di studio i trattamenti vengono forniti in alternativa l'uno all'altro, sfruttando poco meno l'effetto fortemente adiuvante di trattamenti combinati. Non porta effetti su diagnosi e trattamento, invece, l'indice di vecchiaia che caratterizza la sezione di Censimento di residenza.

³⁹ La radioterapia è stata esclusa dall'analisi, perché applicata ad un numero eccessivamente basso di pazienti, non essendo trattamento elettivo per il tumore al colon.

Tab. 6.22 – Stadio alla diagnosi e variabili socio-economiche: modelli di regressione logistica.

| Stadio alla diagnosi | | OR | Sig. |
|----------------------|-----------------------------------|------|----------------------------|
| Stadio I | Coniugato/convivente | 1 | (categoria di riferimento) |
| | Celibe/nubile | 0,88 | n.s. |
| | Separato/divorziato | 0,50 | p<0,05 |
| | Vedovo | 0,64 | p<0,05 |
| | Licenza elementare (o analfabeta) | 0,63 | p<0,05 |
| | Licenza media inferiore | 0,82 | n.s. |
| | Diploma media superiore | 0,94 | n.s. |
| | Laurea | 1 | (categoria di riferimento) |
| | Indice di vecchiaia | 1,00 | n.s. |
| | Costante | 0,29 | p<0,05 |
| Stadio II | Coniugato/convivente | 1 | (categoria di riferimento) |
| | Celibe/nubile | 0,96 | n.s. |
| | Separato/divorziato | 1,24 | n.s. |
| | Vedovo | 0,89 | n.s. |
| | Licenza elementare (o analfabeta) | 1,52 | p<0,05 |
| | Licenza media inferiore | 1,37 | n.s. |
| | Diploma media superiore | 1,23 | n.s. |
| | Laurea | 1 | (categoria di riferimento) |
| | Indice di vecchiaia | 1,00 | n.s. |
| | Costante | 0,21 | p<0,05 |
| Stadio III | Coniugato/convivente | 1 | (categoria di riferimento) |
| | Celibe/nubile | 0,76 | n.s. |
| | Separato/divorziato | 1,24 | n.s. |
| | Vedovo | 0,80 | p<0,05 |
| | Licenza elementare (o analfabeta) | 1,18 | n.s. |
| | Licenza media inferiore | 1,39 | n.s. |
| | Diploma media superiore | 0,99 | n.s. |
| | Laurea | 1 | (categoria di riferimento) |
| | Indice di dipendenza strutturale | 1,00 | n.s. |
| | Costante | 0,32 | p<0,05 |
| Stadio IV | Coniugato/convivente | 1 | (categoria di riferimento) |
| | Celibe/nubile | 1,40 | p<0,05 |
| | Separato/divorziato | 0,97 | n.s. |
| | Vedovo | 1,63 | p<0,05 |
| | Licenza elementare (o analfabeta) | 0,85 | n.s. |
| | Licenza media inferiore | 0,70 | n.s. |
| | Diploma media superiore | 0,89 | n.s. |
| | Laurea | 1 | (categoria di riferimento) |
| | Indice di vecchiaia | 1,00 | n.s. |
| | Costante | 0,57 | p<0,05 |

Nota: OR = Odds ratio.

Tab. 6.23 – Terapie e variabili socio-economiche: modelli di regressione logistica.

| Terapie | | OR | Sig. |
|--------------------------------|-----------------------------------|------|----------------------------|
| Chemioterapia | Coniugato/convivente | 1 | (categoria di riferimento) |
| | Celibe/nubile | 0,60 | p<0,05 |
| | Separato/divorziato | 0,88 | n.s. |
| | Vedovo | 0,41 | p<0,05 |
| | Licenza elementare (o analfabeta) | 0,25 | p<0,05 |
| | Licenza media inferiore | 0,69 | p<0,05 |
| | Diploma media superiore | 0,71 | n.s. |
| | Laurea | 1 | (categoria di riferimento) |
| | Indice di dipendenza strutturale | 1,00 | n.s. |
| | Costante | 0,64 | p<0,05 |
| Immunoterapia | Coniugato/convivente | 1 | (categoria di riferimento) |
| | Celibe/nubile | 0,35 | n.s. |
| | Separato/divorziato | 0,29 | n.s. |
| | Vedovo | 0,30 | p<0,05 |
| | Licenza elementare (o analfabeta) | 0,11 | p<0,05 |
| | Licenza media inferiore | 0,29 | p<0,05 |
| | Diploma media superiore | 0,26 | p<0,05 |
| | Laurea | 1 | (categoria di riferimento) |
| | Indice di dipendenza strutturale | 1,00 | n.s. |
| | Costante | 0,10 | p<0,05 |
| Chirurgia | Coniugato/convivente | 1 | (categoria di riferimento) |
| | Celibe/nubile | 0,69 | p<0,05 |
| | Separato/divorziato | 0,65 | p<0,05 |
| | Vedovo | 0,49 | p<0,05 |
| | Licenza elementare (o analfabeta) | 0,35 | p<0,05 |
| | Licenza media inferiore | 0,52 | p<0,05 |
| | Diploma media superiore | 0,63 | n.s. |
| | Laurea | 1 | (categoria di riferimento) |
| | Indice di dipendenza strutturale | 1,00 | n.s. |
| | Costante | 8,83 | p<0,05 |
| Pazienti trattati/non trattati | Coniugato/convivente | 1 | (categoria di riferimento) |
| | Celibe/nubile | 0,78 | n.s. |
| | Separato/divorziato | 0,88 | n.s. |
| | Vedovo | 0,87 | n.s. |
| | Licenza elementare (o analfabeta) | 1,44 | n.s. |
| | Licenza media inferiore | 1,78 | p<0,05 |
| | Diploma media superiore | 1,46 | n.s. |
| | Laurea | 1 | (categoria di riferimento) |
| | Indice di vecchiaia | 1,00 | n.s. |
| | Costante | 0,24 | p<0,05 |

Nota: OR = Odds ratio.

6.3 Il tumore al retto

Dopo la selezione dei casi presentata nella tabella 6.12 sopra descritta, i pazienti affetti da tumore del retto utilizzati nelle analisi di sopravvivenza sono 2.514.

6.3.1 Le caratteristiche di base

La tabella 6.24 mostra le caratteristiche socio-economiche individuali delle persone affette dal tumore al retto, considerate nello studio.

Tab. 6.24 – Le caratteristiche demografiche e socio-economiche individuali dei pazienti.

| Caratteristiche individuali | | N | % |
|-----------------------------|--|--------------|------|
| Genere | F | 1.043 | 41,5 |
| | M | 1.471 | 58,5 |
| Età all'incidenza | 0-49 | 89 | 3,5 |
| | 50-69 | 779 | 31,0 |
| | 70+ | 1.646 | 65,5 |
| Stato civile | Celibe/nubile | 267 | 10,6 |
| | Coniugato/a o convivente | 1.737 | 69,1 |
| | Separato/a o divorziato/a | 70 | 2,8 |
| | Vedovo/a | 440 | 17,5 |
| Titolo di studio | Licenza elementare o analfabeta | 1.563 | 62,2 |
| | Licenza media inferiore | 495 | 19,7 |
| | Licenza media superiore | 373 | 14,8 |
| | Laurea | 83 | 3,3 |
| Condizione professionale | Professione intellettuale, scientifica o di elevata specializzazione | 2 | 0,1 |
| | Professione qualificata nelle attività commerciali o nei servizi | 71 | 2,8 |
| | Impiegato | 203 | 8,1 |
| | Artigiano, operaio specializzato o agricoltore | 540 | 21,5 |
| | Pensionato | 1.530 | 60,9 |
| | Disoccupato | 2 | 0,1 |
| | Casalinga | 166 | 6,6 |
| Totale | | 2.514 | |

Così come per i tumori al colon, anche per le neoplasie al retto vi sono un po' più uomini che donne, con una proporzione che rispecchia perfettamente quella registrata su tutta la popolazione italiana colpita da tale neoplasia (dati ITACAN, AIRTUM, 2014).

Rispetto alle età, i pazienti sono in maggioranza appartenenti alla fascia anziana (il 66% ha 70 e più anni), mentre il 31% ha avuto diagnosi di tumore nella fascia di età interessata dallo screening organizzato. Molto bassa la percentuale di giovani affetti da tale tumore (attorno al 4%).

Rispetto ai dati italiani per lo stesso periodo (dati ITACAN, AIRTUM, 2014), si osservano anche per questa sede proporzioni molto simili.

Gli individui considerati sono per lo più coniugati (o conviventi), poi vedovi, celibi/nubili e, infine, separati/divorziati in bassa percentuale. Rispetto al dato dell'intera popolazione umbra (dati ISTAT, 2001), si nota una maggior presenza di vedovi (14% circa in tutta la regione), ma una percentuale decisamente più alta di coniugati/conviventi (49% in tutta la regione) e più bassa di celibi/nubili (35% in tutta la regione).

Rispetto al titolo di studio, si osserva una presenza elevata di titoli elementari, seguiti dalle licenze medie inferiori, dai diplomi di scuola superiore e, infine dalle, lauree. Le percentuali dei titoli inferiori sono più alte dei valori regionali e nazionali, mentre diplomi di scuola secondaria superiore e lauree sono molto meno presenti rispetto a regione e nazione (Vercelli et al., 2013).

La condizione professionale dominante è l'essere pensionati, coerentemente con la distribuzione delle età; rilevante pure la presenza di artigiani, operai specializzati e agricoltori, meno presenti le altre occupazioni.

Le caratteristiche sembrano quindi essere connesse alle età dei malati; la tabella 6.25 analizza la distribuzione di genere, stato civile e titolo di studio per i tre gruppi di età.

Tab. 6.25 – Stato civile e titolo di studio per età alla diagnosi.

| | | 0-49 | | 50-69 | | 70+ | | Sig. χ^2 |
|------------------|---------------------------------|------|------|-------|------|-------|------|---------------|
| | | N | % | N | % | N | % | |
| Genere | F | 42 | 47,2 | 263 | 33,8 | 738 | 44,8 | p<0,05 |
| | M | 47 | 52,8 | 516 | 66,2 | 908 | 55,2 | |
| Stato civile | Celibe/nubile | 34 | 38,2 | 96 | 12,3 | 137 | 8,3 | p<0,05 |
| | Coniugato/a o convivente | 47 | 52,8 | 606 | 77,8 | 1.084 | 65,9 | |
| | Separato/a o divorziato/a | 4 | 4,5 | 43 | 5,5 | 23 | 1,4 | |
| | Vedovo/a | 4 | 4,5 | 34 | 4,4 | 402 | 24,4 | |
| Titolo di studio | Licenza elementare o analfabeta | 2 | 2,2 | 213 | 27,3 | 1.348 | 81,9 | p<0,05 |
| | Licenza media inferiore | 37 | 41,6 | 223 | 28,6 | 235 | 14,3 | |
| | Licenza media superiore | 34 | 38,2 | 281 | 36,1 | 58 | 3,5 | |
| | Laurea | 16 | 18,0 | 62 | 8,0 | 5 | 0,3 | |

Stato civile e titolo di studio si distribuiscono in maniera coerente con le età (ad esempio, il livello dei titoli di studio decresce al crescere di essa). La differenza di genere mostra due andamenti fondamentalmente opposti, con gli uomini che presentano una percentuale più alta nella fascia di età interessata dallo screening.

La tabella 6.26 presenta, invece, il rapporto fra condizione professionale e titolo di studio.

Tab. 6.26 – Condizione professionale per titolo di studio.

| | | Licenza elementare o analfabeta | | Licenza media inferiore | | Licenza media superiore | | Laurea | | Sig. χ^2 |
|--------------------------|--|---------------------------------|------|-------------------------|------|-------------------------|------|--------|------|---------------|
| | | N | % | N | % | N | % | N | % | |
| Condizione professionale | Professione intellettuale, scientifica o di elevata specializzazione | 0 | 0,0 | 0 | 0,0 | 0 | 0,0 | 2 | 2,4 | p<0,05 |
| | Professione qualificata nelle attività commerciali o nei servizi | 2 | 0,1 | 2 | 0,4 | 19 | 5,1 | 48 | 57,8 | |
| | Impiegato | 74 | 4,7 | 15 | 3,0 | 98 | 26,3 | 16 | 19,3 | |
| | Artigiano, operaio specializzato o agricoltore | 139 | 8,9 | 202 | 40,8 | 199 | 53,4 | 0 | 0,0 | |
| | Pensionato | 1.198 | 76,6 | 275 | 55,6 | 57 | 15,3 | 0 | 0,0 | |
| | Disoccupato | 1 | 0,1 | 1 | 0,2 | 0 | 0,0 | 0 | 0,0 | |
| | Casalinga | 149 | 9,5 | 0 | 0,0 | 0 | 0,0 | 17 | 20,5 | |

La forte presenza di bassi titoli di studio condiziona la distribuzione delle condizioni professionali: chi è ancora in attività è, per lo più, artigiano, operaio specializzato o agricoltore e, in parte, impiegato, mentre le professione più qualificate, ovviamente,

sono proprie dei titoli di studio più elevati. Si nota anche una forte presenza di pensionati, per lo più con titoli di studio bassi o medio-bassi.

Le caratteristiche della malattia, della diagnosi e della cura, sono presentate nella tabella 6.27 tramite le variabili che, in letteratura, sono più frequentemente utilizzate e che sono state inserite nell'analisi di sopravvivenza. La stessa tabella mostra anche la loro distribuzione nei tre gruppi di età.

Tab. 6.27 – Caratteristiche cliniche dei pazienti, per età.

| | | Totale | | 0-49 | | 50-69 | | 70+ | | Sig. χ^2 |
|--|----------------------------------|--------|-------|------|-------|-------|------|-------|-------|---------------|
| | | N | % | N | % | N | % | N | % | |
| Stadio alla diagnosi | I | 446 | 17,7 | 19 | 21,3 | 178 | 22,8 | 249 | 15,1 | p<0,05 |
| | II | 447 | 17,8 | 9 | 10,1 | 153 | 19,6 | 285 | 17,3 | |
| | III | 637 | 25,3 | 22 | 24,7 | 173 | 22,2 | 442 | 26,9 | |
| | IV | 984 | 39,1 | 39 | 43,8 | 275 | 35,3 | 670 | 40,7 | |
| Grado alla diagnosi | Grado 1 | 416 | 16,5 | 15 | 16,9 | 124 | 15,9 | 277 | 16,8 | n.s |
| | Grado 2 | 1.152 | 45,8 | 35 | 39,3 | 396 | 50,8 | 721 | 43,8 | |
| | Grado 3 | 452 | 18,0 | 27 | 30,3 | 160 | 20,5 | 265 | 16,1 | |
| | Grado 4 | 494 | 19,6 | 12 | 13,5 | 99 | 12,7 | 383 | 23,3 | |
| Modalità diagnosi solo clinica | No | 2.101 | 99,8 | 79 | 100,0 | 696 | 99,6 | 1.326 | 99,8 | n.s |
| | Diagnosi: solo clinica | 5 | 0,2 | 0 | 0,0 | 3 | 0,4 | 2 | 0,2 | |
| Modalità diagnosi RX (clisma opaco) | No | 2.089 | 99,2 | 79 | 100,0 | 694 | 99,3 | 1.316 | 99,1 | n.s. |
| | Diagnosi: RX (clisma opaco) | 17 | 0,8 | 0 | 0,0 | 5 | 0,7 | 12 | 0,9 | |
| Modalità diagnosi ecografia transrettale | No | 2.098 | 99,6 | 79 | 100,0 | 694 | 99,3 | 1.325 | 99,8 | n.s |
| | Diagnosi: ecografia transrettale | 8 | 0,4 | 0 | 0,0 | 5 | 0,7 | 3 | 0,2 | |
| Modalità diagnosi endoscopia | No | 452 | 21,5 | 13 | 16,5 | 129 | 18,5 | 310 | 23,3 | p<0,05 |
| | Diagnosi: endoscopia | 1.654 | 78,5 | 66 | 83,5 | 570 | 81,5 | 1.018 | 76,7 | |
| Modalità diagnosi RMN | No | 2.097 | 99,6 | 79 | 100,0 | 696 | 99,6 | 1.322 | 99,5 | n.s. |
| | Diagnosi: RMN | 9 | 0,4 | 0 | 0,0 | 3 | 0,4 | 6 | 0,5 | |
| Modalità diagnosi Tc spirale | No | 2.041 | 96,9 | 75 | 94,9 | 676 | 96,7 | 1.290 | 97,1 | n.s |
| | Diagnosi: Tc spirale | 65 | 3,1 | 4 | 5,1 | 23 | 3,3 | 38 | 2,9 | |
| Modalità diagnosi SOF | No | 2.105 | 100,0 | 79 | 100,0 | 698 | 99,9 | 1.328 | 100,0 | n.s. |
| | Diagnosi: SOF | 1 | 0,0 | 0 | 0,0 | 1 | 0,1 | 0 | 0,0 | |
| Modalità diagnosi screening | No | 2.100 | 99,7 | 79 | 100,0 | 694 | 99,3 | 1.327 | 99,9 | p<0,05 |
| | Diagnosi: screening | 6 | 0,3 | 0 | 0,0 | 5 | 0,7 | 1 | 0,1 | |

Nota: n.s. = assenza di significatività statistica.

(Continua)

(Continua)

| | | Totale | | 0-49 | | 50-69 | | 70+ | | Sig. χ^2 |
|--------------------------------|------------------|--------|------|------|------|-------|------|-------|------|---------------|
| | | N | % | N | % | N | % | N | % | |
| Comorbidità alla diagnosi | No comorbidità | 2.051 | 81,6 | 78 | 87,6 | 640 | 82,2 | 1.333 | 81,0 | n.s. |
| | Comorbidità | 463 | 18,4 | 11 | 12,4 | 139 | 17,8 | 313 | 19,0 | |
| Chemioterapia | No chemioterapia | 1.890 | 75,2 | 42 | 47,2 | 458 | 58,8 | 1.390 | 84,4 | p<0,05 |
| | Chemioterapia | 624 | 24,8 | 47 | 52,8 | 321 | 41,2 | 256 | 15,6 | |
| Immunoterapia | No immunoterapia | 2.478 | 98,6 | 88 | 98,9 | 751 | 96,4 | 1.639 | 99,6 | p<0,05 |
| | Immunoterapia | 36 | 1,4 | 1 | 1,1 | 28 | 3,6 | 7 | 0,4 | |
| Radioterapia | No radioterapia | 2.093 | 83,3 | 59 | 66,3 | 583 | 74,8 | 1.451 | 88,2 | p<0,05 |
| | Radioterapia | 421 | 16,7 | 30 | 33,7 | 196 | 25,2 | 195 | 11,8 | |
| Chirurgia | No chirurgia | 764 | 30,4 | 13 | 14,6 | 132 | 16,9 | 619 | 37,6 | p<0,05 |
| | Chirurgia | 1.750 | 69,6 | 76 | 85,4 | 647 | 83,1 | 1.027 | 62,4 | |
| Pazienti trattati/non trattati | Non trattato | 681 | 27,1 | 11 | 12,4 | 106 | 13,6 | 564 | 34,3 | p<0,05 |
| | Trattato | 1.833 | 72,9 | 78 | 87,6 | 673 | 86,4 | 1.082 | 65,7 | |

Nota: n.s. = assenza di significatività statistica.

Anche per i tumori al retto, i pazienti considerati nello studio presentano uno stadio piuttosto avanzato della malattia (stadio III o IV, complessivamente 64% circa), ossia situazioni di malattia che hanno una prognosi sfavorevole o pessima. Mentre gli stadi III e IV non seguono un andamento regolare al crescere dell'età, gli stadi I e II presentano un picco rilevante di presenze nella fascia d'età centrale, ossia quella interessata dalle attività di screening organizzato.

Il grado di malignità, non presenta una relazione statisticamente significativa con l'età dei pazienti. Si nota una maggiore presenza di malattia con gradi a malignità medio-alta o alta (1 e 2, ossia neoplasie cellule ben differenziate), mentre sono meno numerosi i casi a malignità medio-alta o alta (3 e 4, ossia situazioni a scarsa differenziazione cellulare).

La modalità di diagnosi più utilizzata è, anche qui, di tipo endoscopico (colonscopia o sigmoidoscopia) da quasi il 79% dei pazienti, con un decremento rilevante al crescere delle età. Meno utilizzate le altre modalità di diagnosi.

Non vi è, invece, relazione fra comorbidità ed età: il 18% dei casi in esame mostra presenza di altre patologie al momento della diagnosi, con percentuali appena crescenti al crescere dell'età.

Per quanto riguarda le terapie, la chirurgia è, come da protocolli, la cura più utilizzata (in circa il 70% dei casi), ma si nota una riduzione della sua applicazione al crescere

dell'età. La chemioterapia è il secondo percorso di cura, spesso assieme alla chirurgia, ma la sua applicazione decresce fortemente al crescere dell'età. Anche la radioterapia è utilizzata, in particolare nel gruppo più giovane di pazienti, mentre l'immunoterapia è residuale nel percorso di cura.

Per questa neoplasia, infine, i pazienti non trattati ammontano al 27% circa, sono per lo più ultrasettantenni che combinano età e stadio avanzati (dato tabulato riportato in appendice). Questi sono ovviamente pazienti per cui la prognosi è tendenzialmente negativa (il 72% circa sopravvive meno di tre anni alla diagnosi), di cui sarà interessante valutare il ruolo e le caratteristiche nei modelli di sopravvivenza.

La tabella 6.28, infine, presenta i valori medi percentuali delle caratteristiche (utilizzate come covariate nel modello di sopravvivenza) delle sezioni di Censimento in cui sono residenti gli individui compresi nell'analisi, rispetto alla loro distribuzione regionale.

Rispetto alle altre due sedi di tumore studiate, gli ammalati di tumore al retto vivono in un insieme di sezioni di Censimento simile al resto della regione. Le differenze rilevanti riguardano il livello di istruzione, tendenzialmente più alto nelle sezioni dei casi considerati, e una maggiore presenza di alloggi di proprietà e in affitto, rispetto agli alloggi senza alcuna specifica.

Tab. 6.28 – Caratteristiche delle sezioni di Censimento dei pazienti a confronto con quelle regionali.

| | Casi in analisi | | | Regione |
|--|-------------------|--------------------------------|-------------------|------------------|
| | Valore medio aree | Intervallo di confidenza (95%) | | Valore regionale |
| | | Estremo inferiore | Estremo superiore | |
| Cittadinanza italiana (%) | 96,2 | 95,9 | 96,6 | 96,6 |
| Cittadinanza straniera (%) | 3,3 | 3,1 | 3,5 | 3,4 |
| Numero persone per famiglia | 2,6 | 2,6 | 2,6 | 2,6 |
| Dipendenti o in altra posizione subordinata (%) | 70,1 | 69,7 | 70,6 | 68,8 |
| Imprenditori (%) | 3,4 | 3,2 | 3,5 | 3,3 |
| Liberi professionisti (%) | 5,1 | 4,8 | 5,4 | 3,4 |
| Lavoratori in proprio (%) | 16,7 | 16,3 | 17,0 | 20,0 |
| Soci di cooperative (%) | 2,0 | 1,9 | 2,1 | 1,9 |
| Coadiuvanti familiari (%) | 1,6 | 1,5 | 1,7 | 2,0 |
| % Occupati | 42,7 | 42,4 | 43,1 | 41,5 |
| % In cerca di prima occupazione | 1,5 | 1,5 | 1,6 | 1,5 |
| % Disoccupati (in cerca di nuova occupazione) | 2,7 | 2,7 | 2,8 | 2,8 |
| % In attesa di iniziare un lavoro già trovato | 0,3 | 0,3 | 0,3 | 0,3 |
| % Studenti | 7,3 | 7,2 | 7,5 | 6,7 |
| % Casalinghe | 14,3 | 14,1 | 14,5 | 15,4 |
| % Ritirati dal lavoro | 25,7 | 25,4 | 26,0 | 26,4 |
| % In servizio di leva o in servizio civile sostitutivo | 0,2 | 0,2 | 0,3 | 0,2 |
| % Inabile al lavoro | 1,4 | 1,4 | 1,5 | 1,7 |
| % In altra condizione | 3,3 | 3,1 | 3,5 | 3,6 |
| Celibe / Nubile (%) | 35,2 | 35,0 | 35,5 | 35,1 |
| Coniugato/a (%) | 51,6 | 51,3 | 51,9 | 52,7 |
| Separato/a di fatto (%) | 0,4 | 0,3 | 0,4 | 0,3 |
| Separato/a legalmente (%) | 1,4 | 1,3 | 1,5 | 1,2 |
| Divorziato/a (%) | 1,2 | 1,1 | 1,2 | 0,8 |
| Vedovo/a (%) | 9,8 | 9,6 | 10,0 | 9,9 |
| Non sa leggere o scrivere (%) | 1,1 | 1,1 | 1,2 | 1,6 |
| Sa leggere e scrivere (%) | 7,6 | 7,4 | 7,8 | 9,8 |
| Licenza di scuola elementare (%) | 25,0 | 24,7 | 25,4 | 27,7 |
| Licenza di scuola media inferiore (%) | 25,9 | 25,6 | 26,2 | 27,7 |
| Diploma di scuola secondaria superiore (%) | 30,5 | 30,1 | 30,9 | 27,6 |
| Diploma non universitario post maturità (%) | 0,5 | 0,4 | 0,5 | 0,4 |
| Diploma universitario (%) | 0,9 | 0,8 | 0,9 | 0,6 |
| Laurea (%) | 8,1 | 7,8 | 8,4 | 4,7 |

(Continua)

(Continua)

| | Casi in analisi | | | Regione |
|---|--------------------------|--------------------------------|--------------------------|-------------------------|
| | Valore medio aree | Intervallo di confidenza (95%) | | Valore regionale |
| | | <i>Estremo inferiore</i> | <i>Estremo superiore</i> | |
| Indice di ricambio | 167,3 | 161,3 | 173,2 | 144,5 |
| Indice di vecchiaia | 239,4 | 230,6 | 248,2 | 220,8 |
| Indice di dipendenza strutturale | 57,4 | 55,6 | 59,1 | 59,2 |
| Tasso di attività | 64,8 | 64,4 | 65,1 | 64,2 |
| Tasso di occupazione | 58,5 | 58,1 | 58,9 | 57,8 |
| Tasso di disoccupazione | 9,1 | 8,9 | 9,3 | 9,4 |
| % alloggi di proprietà | 66,7 | 66,1 | 67,3 | 59,1 |
| % alloggi in affitto | 12,9 | 12,4 | 13,4 | 8,4 |
| % alloggi con altro titolo di godimento | 8,5 | 8,3 | 8,8 | 8,2 |
| % alloggi senza alcuna specifica | 11,2 | 10,7 | 11,7 | 24,2 |

6.3.2 I risultati del modello di sopravvivenza a parametri flessibili ed effetti misti

La tabella 6.29 riporta il risultato del confronto fra i diversi modelli di sopravvivenza implementati.

In essa sono evidenziati i gruppi di variabili utilizzati per ogni modello (a partire dal modello base che comprende età e stadio per quanto riguarda gli effetti fissi e la sezione di Censimento di residenza per quanto riguarda gli effetti casuali), il numero di casi utilizzati in ogni modello e il valore dell'AIC.

Tab. 6.29 – Modelli di sopravvivenza a parametri flessibili ed effetti misti implementati: caratteristiche e confronto.

| Età | Stadio | Variabili diagnostiche | Comorbidità | Variabili terapeutiche | Effetti fissi Caratteristiche socio-economiche individuali | Caratteristiche socio-economiche sezione di Censimento | Effetti casuali Sezione di Censimento di residenza | N | AIC |
|-----|--------|------------------------|-------------|------------------------|---|--|---|-------|---------|
| X | X | | | | | | X | 2.514 | 6.218,2 |
| X | X | | | | X | | X | 2.514 | 6.227,7 |
| X | X | | | | | X | X | 2.514 | 6.251,7 |
| X | X | X | | | | | X | 2.514 | 6.121,6 |
| X | X | X | X | | | | X | 2.514 | 6.122,1 |
| X | X | X | | | X | | X | 2.514 | 6.127,2 |
| X | X | X | | | | X | X | 2.514 | 6.143,7 |
| X | X | X | X | | X | | X | 2.514 | 6.167,9 |
| X | X | X | X | | | X | X | 2.514 | 6.151,0 |
| X | X | | | | X | X | X | 2.514 | 6.252,0 |
| X | X | X | | | X | X | X | 2.514 | 6.195,2 |
| X | X | X | X | | X | X | X | 2.514 | 6.153,1 |
| X | X | | | X | | | X | 2.514 | 6.207,3 |
| X | X | | | X | X | | X | 2.514 | 6.209,5 |
| X | X | | | X | | X | X | 2.514 | 6.239,1 |
| X | X | X | | X | | | X | 2.514 | 6.132,5 |
| X | X | X | X | X | | | X | 2.514 | 6.098,4 |
| X | X | X | | X | X | | X | 2.514 | 6.101,4 |
| X | X | X | | X | | X | X | 2.514 | 6.124,4 |
| X | X | X | X | X | X | | X | 2.514 | 6.103,4 |
| X | X | X | X | X | | X | X | 2.514 | 6.126,3 |
| X | X | | | X | X | X | X | 2.514 | 6.238,9 |
| X | X | X | | X | X | X | X | 2.514 | 6.101,3 |
| X | X | X | X | X | X | X | X | 2.514 | 6.091,1 |

Il modello che presenta il valore di AIC è più basso è quello che comprende tutte le aree in cui sono raggruppabili le covariate. Se si osserva l'andamento dei valori dell'AIC, si può già fare qualche considerazione iniziale su quali siano le aree che, a partire dal modello di base, incidono in maniera più rilevante sul calcolo della sopravvivenza. Si può, infatti, facilmente notare che le riduzioni più marcate si ottengono quando nel modello sono inserite le variabili relative alla diagnosi e alle caratteristiche socio-economiche individuali, mentre effetti di riduzione minori si registrano quando si inseriscono le comorbidità, le terapie e le caratteristiche dell'area di residenza.

Il modello di sopravvivenza illustrato nella tabella 6.30, che corrisponde a quello identificato dall'AIC come il più adeguato, riporta solo le covariate che hanno una relazione statisticamente significativa con il rischio di morte, escludendo tutte le variabili non statisticamente significative e quelle escluse dal test di collinearità. Similmente, si sono riportate solo le interazioni a due vie statisticamente significative e non escluse dal medesimo test.

Il modello presenta un valore di verosimiglianza pari a $-3.006,7$, statisticamente significativo, raggiunto utilizzando una quadratura di Gauss-Hermite a 16 nodi. Anche in questo caso il modello va quindi considerato attendibile.

Tab. 6.30 – Il modello di sopravvivenza relativa a parametri flessibili ed effetti misti.

| | | | Intervallo di confidenza (95%) | | | |
|----------------------|---|-----------------------|---|-----------------------------------|-----------------|-------|
| | | HR | Sig. | Est. Inf. | Est. Sup | |
| Effetti fissi | Stadio | Stadio I | 1 | <i>(categoria di riferimento)</i> | | |
| | | Stadio II | 3,97 | 0,00 | 1,83 | 8,60 |
| | | Stadio III | 5,45 | 0,00 | 2,57 | 11,57 |
| | | Stadio IV | 35,49 | 0,00 | 16,92 | 74,42 |
| | Diagnosi | Grado 1 | 1 | <i>(categoria di riferimento)</i> | | |
| | | Grado 2 | 1,78 | 0,00 | 1,30 | 2,46 |
| | | Grado 3 | 3,70 | 0,00 | 2,67 | 5,14 |
| | | Grado 4 | 3,35 | 0,00 | 2,41 | 4,67 |
| | Terapia | No chemioterapia | 1 | <i>(categoria di riferimento)</i> | | |
| | | Si chemioterapia | 0,65 | 0,00 | 0,54 | 0,80 |
| | | No terapia chirurgica | 1 | <i>(categoria di riferimento)</i> | | |
| | | Si terapia chirurgica | 0,60 | 0,00 | 0,33 | 0,91 |
| | SE individuali | Coniugata/convivente | 1 | <i>(categoria di riferimento)</i> | | |
| | | Celibe/Nubile | 1,56 | 0,00 | 1,13 | 2,15 |
| | | Separato/divorziato | 0,96 | 0,86 | 0,63 | 1,46 |
| | | Vedovo | 1,09 | 0,40 | 0,89 | 1,33 |
| | SE d'area | Indice di vecchiaia | 1,02 | 0,05 | 1,00 | 1,05 |
| | Effetti fissi di interazione a due vie | F # Stadio III | 1,85 | 0,02 | 1,53 | 2,12 |
| | | F # Stadio IV | 1,95 | 0,01 | 1,76 | 2,27 |
| | Effetti casuali | | | 0,01 | 1,00 | 0,00 |

Nota: HR = hazard ratio(eccesso di rischio di morte); SE = variabili socio-economiche.

Osservando gli effetti principali, si nota subito che lo stadio è sempre il principale predittore del rischio di morte, con eccessi di rischio fortemente crescenti quando la malattia è diagnosticata in stadio sempre più avanzato, confermando quindi quanto noto in letteratura e commentato in precedenza.

Per quanto riguarda le variabili legate alla diagnosi, solo il grado di malignità è mantenuto nel modello: gli eccessi di rischio di morte si incrementano al crescere della malignità della malattia.

Fra le terapie applicate, chirurgia e chemioterapia mostrano un effetto di forte riduzione degli eccessi di rischio (circa il 40% per la prima e circa il 35% per la seconda). Radioterapia e immunoterapia, invece, vengono escluse dal modello.

Fra le variabili socio-economiche individuali, solo lo stato civile presenta un effetto statisticamente significativo: i celibi/nubili fanno registrare un eccesso di rischio pari al

56%, i cui effetti sulla diagnosi e la cura (quindi effetti indiretti rispetto alla sopravvivenza) verranno analizzati nel paragrafo successivo.

Gli effetti socio-economici relativi alla sezione di Censimento di residenza mostrano un ruolo statisticamente significativo dell'indice di vecchiaia, anche se con un effetto debole, attorno al 2% di eccessi di rischio al crescere dell'indice.

L'indice di vecchiaia è l'unica covariata che, in questo modello, identifica un ruolo specifico dell'età sulla sopravvivenza.

Gli effetti delle interazioni a due vie, infatti, rimettono in gioco, almeno in parte, la differenza di genere che, combinando il genere femminile agli stadi avanzati 3 e 4 porta ad un marcato effetto di peggioramento delle probabilità di sopravvivenza. In realtà, si è già detto sopra come lo stadio avanzato sia di per sé il predittore principale di una bassa sopravvivenza, per cui la combinazione con il genere, in realtà, non aggiunge un'informazione particolarmente significativa al modello.

Infine, la sezione di Censimento di residenza, ossia la variabile di *clustering* responsabile per gli effetti casuali, non ha un ruolo statisticamente significativo nel modello.

6.3.3 L'analisi degli effetti dello stato civile sulla diagnosi e sulle cure

Nel tumore al retto, lo stato civile presenta un ruolo rilevante fra le covariate che producono gli effetti fissi.

In questo paragrafo esamineremo, quindi, il ruolo di tale caratteristica nell'influenzare la tempestività della diagnosi (ossia lo stadio) e le cure a cui i pazienti hanno avuto accesso.

La tabella 6.31 mostra come si distribuiscono stadio alla diagnosi e terapie rispetto allo stato civile, valutandone la connessione tramite significatività statistica del test del χ^2 ($p < 0,05$).

Tab. 6.31 – Stato civile, stadio alla diagnosi e terapie seguite.

| | | Stato civile | | | | | | | | Sig. χ^2 |
|--------------------------------|------------------|---------------|------|--------------------------|------|---------------------------|------|----------|------|---------------|
| | | Celibe/nubili | | Coniugato/a o convivente | | Separato/a o divorziato/a | | Vedovo/a | | |
| | | N | % | N | % | N | % | N | % | |
| Stadio alla diagnosi | I | 43 | 16.1 | 342 | 19.7 | 9 | 12.9 | 52 | 11.8 | p<0,05 |
| | II | 44 | 16.5 | 321 | 18.5 | 12 | 17.1 | 70 | 15.9 | |
| | III | 67 | 25.1 | 455 | 26.2 | 14 | 20.0 | 101 | 23.0 | |
| | IV | 113 | 42.3 | 619 | 35.6 | 35 | 50.0 | 217 | 49.3 | |
| Chemioterapia | No chemioterapia | 190 | 71.2 | 1.254 | 72.2 | 49 | 70.0 | 397 | 90.2 | p<0,05 |
| | Chemioterapia | 77 | 28.8 | 483 | 27.8 | 21 | 30.0 | 43 | 9.8 | |
| Immunoterapia | No immunoterapia | 265 | 99.3 | 1.707 | 98.3 | 67 | 95.7 | 439 | 99.8 | n.s. |
| | Immunoterapia | 2 | 0.7 | 30 | 1.7 | 3 | 4.3 | 1 | 0.2 | |
| Radioterapia | No radioterapia | 224 | 83.9 | 1.408 | 81.1 | 54 | 77.1 | 407 | 92.5 | p<0,05 |
| | Radioterapia | 43 | 16.1 | 329 | 18.9 | 16 | 22.9 | 33 | 7.5 | |
| Chirurgia | No chirurgia | 83 | 31.1 | 459 | 26.4 | 28 | 40.0 | 194 | 44.1 | p<0,05 |
| | Chirurgia | 184 | 68.9 | 1.278 | 73.6 | 42 | 60.0 | 246 | 55.9 | |
| Pazienti trattati/non trattati | Non trattato | 74 | 27,7 | 391 | 22,5 | 28 | 40,0 | 188 | 42,7 | p<0,05 |
| | Trattato | 193 | 72,3 | 1.346 | 77,5 | 42 | 60,0 | 252 | 57,3 | |

Lo stadio alla diagnosi vede le persone non in coppia con situazioni di malattia più problematiche: separati/divorziati, vedovi e celibi/nubili, in quest'ordine, hanno una maggiore frequenza di stadi avanzati (e conseguente prognosi negativa). Meno evidente la relazione fra gli altri stadi e la situazione familiare.

Rispetto alle terapie, si osserva come, tendenzialmente, i vedovi e, in parte, i celibi/nubili vedano una minore adesione ai protocolli di cura. Nella variabile dicotomica fra pazienti trattati e non trattati, è di nuovo il gruppo dei vedovi che, tendenzialmente, non ha ricevuto alcun trattamento terapeutico.

I risultati dei modelli di regressione logistica, presentati nelle tabelle 6.32 e 6.33, permettono di dare una lettura più approfondita del rapporto fra stato civile, diagnosi e terapie. In essi è stato anche utilizzato l'indice di vecchiaia, unica variabile socio-economica d'area che ha mostrato un effetto statisticamente significativo sulla sopravvivenza.

Tab. 6.32 – Stadio alla diagnosi e variabili socio-economiche: modelli di regressione logistica.

| Stadio alla diagnosi | | OR | Sig. |
|----------------------|----------------------|------|----------------------------|
| Stadio I | Coniugato/convivente | 1 | (categoria di riferimento) |
| | Celibe/nubile | 0,78 | n.s. |
| | Separato/divorziato | 0,60 | n.s. |
| | Vedovo | 0,55 | p<0,05 |
| | Indice di vecchiaia | 1,00 | n.s. |
| | Costante | 0,24 | p<0,05 |
| Stadio II | Coniugato/convivente | 1 | (categoria di riferimento) |
| | Celibe/nubile | 0,88 | n.s. |
| | Separato/divorziato | 0,92 | n.s. |
| | Vedovo | 0,84 | n.s. |
| | Indice di vecchiaia | 1,00 | n.s. |
| | Costante | 0,24 | p<0,05 |
| Stadio III | Coniugato/convivente | 1 | (categoria di riferimento) |
| | Celibe/nubile | 0,94 | n.s. |
| | Separato/divorziato | 0,70 | n.s. |
| | Vedovo | 0,84 | n.s. |
| | Indice di vecchiaia | 1,00 | n.s. |
| | Costante | 0,36 | p<0,05 |
| Stadio IV | Coniugato/convivente | 1 | (categoria di riferimento) |
| | Celibe/nubile | 1,32 | p<0,05 |
| | Separato/divorziato | 1,78 | p<0,05 |
| | Vedovo | 1,76 | p<0,05 |
| | Indice di vecchiaia | 1,00 | n.s. |
| | Costante | 0,53 | p<0,05 |

Nota: OR = Odds ratio.

Rispetto allo stadio alla diagnosi, si nota come solo lo stadio IV sia fortemente influenzato dallo stato civile, rispecchiando quanto visto nella tabella 25. Solo lo stadio I mostra un legame con l'essere vedovi, che sperimentano il 55% di probabilità in meno di essere così diagnosticati.

Tab. 6.33 – *Terapie e variabili socio-economiche: modelli di regressione logistica.*

| Terapie | | OR | Sig. |
|--------------------------------|----------------------|-----------|----------------------------|
| Chemioterapia | Coniugato/convivente | 1 | (categoria di riferimento) |
| | Celibe/nubile | 1,05 | n.s. |
| | Separato/divorziato | 1,11 | n.s. |
| | Vedovo | 0,28 | p<0,05 |
| | Indice di vecchiaia | 1,00 | n.s. |
| | Costante | 0,39 | p<0,05 |
| Immunoterapia | Coniugato/convivente | 1 | (categoria di riferimento) |
| | Celibe/nubile | 0,43 | n.s. |
| | Separato/divorziato | 2,46 | n.s. |
| | Vedovo | 0,13 | p<0,05 |
| | Indice di vecchiaia | 1,01 | p<0,05 |
| | Costante | 0,01 | p<0,05 |
| Radioterapia | Coniugato/convivente | 1 | (categoria di riferimento) |
| | Celibe/nubile | 0,83 | n.s. |
| | Separato/divorziato | 1,28 | n.s. |
| | Vedovo | 0,35 | p<0,05 |
| | Indice di vecchiaia | 1,00 | n.s. |
| | Costante | 0,26 | p<0,05 |
| Chirurgia | Coniugato/convivente | 1 | (categoria di riferimento) |
| | Celibe/nubile | 0,80 | p<0,05 |
| | Separato/divorziato | 0,54 | p<0,05 |
| | Vedovo | 0,46 | p<0,05 |
| | Indice di vecchiaia | 1,00 | n.s. |
| | Costante | 2,86 | p<0,05 |
| Pazienti trattati/non trattati | Coniugato/convivente | 1 | (categoria di riferimento) |
| | Celibe/nubile | 0,97 | n.s. |
| | Separato/divorziato | 0,82 | n.s. |
| | Vedovo | 1,10 | n.s. |
| | Indice di vecchiaia | 1,00 | n.s. |
| | Costante | 0,36 | p<0,05 |

Nota: OR = Odds ratio.

Per quanto riguarda le cure seguite⁴⁰, si osserva una generale riduzione delle probabilità di essere trattati chirurgicamente se si è celibi/nubili, separati/divorziati o vedovi.

⁴⁰ La radioterapia è stata esclusa dall'analisi, perché applicata ad un numero eccessivamente basso di pazienti, non essendo trattamento elettivo per il tumore al colon.

Quest'ultima condizione riduce la probabilità di ricevere anche le altre cure. Non vi è, invece, una relazione fra l'essere trattati o meno e lo stato civile.

L'indice di vecchiaia che caratterizza la sezione di Censimento di residenza fa registrare un'interazione con l'immunoterapia: al crescere di tale indice, cresce debolmente la probabilità di ricevere tale trattamento. Visto però il basso numero di casi che ricevono tale cura e l'effetto molto debole indicato dalla regressione, è da considerarsi dubbia la reale attendibilità di tale relazione.

6.4 Prime considerazioni sui risultati

I risultati sopra descritti mostrano aspetti interessanti, che verranno discussi e approfonditi nel capitolo seguente, ma che vengono qui sintetizzati per sottolineare anticipatamente alcuni aspetti particolarmente significativi del ruolo che le caratteristiche socio-economiche hanno mostrato nell'interagire in maniera statisticamente significativa con la sopravvivenza.

Il risultato principale è il ruolo rivestito dallo stato civile come *proxy* del supporto sociale e familiare nell'influenzare la sopravvivenza, in particolare condizionando la diagnosi precoce e l'accesso alle cure: persone *single* o vedove hanno peggiori probabilità di sopravvivenza, di avere una diagnosi più tempestiva (ossia un più basso stadio alla diagnosi), di accedere alle cure corrette.

Vi sono invece effetti ridotti del livello di istruzione, considerato in letteratura come uno degli indicatori più stabili ed affidabili per analizzare il rapporto fra le caratteristiche socio-economiche individuali e lo stato di salute di un individuo.

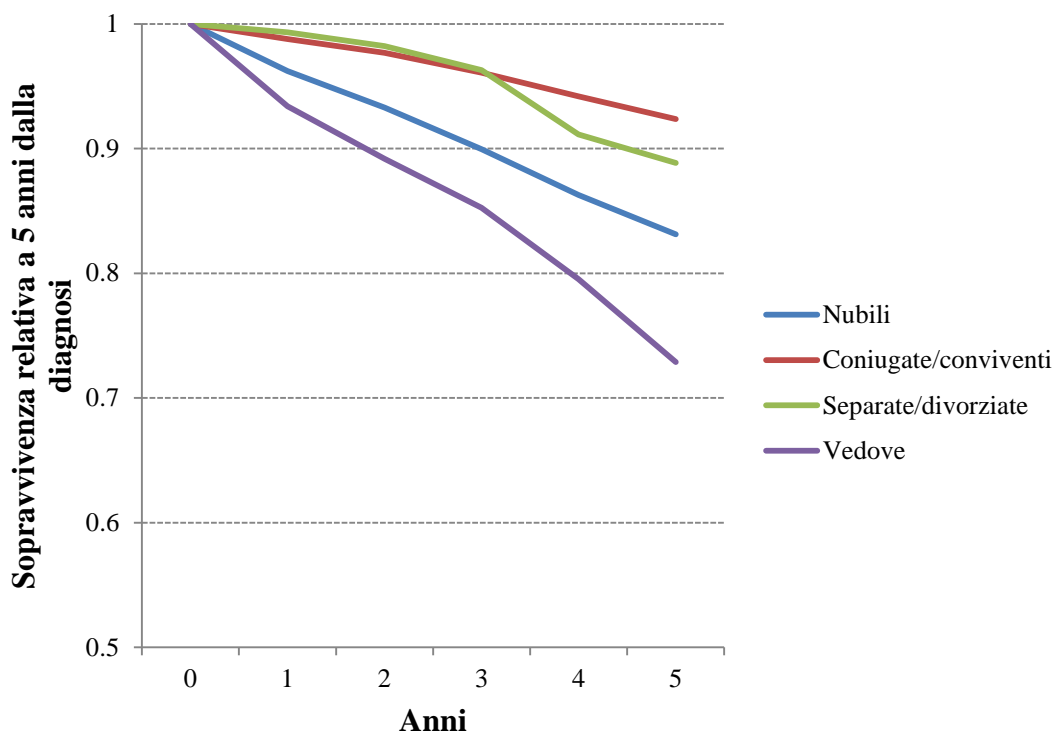
Questo risultato sembra sottolineare, di fatto, la buona qualità del sistema sanitario regionale di cui si sono analizzati i dati, perché sembra in grado di gestire le disuguaglianze di tipo economico e culturale che il titolo di studio incorpora, garantendo a quasi tutti i pazienti le stesse opportunità di diagnosi e cura.

Gli aspetti problematici sembrano, invece, emergere quando si analizza il ruolo del supporto familiare e sociale che, come vedremo meglio nel prossimo capitolo, riveste un significato e un ruolo importante sotto tanti aspetti (pratico, emotivo, motivazionale, ecc.) perché un paziente oncologico ponga attenzione ai segnali del proprio corpo, a campagne di prevenzione che permettano una diagnosi precoce, a seguire attentamente i complessi percorsi di cura quando la malattia si è ormai manifestata.

Se si osservano gli andamenti della sopravvivenza relativa a cinque anni dalla diagnosi, calcolati con i modelli sopra descritti, lo stato civile degli individui risulta essere discriminante in maniera significativa.

Il grafico 6.1, che riporta gli andamenti per il tumore alla mammella, evidenzia come nubili e vedove sperimentano costantemente una sopravvivenza significativamente inferiore alle coniugate e alle separate per tutto il quinquennio.

Graf. 6.1 – Sopravvivenza relativa al tumore alla mammella, a cinque anni dalla diagnosi per stato civile.



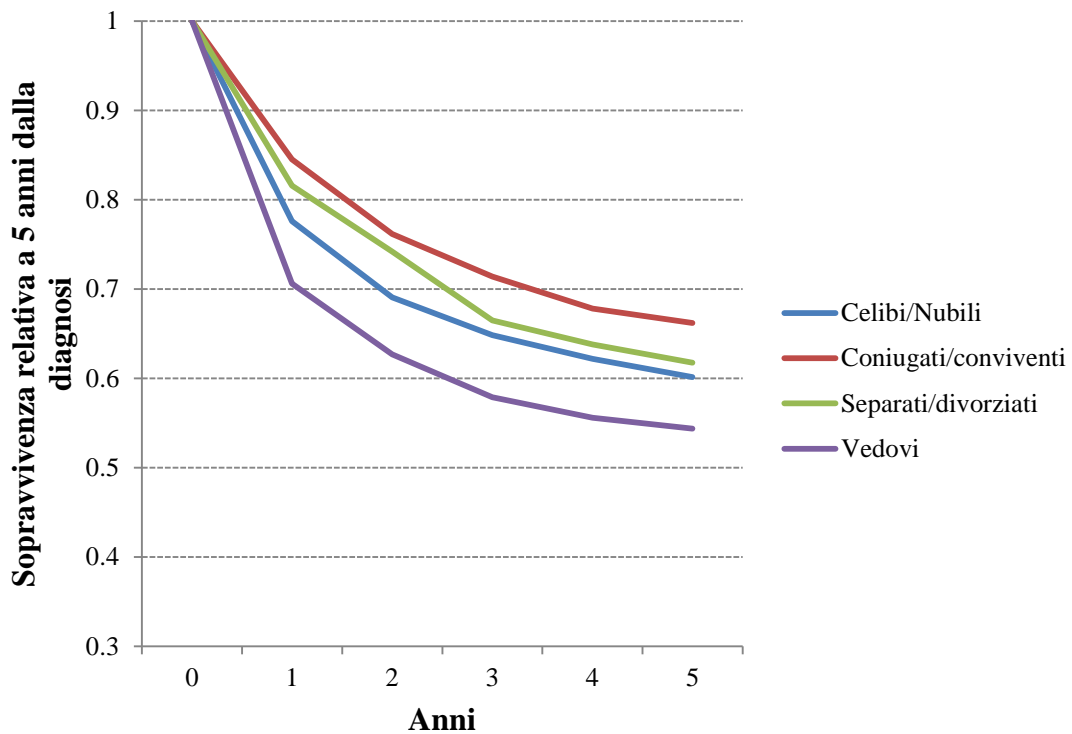
Le vedove mostrano addirittura una riduzione del 20% della sopravvivenza al quinto anno rispetto alle coniugate (73% vs. 93%), ma anche le nubili presentano una prognosi piuttosto sfavorevole con valori a cinque anni attorno all'83%.

Anche per gli individui con neoplasie al colon e al retto le curve di sopravvivenza mostrano importanti differenze legate allo stato civile.

Gli ammalati di tumore al colon presentano una sopravvivenza significativamente bassa più per i vedovi rispetto agli altri tre gruppi (Graf. 6.2): a cinque anni presentano una differenza con i coniugati di 12 punti percentuali (54% vs. 66%).

E' interessante l'andamento dei separati, che nei primi due anni si comportano in maniera simile ai coniugati e nei successivi tre anni come i *single*.

Graf. 6.2 – Sopravvivenza relativa al tumore al colon, a cinque anni dalla diagnosi per stato civile.

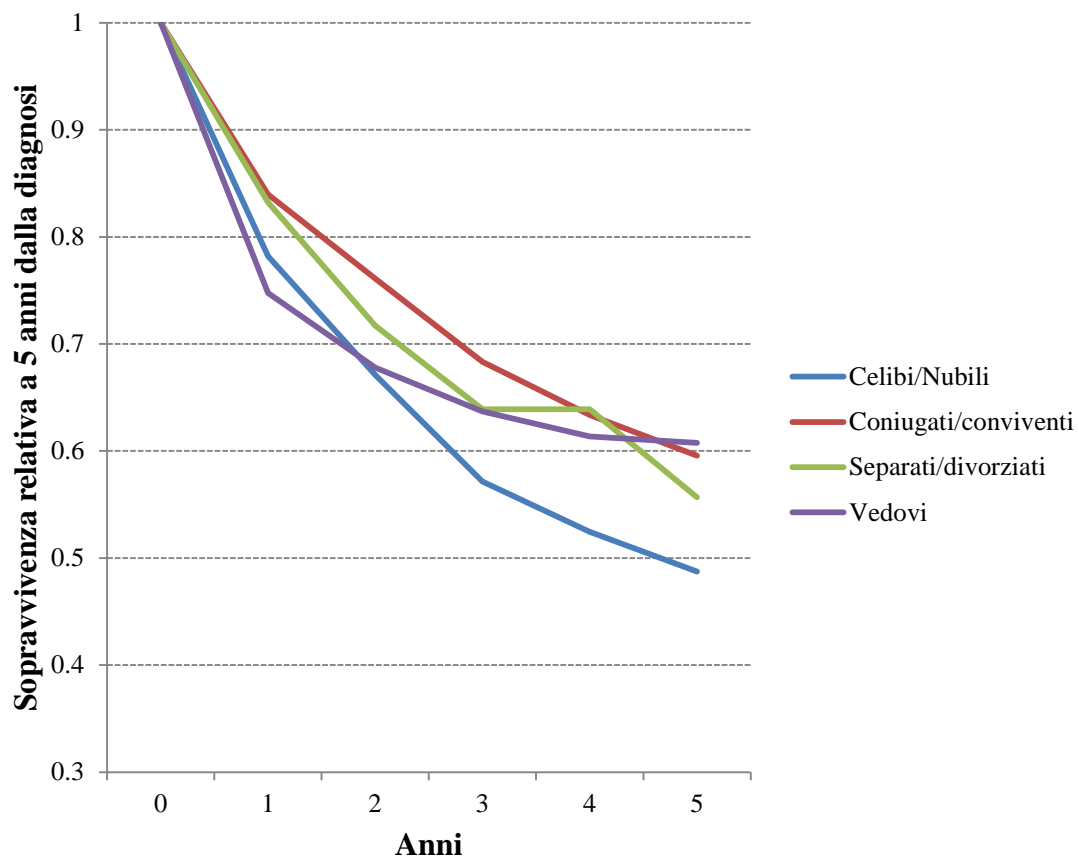


Anche per gli individui affetti dal tumore al retto la sopravvivenza si differenzia per stato civile (Graf. 6.3).

I *single* presentano probabilità di sopravvivenza significativamente più basse di tutti gli altri gruppi a partire dal secondo anno alla diagnosi, con una differenza fra essi e i coniugati alla fine del quinquennio di circa il 12% (49% vs. 61%).

E' interessante notare come negli altri tre gruppi i valori di sopravvivenza tendano ad uniformarsi a partire dal terzo anno, con i soli vedovi che, nei primi due anni, sperimentano una sopravvivenza peggiore dei celibi.

Graf. 6.3 – Sopravvivenza relativa al tumore al retto, a cinque anni dalla diagnosi per stato civile.



I risultati sopra descritti sembrano quindi indicare in generale un bisogno di intervenire sul problema del supporto a prescindere dall'età (celibi/nubili e vedovi hanno età differenti), anche se il ruolo dell'invecchiamento della popolazione, espresso dall'indice di dipendenza strutturale e dall'indice di vecchiaia delle sezioni di Censimento in cui risiedono i pazienti, indubbiamente sottolinea come il sistema sanitario locale dovrebbe considerare entrambe le problematiche in un'ottica di azioni per la riduzione delle disuguaglianze.

Capitolo 7

Discussione dei risultati e conclusioni

I risultati presentati nel capitolo precedente vengono discussi in questo capitolo secondo il seguente schema:

- dapprima verrà commentata la scelta del modello e il ruolo che le variabili socio-economiche individuali e di area hanno avuto nel definire tale selezione;
- successivamente saranno esaminate le caratteristiche dei modelli scelti per le differenti sedi di tumore, con particolare riferimento al ruolo delle variabili socio-economiche;
- infine, nelle conclusioni si discuterà se e come i risultati abbiano raggiunto gli obiettivi dello studio, quali siano i suoi limiti e quale contributo possa aver dato alla discussione scientifica in materia.

7.1 Considerazioni sulla scelta dei modelli di sopravvivenza

Come detto in precedenza, per individuare il modello che meglio descrive il ruolo delle covariate diagnostico-terapeutiche e socio-economiche nella sopravvivenza dei pazienti si è seguito un processo additivo dei gruppi di variabili in diversi passi (Wang et al., 2015).

In questo modo è stato possibile valutare 24 modelli differenti per ogni sede di tumore (che hanno coperto tutte le possibili permutazioni fra i diversi gruppi di variabili) tramite i mutamenti del valore dell'AIC (Akaike, 1974 e 1985; Sugiura, 1978).

Tali mutamenti hanno permesso di identificare quali gruppi di variabili sono risultati maggiormente rilevanti nel definire l'adattamento del modello dopo il loro inserimento (maggiore variazione equivale a maggiore rilevanza) e, di conseguenza, avere un maggiore impatto (diretto o indiretto) sulla sopravvivenza.

Come era attendibile, i gruppi di covariate che hanno un effetto più rilevante sono, tendenzialmente, quelle relative agli aspetti della gestione della malattia (diagnosi e terapie) con un effetto meno rilevante delle comorbidità. Ma per questo studio è particolarmente interessante esaminare cosa accade quando entrano in gioco le

caratteristiche socio-economiche individuali (stato civile, titolo di studio, condizione professionale).

Infatti, nel momento in cui tali covariate sono inserite in modelli in cui sono già presenti le informazioni diagnostico-terapeutiche, si assiste sempre ad una riduzione rilevante dell'AIC; tale riduzione non è invece altrettanto marcata (in alcuni modelli si assiste addirittura ad un incremento) quando sono assenti uno o più blocchi di covariate di tipo medico.

Questo sembra confermare come, in generale a livello di modello, le caratteristiche socio-economiche individuali abbiano un ruolo non trascurabile nel determinare la sopravvivenza al cancro, anche se devono essere considerate principalmente come portatrici di effetti indiretti che agiscono tramite un condizionamento dell'accesso alla diagnosi e alle opportunità di cura (come vedremo meglio nel prossimo paragrafo).

Tale risultato è in linea con quanto mostrato dai (non numerosi) studi che hanno affrontato tale problematica sempre a livello socio-economico individuale, come lo studio di Braaten et al. (2009), che ha verificato con un'indagine campionaria il ruolo di stato civile, titolo di studio, numerosità del nucleo familiare e abitudine al fumo sulle differenze nella sopravvivenza ad alcuni tumori, verificando come tali significatività permanessero anche dopo l'inserimento delle variabili diagnostiche e terapeutiche.

Diverso il comportamento osservato per le variabili che descrivono le caratteristiche socio-economiche dell'area di residenza. In generale esse non portano ad una riduzione particolarmente rilevante dell'AIC, evidenziando quindi un effetto debole sulla stima della sopravvivenza, ma tale effetto si registra anche in assenza di alcuni blocchi di covariate cliniche, a sottolineare una relazione comunque non trascurabile fra la situazione dell'area in cui vive il singolo individuo e la probabilità di sopravvivere ad un tumore.

Tale relazione è, come visto nel capitolo 2 di questa tesi, l'ipotesi su cui si basa una gran parte degli studi che cercano di valutare l'esistenza di causalità fra condizioni socio-economiche dell'area di residenza dei malati e probabilità di sopravvivenza.

Nello specifico dei tumori osservati in questo studio, possiamo ricordare ad esempio i risultati conseguiti dallo studio di Byers et al. (2008), che ha mostrato come negli Stati Uniti le caratteristiche socio-economiche dell'area di residenza incidessero sulla sopravvivenza di malati di tumore alla mammella, alla prostata e al colon-retto, con

sopravvivenze decrescenti al peggiorare delle caratteristiche dell'area. Tale lavoro ha anche verificato come questi effetti si indebolivano, inserendo nei modelli modalità di diagnosi e terapie applicate, pur rimanendo sempre con un effetto statisticamente significativo e non trascurabile.

Similarmente, quando nel presente studio si va ad osservare l'effetto combinato di tutte le variabili individuali, si nota subito come l'effetto sull'AIC delle variabili socio-economiche d'area risulti essere molto meno marcato, ma mantenga comunque una sua significatività statistica.

In più, come detto nel capitolo precedente, per ogni sede di tumore il modello che presenta l'AIC più basso (e, quindi, una migliore attendibilità per la stima della sopravvivenza) è sempre quello in cui tutti i gruppi di covariate sono presenti, ossia in cui tutti i gruppi hanno un effetto statisticamente significativo (più o meno marcato) sulla sopravvivenza e ciò permette le prime considerazioni rispetto alle ipotesi e agli obiettivi del presente studio.

In generale, infatti, i risultati relativi alla scelta dei modelli mostrano che, se si vogliono valutare correttamente le cause che stanno alla base della sopravvivenza ai tumori, non si possono escludere gli aspetti socio-economici che caratterizzano i pazienti e l'area in cui vivono, prendendo in particolare in considerazione il fatto che essi producono principalmente effetti indiretti, come emerge dal modo in cui viene modificato il valore dell'AIC.

In più, il livello socio-economico individuale e quello aggregato d'area non sono fra di loro sovrapponibili, ma sembrano esprimere aspetti differenti di questa relazione, visto che entrambi forniscono un contributo statisticamente significativo alla bontà dell'adattamento del modello.

Un'altra considerazione è particolarmente rilevante: tutto questo dimostra che le differenze in sopravvivenza legate alle caratteristiche socio-economiche permangono anche in una società il cui servizio sanitario pubblico offre azioni di prevenzione e diagnosi precoce, nonché regimi di trattamento potenzialmente standardizzati a tutta la popolazione, indipendentemente dal suo status socio-economico.

Questo aspetto, comune anche ad altri studi che hanno affrontato il problema in altre nazioni caratterizzate da un sistema sanitario pubblico universalistico (si veda, ad esempio, il lavoro di Lagerlund et al., 2005), sposta quindi l'attenzione sulle

caratteristiche specifiche e sul ruolo che esse hanno nel condizionare, per i diversi tumori, la possibilità di arrivare ad una diagnosi precoce e/o ad aderire correttamente ai percorsi di cura e di riabilitazione.

7.2 Considerazioni specifiche sui modelli di sopravvivenza relative alle singole sedi

Partendo da questa ultima considerazione, occorre quindi valutare cosa succede all'interno dei modelli di sopravvivenza scelti per i tumori studiati.

Si osservano, innanzitutto, alcuni elementi comuni:

- l'età alla diagnosi, uno dei fattori tradizionalmente più rilevanti della sopravvivenza, esce dagli effetti principali e resta solo in alcune interazioni a due vie;
- fra le variabili socio-economiche individuali solo lo stato civile mostra un effetto statisticamente significativo legato agli effetti principali; il titolo di studio entra in gioco solamente nelle interazioni a due vie con le variabili di tipo diagnostico nei pazienti affetti da tumore alla mammella e al colon, mentre la condizione professionale non presenta interazioni statisticamente significative con la sopravvivenza;
- in tutti i modelli selezionati emerge come statisticamente significativa una sola variabile relativa all'area di residenza degli individui. Per la sopravvivenza al tumore della mammella è l'indice di dipendenza strutturale, mentre per quanto riguarda colon e retto è l'indice di vecchiaia.

L'assenza dell'età alla diagnosi fra le covariate responsabili degli effetti principali è facilmente spiegabile con la presenza di altre variabili che sono fortemente legate ad essa ed esprimono, in parte, anche i suoi effetti. In particolare, è stata verificata in molti lavori la relazione fra età crescenti e diagnosi tardiva in tumori che fanno registrare una buona probabilità di sopravvivenza, come quelli qui considerati. Ma gli stessi studi hanno anche evidenziato come nei gruppi di età più giovane (comparabili al gruppo 0-49 anni utilizzato in questo lavoro), non sottoposti ad attività di screening organizzato, possano presentarsi percentuali di stadi avanzati paragonabili a quelli dei gruppi più anziani (anche se, ovviamente, con valori assoluti in termini di casi incidenti decisamente inferiori). In merito si può, ad esempio, citare la ricerca sviluppata da

Holmes et al. nel 1981, di riferimento per questo tipo di problematiche.

I dati qui analizzati hanno confermato questa relazione e di conseguenza lo stadio alla diagnosi ha sostituito l'effetto dell'età nei modelli di sopravvivenza come covariata diagnostica fondamentale, sia da un punto di vista puramente clinico, sia come *proxy* dell'età alla diagnosi.

Sulle ragioni che portano, nei tumori da screening, a questa sorta di parallelismo fra gruppi di pazienti più giovani e più anziani sono stati sviluppati diversi lavori che hanno portato a spiegazioni differenti, fortemente legate alle caratteristiche del sistema sanitario.

In regimi sanitari di tipo privatistico, come quello statunitense, il rapporto fra età e diagnosi tardiva è fondamentalmente legato ad avere la disponibilità di una buona assicurazione sanitaria (Walker et al., 2014) e alla possibilità di seguire uno stile di vita corretto, in particolare per quanto riguarda l'alimentazione. Un recente studio, sviluppato da Mobley et al. (2015), ha verificato questo tipo di associazione, sottolineando contemporaneamente il ruolo di supporto che, anche in un sistema sanitario come quello americano, possono avere le politiche sanitarie, implementate diversamente in ogni singolo stato federale.

In regimi sanitari di tipo pubblico, come quello italiano, le politiche sanitarie cercano (o dovrebbero cercare) di indirizzare con opportuni strumenti di prevenzione primaria e secondaria i comportamenti degli individui rispetto alla propria salute, fra cui una maggiore attenzione per i segnali della propria salute che possono portare ad una diagnosi precoce della malattia neoplastica. In realtà, i dati italiani mostrano come la diagnosi precoce è più frequente negli individui in età da screening organizzato (nel presente studio, il gruppo 55-69 anni), mentre solo una parte dei non numerosi giovani affetti da neoplasia arriva ad una diagnosi precoce (si veda, ad esempio, Paci et al., 2011 per il tumore alla mammella e Caldarella et al., 2014 per i tumori coloretali).

La spiegazione è duplice; da un punto di vista puramente clinico, in particolare per i tumori coloretali, spesso le neoplasie nei pazienti negli individui entro i 50 anni tendono ad essere più aggressive, con gradi di malignità elevati (dato in parte confermato da questo lavoro nel capitolo precedente); da un punto di vista comportamentale si è rilevato in più lavori come in molti casi vi sia una minore attenzione ad eventuali segnali deboli di malattia nella fascia di età più giovane, in

particolare in associazione con titoli di studio bassi o medio-bassi (come discusso, ad esempio, in ONDA, 2014).

Una seconda caratteristica comune ai modelli di sopravvivenza selezionati, di interesse fondamentale per questa tesi, è la presenza di un'unica variabile socio-economica individuale fra le covariate responsabili degli effetti principali: lo stato civile.

Questo risultato non è inconsueto negli studi che affrontano lo studio della sopravvivenza tramite caratteristiche a livello individuale. Come si è già discusso nel capitolo 2 di questa tesi, infatti, la situazione familiare è fortemente indicativa della possibilità di avere un supporto nella corretta adesione del paziente ai protocolli di cura, non solo da un punto di vista medico, ma anche psicologico. Studi anche molto recenti sviluppati all'estero, infatti, hanno mostrato come il vivere in famiglia o l'essere single, separati o vedovi abbia una forte rilevanza per spiegare le disuguaglianze in termini di sopravvivenza ai tumori come effetto indiretto (ad esempio, Levitz et al., 2015; Li et al., 2015) e tutti puntano sul ruolo del supporto che il vivere in famiglia può fornire, anche a livello motivazionale, per seguire pratiche di diagnosi precoce e attenzione alla propria salute.

L'impatto di una diagnosi di cancro sul paziente e sul suo nucleo familiare è, infatti, devastante, non solo da un punto di vista medico, ma anche da un punto di vista psico-sociale. Di conseguenza, per il paziente diventa fondamentale poter affrontare tutte queste problematiche con un supporto di tipo familiare che le condivide e che permetta di affrontare il complesso percorso di cura adeguato (Biondi et al., 1995).

Si è infatti già detto come le terapie siano lunghe e presentino forti elementi di complessità che impongono sovente la necessità di contare su di un aiuto concreto per poter giungere correttamente alla fine del percorso, malgrado in un sistema di sanità pubblica il supporto base per accedere alle cure è (o dovrebbe essere) garantito in maniera universalistica.

Un tumore che presenta buone opportunità di cura è infatti una malattia lunga da gestire, specialmente se si ha una situazione familiare in cui si è soli o con scarso supporto. Diagnosi e inizio delle cure rappresentano una fase la cui gestione da parte del malato è estremamente complessa, sia da un punto di vista emotivo sia da un punto di vista squisitamente pratico, vista la continuità prolungata di appuntamenti medici a cui, in

genere, il paziente deve sottoporsi e che, frequentemente, non è in grado di gestire da solo a causa delle conseguenze che i trattamenti stessi hanno. A questo si deve aggiungere tutta la fase di riabilitazione alla fine delle cure che, sovente, è altrettanto complessa da gestire della fase terapeutica e anche essa ha una pesante influenza sulla sopravvivenza (Baili et al., 2013).

In più, tutto il percorso della malattia implica un'interruzione del progetto di vita del malato e una completa ridefinizione dei propri obiettivi, del proprio ruolo e posizione sociale, obbligandolo a trovare nuove e non previste strategie di trasformazione della propria esistenza e conseguente gestione dello stress. Come si è già visto nel capitolo 2 di questa tesi, tali situazioni hanno potenzialmente un effetto negativo sulla sopravvivenza, quando vengono affrontate senza un adeguato supporto che è prima familiare che professionale (in merito, si veda anche Scabini et al., 2003 e 2006).

Il paziente oncologico, infatti, affronta fundamentalmente quattro problematiche che possono condizionare pesantemente il corso delle cure e il periodo di riabilitazione (con evidenti effetti negativi sulla sopravvivenza): il *distress* emozionale, il decadimento delle funzioni cognitive, la spossatezza legata all'esperienza della malattia, le difficoltà legate all'accettazione della nuova immagine di sé (fisica e di ruolo) (Annunziata et al., 2011). Queste problematiche hanno caratteristiche tali per cui sono proprio i soggetti che sono in una situazione di famiglia monocomponente ad avere una maggiore difficoltà a gestirle e, quindi, a vedere significativamente ridotta la probabilità di sopravvivenza.

Esaminando più dettagliatamente come queste quattro problematiche influenzano le probabilità di sopravvivenza rispetto al supporto che la famiglia può offrire, si osserva che per *distress* emozionale si intende l'esperienza multifattoriale di natura sociale e psicologica, emozionalmente spiacevole, che accompagna il paziente dal momento della diagnosi alla fine delle cure. Si tratta di una condizione definita da stati di tipo ansioso o depressivo, che possono incidere anche pesantemente in maniera negativa sull'adesione alle pratiche terapeutiche (Holland et al., 2010). Accanto alla necessità di un aiuto di tipo professionale, diversi studi hanno mostrato come, in realtà, sembra essere più efficace l'intervento da parte di appartenenti al nucleo familiare, in particolare il coniuge, per gestire adeguatamente tale situazione (Bultz et al., 2006).

Il decadimento delle funzioni cognitive temporaneo o a lungo termine è,

fondamentalmente, legato agli effetti collaterali delle terapie, in particolare di chemioterapie e radioterapie (si veda, ad esempio, Weiss, 2008). Tale decadimento, spesso amplificato dall'età, comporta una difficoltà a completare correttamente compiti e funzioni localizzati nel tempo e nello spazio, a mantenere concentrazione e coordinazione psico-motoria, ecc. La funzione di un coniuge o, comunque, di un familiare (o altra forma di supporto sociale) è quindi quella di garantire che, comunque, le cure vengano adeguatamente seguite pur in presenza di tale decadimento, migliorando di conseguenza le probabilità di sopravvivenza del malato.

La spossatezza legata all'esperienza della malattia è invece definita come una sensazione soggettiva, persistente e disturbante di stanchezza fisica, emozionale e cognitiva, legata alle cure e non connessa alle attività del vivere quotidiano. Essa agisce sulla capacità dell'individuo di svolgere le proprie attività personali, sociali e lavorative, danneggiando profondamente la qualità della vita (Dittner et al., 2004) e, in maniera indiretta, la sopravvivenza di lungo periodo (ossia superiore ai 5 anni) (Bower et al., 2006; Hofman et al., 2007). Diversi studi hanno però evidenziato come l'essere sposati (o vivere in coppia) tende a ridurre gli effetti della spossatezza da tumore, migliorando la sopravvivenza stessa (ad esempio, Wang et al., 2011).

La difficoltà di accettare le modificazioni fisiche e funzionali imposte dalla malattia e dalle cure coinvolge la percezione di sé stessi da un punto di vista fisico e, quindi, implica aspetti cognitivi ed emotivi che influenzano il modo di rapportarsi sia all'altro che alla malattia (Cash, 2004). La letteratura in materia si è focalizzata, in particolare, sulle donne affette da tumore alla mammella, in particolare su coloro che hanno subito un intervento totalmente o parzialmente non conservativo. Costoro mostrano, in genere, un basso livello di autostima in relazione alla propria immagine corporea (Helms et al., 2008; Ogden et al., 2008) e una minore efficacia nell'affrontare la malattia, con peggiori scelte terapeutiche e minore capacità di seguire adeguatamente le cure, con una riduzione della sopravvivenza (Helms et al., 2008). Anche rispetto a questa problematica, però, si è osservato come l'essere sposati o conviventi modifica positivamente tale atteggiamento e influisce, quindi, su come sono affrontate le cure.

Tutto ciò sembra quindi spiegare perché anche nei modelli di sopravvivenza sviluppati in questa tesi lo stato civile assume un'importanza così rilevante rispetto alle altre variabili socio-economiche: il bisogno di supporto sociale e familiare per affrontare

correttamente la malattia è tale che, quando è potenzialmente assente o insufficiente perché l'individuo è *single* o vedovo, esso condiziona in maniera significativa la sopravvivenza generando una situazione di disuguaglianza che deve essere attentamente presa in considerazione.

L'assenza del titolo di studio fra le variabili socio-economiche individuali portatrici degli effetti principali in entrambe le sedi è, invece, un risultato piuttosto inatteso perché, come ampiamente discusso nel capitolo 2, numerosi studi internazionali e nazionali hanno verificato il ruolo rilevante di tale variabile nel condizionare la sopravvivenza, sia in termini di diagnosi precoce, sia in termini di adesione corretta ai protocolli di cura. E questi risultati sono stati ottenuti in ricerche che hanno lavorato sulle informazioni sia a livello di area di residenza, sia a livello di informazioni individuali nei casi in cui è stato possibile (ad esempio, in Hussain et al., 2008).

Qui la situazione è decisamente diversa e sembra sottolineare una buona capacità del sistema sanitario pubblico regionale umbro a venire incontro a questo tipo di disuguaglianza.

I risultati in letteratura hanno infatti più volte evidenziato come il rapporto livello di istruzione-reddito condizioni la sopravvivenza, laddove il paziente ha bisogno di pagarsi la copertura sanitaria in termini assicurativi (si veda, fra i tanti, il lavoro recente di Hastert et al., 2016, sul contributo dalle caratteristiche socio-economiche d'area e la sopravvivenza ai tumori); ma anche in nazioni ove vi è completa copertura pubblica dei bisogni della popolazione si è più volte osservata tale associazione (si veda proprio il lavoro sopracitato di Hussain et al., 2008, che ha studiato l'influenza del titolo di studio sulla sopravvivenza a diversi tumori).

Anche in Italia è stato più volte rilevato come il livello di istruzione influisca sulla sopravvivenza (ad esempio, Caiazzo et al., 2004; Marinacci et al., 2013), ma alcuni studi più recenti hanno spostato l'attenzione su altri aspetti delle disuguaglianze. In particolare è interessante la ricerca di Pacelli et al. (2014) che, lavorando sulla popolazione dell'Emilia-Romagna e sul tumore della mammella, ha mostrato come l'effetto del titolo di studio (ottenuto attribuendo ad ogni individuo il titolo di studio più presente nella sezione di Censimento di residenza) sulla sopravvivenza si sia ridotto drasticamente nel gruppo di età inferiore a 50 anni, fino a scomparire completamente nel gruppo di età compresa fra i 50 e i 69 anni, a seguito dell'attivazione di un

programma di screening, completamente gratuito anche per le donne giovani in età non compresa nello screening organizzato.

La situazione umbra è, probabilmente, simile a quella emiliana, con un sistema sanitario ben distribuito sul territorio e che riesce ad agire in maniera adeguata sia in termini di prevenzione (i programmi di screening oncologico organizzato nella regione sono ormai attivi da molti anni), che di cura offerta ai malati.

In realtà, come vedremo nei paragrafi successivi dedicati ai modelli per sede di tumore, un effetto legato al titolo di studio si vede anche qua, almeno per quanto riguarda la sopravvivenza al tumore alla mammella e al colon, anche se tale effetto è mediato da altre variabili legate a come si presenta la neoplasia. Sarà quindi interessante discutere di come tale variabile socio-economica influenzi differenzialmente l'adesione alle terapie, come mostrato dai risultati presentati nel capitolo precedente.

Non è sorprendente, invece, che non si registrino effetti legati alla condizione professionale; questa è infatti fortemente connessa con il titolo di studio, come mostrato nel capitolo 6, e all'interno dei modelli essa viene esclusa dal calcolo per la forte collinearità con il livello di istruzione stesso e con l'età alla diagnosi.

Solo due variabili, fra quelle che descrivono le caratteristiche socio-economiche della sezione di Censimento di residenza, risultano statisticamente significative all'interno dei modelli: l'indice di dipendenza strutturale per il tumore alla mammella e l'indice di vecchiaia per i tumori coloretali. Si è già osservato nel capitolo precedente che l'effetto osservato è debole, anche se statisticamente significativo, ma l'aspetto particolarmente interessante è che entrambi gli indici sottolineano nuovamente il rapporto fra sopravvivenza e necessità di supporto sociale, reinserendo oltretutto l'elemento dell'invecchiamento della popolazione.

Valori di questi due indici particolarmente alti identificano, infatti, aree in cui la popolazione non è più attiva e, tendenzialmente, anziana, con uno sbilanciamento verso i bisogni che la popolazione attiva e giovane deve provvedere alla parte non attiva e anziana della popolazione (e viceversa, al loro calare). In termini di supporto sociale e sanitario, questo significa che chi vive in aree caratterizzate da alti valori di questi indici ha una maggiore probabilità di essere in condizioni di bisogno di supporto e, se non esso viene fornito a livello familiare e/o pubblico, il malato può vedere ridotta la propria

probabilità di sopravvivenza.

Considerando che, come visto nel capitolo 3, la popolazione umbra è tendenzialmente più anziana della media nazionale (con una ulteriore tendenza all'invecchiamento), sembra venire ulteriormente sottolineata la necessità di riflettere sulla progettazione di azioni che vadano nella direzione di fornire gli strumenti ad una popolazione che, sempre più spesso, avrà bisogno di supporto per accedere a procedure di diagnosi e cura adeguate, al fine di evitare che famiglie monocomponenti e anziane numericamente in crescita siano causa di situazioni di forte disuguaglianza socio-sanitaria

7.2.1 Il modello di sopravvivenza al tumore della mammella

Esaminando in dettaglio le caratteristiche del modello che descrive la relazione fra le covariate e la sopravvivenza al tumore della mammella, la variabile relativa allo stato civile risulta essere il principale indicatore socio-economico individuale. Più precisamente, abbiamo visto come l'essere nubili o vedove incrementa in maniera significativa i rischi di morte, sottolineando come l'essere tendenzialmente privi di un compagno (ossia di un supporto familiare) sia il maggiore determinante di disuguaglianza sanitaria per tale patologia, all'interno di un sistema pubblico che, come osservato nel paragrafo precedente, sembra in realtà essere in grado di contrastare altri determinanti di disuguaglianza socio-economica tradizionali (come il titolo di studio e la condizione professionale).

Questo risultato è in linea con quanto trovato da altri studi; ad esempio, un recente e importante lavoro di Aizer et al. (2013), ha verificato un simile ruolo del supporto sociale di tipo familiare nel garantire diagnosi e adesione adeguata alle terapie e, di conseguenza, una migliore sopravvivenza. Tale lavoro ha preso in considerazione più di un milione di pazienti statunitensi con tumore diagnosticato fra il 2004 e il 2008, considerando come nel nostro caso le principali caratteristiche socio-economiche individuali, accanto a tutti dati clinici dei pazienti⁴¹ (utilizzando però una diversa tecnica per l'analisi del rapporto fra covariate e sopravvivenza). Dopo l'inserimento nei modelli di sopravvivenza delle variabili di tipo sanitario, anche qui lo stato civile è

⁴¹ Come detto in precedenza, negli Stati Uniti il Surveillance, Epidemiology and End Results Program (SEER, l'equivalente dell'italiana AIRTUM) raccoglie i dati socio-economici dei pazienti a livello individuale, rendendo più semplice (anche se non frequente) l'esecuzione di studi come quello della presente tesi.

emerso come il fattore socio-economico statisticamente significativo e rilevante, in particolare proprio nel tumore femminile della mammella⁴²: nubili e vedove hanno più alta mortalità e peggiore tempo di sopravvivenza. Ad un'analisi più approfondita, sia quantitativa che qualitativa, è emerso come le pazienti non sposate sono risultate essere a rischio significativamente più alto di presentazione con cancro metastatico e di sottotrattamento, evidenziando l'impatto potenzialmente significativo che il sostegno sociale può avere sulla diagnosi precoce e sul trattamento.

Se si considerano i risultati relativi ai modelli di regressione logistica che valutano l'impatto dello stato civile sullo stadio alla diagnosi e sulle terapie seguite, presentati nel capitolo precedente, si trovano risultati molto simili.

Chi è single o vedova arriva più tardivamente alla diagnosi, probabilmente a causa di una minore attenzione ai messaggi legati alla prevenzione e ad una costruzione condivisa nella coppia della propria salute personale (Ren, 1997; Waxler-Morrison et al., 1991), ma anche per la preoccupazione di dover affrontare da soli (o, comunque, senza un compagno di vita al fianco) le conseguenze di una diagnosi di tumore (Canestrini, 2013).

Come già osservato, infatti, si tratta di un'esperienza totale che coinvolge la persona nella sua globalità: le donne con tumore alla mammella sperimentano una sofferenza connessa a tutte le dimensioni della vita, fisica, emotiva, psicologica, spirituale, sociale (Arman et al., 2002; Perreault, et al., 2005). La particolarità di tale tumore ci pone di fronte, oltre al disagio fisico, anche ad un trauma più profondo, legato ai grandi mutamenti che riguardano il corpo e l'identità femminile. Le donne con tumore al seno, infatti, devono fare i conti con l'alterazione del proprio corpo e dei significati ad esso connessi (Lindwall et al., 2009) e, per le donne nubili in particolare, con la paura di avere difficoltà future nel ricoprire i potenziali ruoli di madri e mogli.

I cambiamenti e le conseguenze che devono affrontare sono spesso radicali, in quanto anche nelle sopravvissute al cancro rimangono i segni e le ferite nel corpo di donna che complicano in maniera non trascurabile il percorso di sopravvivenza durante e dopo le cure, come hanno verificato diversi studi e articoli che si sono occupati della sopravvivenza tra le donne con tumore alla mammella e delle conseguenze della malattia anche nel lungo termine (si veda, a tal proposito, Foley et al. 2006; Kaiser,

⁴² E per i tumori del colon-retto, di cui parleremo nel prossimo paragrafo.

2008).

La prospettiva di affrontare una simile ordalia da sola potrebbe quindi essere un fattore importante nello scoraggiare l'adesione a pratiche di diagnosi precoce, temendo la possibilità di affrontare un percorso di cure così faticoso senza avere alle spalle un adeguato supporto familiare (Bloom, et al., 2001). In questo modo, la persona arriva alla diagnosi solo quando la malattia non è più negabile, ossia con stadio avanzato, eventuali metastasi e una prognosi decisamente negativa.

Le pazienti considerate in questi studio presentano, coerentemente a quanto appena detto, una maggiore probabilità di diagnosi in stadio III e IV proprio se nubili o vedove, a prescindere dall'età.

Similarmente, lo stato civile influenza la probabilità di ricevere cure adeguate: i risultati riportati nel capitolo precedente mostrano, infatti, come nubili e vedove abbiano minori probabilità di essere trattate con chirurgia e radioterapia, mentre per quanto riguarda la chemioterapia solo le vedove sembrano essere sotto-trattate (probabilmente perché prevalentemente anziane e, quindi, con maggiori problemi nel sopportare gli effetti collaterali, ma non solo).

Trattamenti inadeguati e necessità di supporto per seguire le cure sembrerebbero essere strettamente legati alla loro durata e agli effetti collaterali, che impongono alla paziente di essere adeguatamente seguita da persone che possano dare un tipo di accudimento che va oltre le mere necessità cliniche. E questo avviene in un'area geografica come quella studiata, dove non sembra esserci una problematica territoriale (la variabile che esprime gli effetti casuali legati alla località residenza non esprime effetti statisticamente significativi).

Si è discusso nel paragrafo precedente del costo pesante che un paziente affetto da tumore paga per portare a compimento il ciclo terapeutico, nonché di quanto sia difficile per una persona senza una relazione familiare stabile alle spalle mantenere la giusta costanza nelle pratiche da seguire e il tumore alla mammella non fa ovviamente eccezione. Uno studio inglese del 2001, fra gli altri, ha verificato (con una metodologia mista quali-quantitativa) come sia fortemente percepito il bisogno di aiuto continuo da parte della famiglia, e quindi come nubili e vedove facciano grande fatica a seguire tutte le pratiche terapeutiche se attorno non hanno una struttura familiare che si faccia in parte carico di aiutarle nei bisogni quotidiani di un periodo così difficile (Lindop et al.,

2001).

A fronte di queste problematiche legate al supporto sociale, abbiamo già sottolineato come sia interessante il ruolo marginale del titolo di studio come indicatore di disuguaglianza di sopravvivenza legata al capitale socio-economico, in particolare in questo tipo di tumore per il quale altri studi hanno mostrato relazione rispetto a diagnosi precoce e cure (ad esempio, Maturò 2009)⁴³.

Nel caso umbro, in particolare per questa neoplasia, si può ipotizzare che alla base della mancanza di tali differenze vi sia un buon lavoro sul territorio da parte di ASL e medici di famiglia per la diffusione delle azioni di screening mammografico organizzato o volontario, che riesce a contenere significativamente gli effetti delle disuguaglianze legate al capitale culturale ed economico⁴⁴. Il buon funzionamento del sistema regionale umbro si applica anche alla capacità di proporre le cure in maniera adeguata rispetto alla capacità delle pazienti di comprendere e accettare la malattia e come essa vada combattuta. Infatti, l'applicazione della regressione logistica al rapporto fra stadio alla diagnosi, terapie applicate e titolo di studio delle pazienti non mostra alcuna relazione statisticamente significativa (risultato riportato in Appendice).

L'unico effetto interessante attribuibile al titolo di studio è legato all'interazione fra licenza elementare e grado di malignità: in questo caso si osserva come l'essere in possesso di licenza elementare costituisca in effetti uno svantaggio significativo, perché rende simili le probabilità di sopravvivenza di chi ha un tumore con grado di malignità 2, ossia con una prognosi piuttosto favorevole al malato, a quelle di chi ha un tumore con grado di malignità 3, ossia con una prognosi decisamente meno favorevole.

Non essendo mai stata verificata alcuna relazione diretta fra grado di malignità e livello di istruzione, l'interazione di quest'ultimo con il grado di malignità evidenzia come, probabilmente, l'effetto trainante sulla sopravvivenza è del titolo di studio che, in questo caso specifico, agisce effettivamente come indicatore di disuguaglianza. Va anche osservato che il livello di studio elementare è, nelle pazienti considerate, prevalentemente associato all'età anziana (70 e più anni) la quale, infatti, agisce in

⁴³ Nel rapporto PASSI 2015 si può notare come il livello di istruzione incida notevolmente su tale attività preventiva: le donne meno istruite usufruiscono prevalentemente dello screening organizzato mentre, le donne più colte si sottopongono più spesso agli esami come scelta.

⁴⁴ Si ricorda che la condizione professionale è fortemente correlata al titolo di studio

interazione nel modello di nuovo con il grado di malignità 3⁴⁵.

Si può quindi concludere che, malgrado l'efficacia del sistema sanitario regionale, resta comunque un gruppo ristretto di pazienti anziane e con basso titolo di studio che fa registrare uno svantaggio di sopravvivenza legato al capitale culturale. Per esse si dovrebbero valutare azioni specifiche di intervento sanitario.

L'indice di dipendenza strutturale è, infine, l'unica covariata socio-economica relativa alla sezione di Censimento che interagisce in maniera statisticamente significativa con la sopravvivenza. Essa rinforza le considerazioni sulla rilevanza del supporto sociale necessario per aspirare ad avere migliori probabilità di superare la malattia.

Tale indice esprime, infatti, il rapporto tra la parte di popolazione non attiva (sia giovane che anziana) e la parte di popolazione attiva: di fatto esprime il bisogno di supporto economico (e non solo) presente in ogni sezione di Censimento. Tendenzialmente, considerando l'andamento demografico regionale, individua al suo crescere la parte di popolazione anziana che ha bisogno di essere sostenuta dalla parte attiva a livello lavorativo della popolazione.

Nel presente lavoro, questo risultato rimette in gioco il ruolo dell'età, che le caratteristiche socio-economiche individuali avevano drasticamente ridotto. Infatti, aree caratterizzate da un alto valore dell'indice presenteranno un maggior numero di anziani e, di conseguenza, una maggior probabilità di donne affette da tale patologia. In più, zone caratterizzate da un alto valore dell'indice avranno anche meno risorse economiche e umane da mettere in gioco direttamente per dare aiuto a chi incorre in tale patologia e faranno registrare bisogni di supporto più elevati.

Se in uno studio come il presente, che associa caratteristiche socio-economiche individuali e aggregate, è lecito attendersi da un punto di vista statistico un contributo non particolarmente elevato di questa variabile al modello di sopravvivenza (in merito si veda, ad esempio, Biolcati Rinaldi et al., 2013; Lucchini et al., 2009; Marinacci et al., 2010), da un punto di vista interpretativo invece può essere considerato un indicatore socio-economico molto interessante per individuare aree dove esiste una disuguaglianza in termini di esiti di salute legata a differenze di risorse di supporto che chi è malato può mettere in gioco e per le quali politiche di intervento opportunamente mirate potrebbero

⁴⁵ Nonché con lo stadio IV, confermando anche a livello di modello l'associazione fra stadio avanzato ed età anziana.

essere applicate per ridurre lo svantaggio.

7.2.2 Il modello di sopravvivenza al tumore del colon-retto

Abbiamo visto come i pazienti affetti da tumori del colon-retto siano stati considerati separatamente, prendendo in considerazione le due sedi come sono individuate dalla Classificazione Internazionale delle Malattie (ICD-10; Ministero della Sanità, 2001), per tenere conto delle differenze di tipo clinico e biologico che contraddistinguono le due neoplasie (ad esempio, Hong et al., 2012).

Se si osservano però i risultati dei due modelli di sopravvivenza, si osserva la presenza di forti similitudini, almeno per quanto riguarda gli effetti principali. Entrambe le sedi, infatti, presentano come covariata socio-economica rilevante lo stato civile e come covariata d'area l'indice di vecchiaia, che recupera il ruolo dell'età in questo tumore. Non emergono, invece, differenze di genere statisticamente significative.

Lo stato civile è però declinato diversamente: per il colon è rilevante l'essere vedovi, mentre per il retto l'essere single.

Viene però ancora una volta confermata come l'essere in una condizione familiare senza un appoggio diretto da parte del coniuge sia un elemento rilevante per avere peggiori probabilità di sopravvivenza, il che vuol dire arrivare ad una diagnosi più tardiva e ad una adesione meno accurata ai protocolli di cura.

Tutte le considerazioni fin qui fatte sia in generale, sia per il tumore alla mammella sono quindi valide anche per i pazienti di cancro al colon-retto, con alcuni elementi che rafforzano la necessità di supporto: sia le pratiche di diagnosi (che coinvolgono la colonscopia) che di cura per queste neoplasie sono particolarmente invasive (si pensi all'uso di chirurgia demolitiva che può portare alla stomia temporanea o definitiva, accompagnata da chemioterapia e/o radioterapia in cicli piuttosto lunghi).

Quindi, poiché il sottoporsi alle terapie antitumorali risulta particolarmente faticoso, la rete sociale e familiare può aiutare nella gestione del dolore e delle problematiche individuali che tale malattia genera, nell'alleviare la fatica delle cure, nel comprendere l'utilità e motivare alla partecipazione delle scelte di trattamento (come già verificato da Johansen et al. in uno studio del 1996 su circa 10000 pazienti danesi).

Anche nel presente lavoro, i risultati hanno mostrato l'effetto dell'essere single o vedovi: nei pazienti affetti da tumore al colon i vedovi hanno basse probabilità di avere

una diagnosi precoce, mentre hanno altre probabilità di essere diagnosticati ad uno stadio avanzato, metastatico; similamente, per la sede del retto, i risultati mostrano come non solo i vedovi ma anche i single sperimentino la medesima diagnosi tardiva.

Diverse le ragioni: le neoplasie del colon e del retto, a causa della mancanza di sintomi rilevanti in fase iniziale, sono malattie la cui diagnosi è spesso tardiva e tende a colpire persone anziane che possono trascurare gli eventuali segnali deboli, specie se non hanno a fianco il coniuge con cui scambiare informazioni e opinioni su quanto sta avvenendo, vista anche la difficoltà che si può incontrare nel parlare di certi argomenti (come l'aver notato presenza di sangue nelle feci) con persone altre dal coniuge stesso. Altri lavori in materia, infatti, hanno mostrato come di questi argomenti le persone, in particolare gli anziani, abbiano difficoltà a parlarne con altri che non siano la persona con cui condividono alla pari la vita quotidiana (si veda, ad esempio, Maturo, 2009). Celibi/nubili e vedovi, quindi, a prescindere dall'età, tendono quindi a sottovalutare tali segnali, non avendo una figura di riferimento con cui condividere in maniera confidenziale tale situazione (se viene notata).

È possibile, dunque, che l'aver una rete sociale più densa o allargata, ma comunque focalizzata sul rapporto con il coniuge, favorisca il passaggio di informazioni sanitarie corrette o stili di vita sani e l'attitudine a sottoporsi a esami di screening. Nel caso specifico va però anche considerato un possibile *bias* causato dall'età, ossia al crescere dell'età si riduce il numero dei legami sociali posseduti. Ma anche la presenza di tale distorsione non cancellerebbe comunque il bisogno di supporto per l'accesso a programmi di diagnosi precoce organizzati (peraltro presenti in Umbria anche per le neoplasie del colon-retto) o volontari.

L'effetto dello stato civile si manifesta anche nelle terapie applicate: *single* e vedovi (e anche separati/divorziati) sperimentano una probabilità di accesso alle cure significativamente più bassa dei coniugati, sia per il colon che per il retto, verificando quanto emerge dalla letteratura internazionale in tema (ad esempio, Ayanian et al., 2003; Langenbach et al., 2003).

Torna quindi rilevante il problema di come persone con scarso supporto familiare abbiano meno possibilità di ricevere cure adeguate, quando queste sono particolarmente debilitanti e problematiche sotto tanti punti di vista se affrontate da soli, portando ad una rinuncia più o meno consapevole della cura adeguata (Biondi et al., 1995;

Canestrini, 2013).

Per il solo tumore al colon è interessante discutere il ruolo del titolo di studio, che entra in gioco quando è basso (licenza elementare o media inferiore) negli effetti di interazione con lo stadio II, ossia uno stadio di diagnosi che permette di formulare in genere una prognosi ragionevolmente favorevole. Quando però questo stadio si combina con il livello di istruzione sopra citato, le probabilità di sopravvivere per il paziente si abbassano drasticamente.

Come mostrato nel capitolo precedente, questo effetto del titolo di studio sulla sopravvivenza è spiegabile con le cure a cui accedono i pazienti: coloro che hanno tale livello di istruzione hanno probabilità decisamente inferiore di ricevere un qualsiasi tipo di terapia e, in parte, vedono aumentare significativamente la probabilità di non essere trattati affatto.

Per i tumori al colon, questo risultato conferma quanto la letteratura in materia ha trovato in termini di disuguaglianze nella sopravvivenza legate a differenze nel capitale socio-economico, inteso principalmente come capitale culturale (Ross et al., 2010).

In particolare, si può ricordare un lavoro svedese del 2011 (Cavalli-Björkman et al., 2011) che, su circa quattromila malati di tumore al colon cui è stata diagnosticata la malattia nel periodo 1995-2006, ha verificato come la presa in carico del paziente e, in particolare, le cure somministrate fossero condizionate dal loro titolo di studio (oltreché dallo stato civile), sottolineando come pazienti con un più basso livello di istruzione fossero sotto-trattati, anche in un sistema sanitario pubblico universalistico come quello svedese. In più, tale studio ha anche permesso di evidenziare come il livello di istruzione riducesse gli effetti dell'età, agendo in maniera non collineare ad essa, ossia contribuendo a migliorare i trattamenti somministrati in caso di titoli di studio elevati e a ridurli in caso contrario. Ovviamente, la sopravvivenza si comportava di conseguenza. Le ragioni di questo sotto-trattamento sono state esplorate ampiamente (in Italia si possono, ad esempio, ricordare i lavori di Ardissonne et al., 2013; Cipolla et al., 2008; Iseppato, 2008; Maturo et al., 2006) e sono fondamentalmente riconducibili alla gestione del rapporto non facile fra il paziente e il mondo medico, che molte volte non riesce a trasmettere in maniera sufficientemente comprensibile il messaggio relativo alle cure, ingenerando errori nella scelta del percorso da seguire, nella determinazione a volerlo seguire, ecc.

Esiste anche il problema logistico, di cui in parte si è già parlato in precedenza, e che può esprimersi, per persone meno preparate, in una difficoltà ad organizzarsi adeguatamente in presenza di cicli terapeutici lunghi e faticosi. In più, i risultati hanno mostrato come a titoli di studio bassi si associno condizioni professionali dal reddito non particolarmente elevato, aspetto che sicuramente può limitare il ricorso ad aiuti a pagamento per il supporto alle cure o per fare scelte di mobilità sanitaria per la terapia (fenomeno comunque non particolarmente rilevante in Umbria).

Infine non va comunque trascurata la combinazione fra titoli di studio non elevati ed età più elevate, che sicuramente complica ulteriormente l'adesione ai protocolli di cura da parte di tali pazienti.

Per il tumore al colon, quindi, siamo in presenza di una situazione in cui un sistema sanitario regionale particolarmente efficiente non riesce comunque a intervenire in maniera completamente adeguata per limitare gli effetti di questo tipo di disuguaglianza socio-economica, peraltro limitata ad un numero ridotto di individui, rendendo necessaria una riflessione in termini di azioni da mettere in atto per intervenire su questo gruppo di popolazione.

La presenza dell'indice di vecchiaia come covariata d'area conferma a livello ecologico quanto visto sopra, focalizzando però l'attenzione sui bisogni specifici di supporto di una popolazione anziana.

Si è già detto come i tumori coloretali siano particolarmente frequenti in popolazioni anziane che, come discusso in precedenza, vedono amplificate tutte le problematiche legate alle proprie caratteristiche socio-economiche.

Un'area caratterizzata da un indice di vecchiaia crescente vuole quindi dire un'area in cui aumentano le problematiche che il sistema sanitario deve affrontare, perché aumentano gli effetti negativi sulla sopravvivenza di tali caratteristiche individuali.

Perciò, come accadeva per l'indice di dipendenza strutturale connesso alla sopravvivenza del tumore alla mammella, anche questo indice può rivelarsi un ottimo indicatore per individuare aree verso cui indirizzare opportune azioni di intervento per ridurre gli squilibri nella popolazione malata.

7.3 Conclusioni

In un editoriale del 2013, Kissane, ribadisce la fondamentale importanza di una relazione stabile nel condividere la fatica della cura e migliorare la capacità di gestire la malattia nel suo insieme complesso. In questo, sottolinea la necessità di riconoscere lo stato di single o vedovo del paziente come segno di supporto sociale sfavorevole.

Questa tesi è arrivata fondamentalmente alla medesima conclusione: in un sistema sanitario pubblico di tipo universalistico efficiente, lo stato civile dei pazienti è il principale indicatore socio-economico che sottolinea una disuguaglianza nella sopravvivenza a tumori per cui sono disponibili percorsi di diagnosi e cura efficaci.

Lo stato civile incorpora in sé tutta una serie di aspetti legati alla necessità di supporto sociale e familiare necessario per:

- gestire gli aspetti pratici della diagnosi e della presa in carico del malato per assicurare una corretta adesione ai protocolli di cura;
- assicurare il supporto per una corretta comunicazione con le strutture sanitarie, in un momento di particolare stress per l'individuo, che si trova a combattere con una patologia particolarmente devastante per il proprio percorso di vita;
- aiutare ad affrontare e sopportare fatica, stress, dolore, demotivazione e potenziale depressione, conseguenti alla malattia e alle cure e che possono compromettere la corretta adesione alle stesse;
- favorire pratiche di diagnosi precoce, come la partecipazione ai programmi di screening.

Tali effetti devono essere considerati con particolare attenzione quando si è in presenza di una popolazione tendenzialmente anziana come quella umbra, come rilevato dall'indice di dipendenza strutturale e dall'indice di vecchiaia che sottolineano dove può essere maggiormente necessario il bisogno di un supporto che le caratteristiche familiari e demografiche potrebbero non garantire.

Il primo obiettivo del presente lavoro sembra quindi essere stato raggiunto: esiste un effetto sulla sopravvivenza delle caratteristiche socio-economiche individuali e relative all'area in cui vive il malato, il cui ruolo non viene annullato dal considerare tutte le possibili variabili diagnostiche terapeutiche.

Un aspetto specifico è anche più interessante; questo studio sottolinea come il sistema

sanitario pubblico considerato funzioni particolarmente bene, tanto da ridimensionare il ruolo delle caratteristiche socio-economiche che, tradizionalmente, si associano alle disuguaglianze di salute, ossia il titolo di studio e la condizione professionale, espressione di disuguaglianze nel capitale culturale ed economico che l'individuo può mettere in gioco per contrastare la malattia.

E' infatti missione di un simile sistema fare in modo che tutti possano avere le medesime opportunità di accedere a procedure di diagnosi e presa in carico della malattia il più corrette e adeguate possibili, a prescindere dalla propria condizione socio-economica.

Viene però sottolineata anche un'altra problematica che, come anticipato nel primo capitolo di questa tesi, potrebbe essere destinata a diventare una delle principali sfide per i paesi caratterizzati da un sistema di questo tipo, ossia la necessità di provvedere alla riduzione delle reti familiari di supporto in conseguenza delle trasformazioni dell'istituzione famiglia e del contemporaneo invecchiamento della popolazione.

Avere individuato quale sia ruolo dell'essere o meno parte di una convivenza familiare fornisce quindi un'indicazione interessante sulle priorità che, almeno per quanto riguarda le due patologie esaminate in questo studio, potrebbero indirizzare le scelte di intervento sociale e sanitario da parte delle istituzioni locali, al fine di migliorare le probabilità di sopravvivenza dei pazienti svantaggiati in termini di supporto socio-familiare.

Da questo punto di vista, tutto quanto detto finora ha mostrato come lavorare su queste tematiche a livello d'area e a livello individuale porti a risultati che si integrano, ma ha anche sottolineato quanto l'informazione individuale sia più rilevante rispetto a quella aggregata, anche quando si lavora su sezioni di Censimento ridotte come quelle italiane. Il lavorare solo a livello ecologico, quindi, presenta il problema di arrivare ad un'informazione che, pur descrivendo una parte delle problematiche che si troverebbero anche a livello individuale, è decisamente meno approfondita e non riesce a descrivere altrettanto bene le caratteristiche socio-economiche specifiche che individuano in maniera dettagliata gruppi di individui con peggiori performance di salute (statisticamente questo è stato verificato nel corso del processo di selezione del modello di sopravvivenza migliore).

Per le istituzioni che si occupano di raccogliere le informazioni con finalità di studio e intervento nell'area socio-economica e della salute, questi risultati danno una forte indicazione sull'importanza di espandere i propri database e la propria raccolta dati, inserendo variabili a livello individuale che, almeno in Italia, non sono abitualmente raccolte in maniera routinaria e dettagliata da enti che si occupano di salute, come i Registri Tumori.

Proprio i Registri Tumori, come in parte dimostrato da questo studio e in parte dai lavori esteri che, si è visto, fanno già uso di tale livello di dettaglio, troverebbero particolare utilità in tale procedura. Potrebbero infatti fornire ai decisori pubblici informazioni molto più dettagliate per stratificare più accuratamente la popolazione malata a seconda delle sue caratteristiche socio-economiche e degli svantaggi osservati o potenziali in termini di condizioni di salute e aiutando, così, a meglio orientare le scelte di intervento.

Un'alternativa a tale soluzione potrebbe essere il tipo di procedura di *data linkage* che è stata utilizzata per costruire i *dataset* utilizzati in questa tesi. Si è infatti riusciti a mettere assieme fonti informative che già raccolgono enormi quantità di informazioni sulla popolazione e sul territorio, ma che nell'attività quotidiana non interagiscono a questo livello di dettaglio, limitandosi a scambiarsi solo i dati strettamente necessari per l'attività routinaria.

Questo avviene per tante ragioni che, come visto nel capitolo 4, vanno da problematiche di tipo normativo legate alla gestione di dati sensibili, alla possibilità tecnica di operare *linkage* con ridotta perdita di informazioni, alla volontà politica/amministrativa e all'interesse specifico a sviluppare tali procedure.

In questo caso si è verificato che, laddove esistano già processi regolamentati di trasmissione delle informazioni all'interno del sistema informativo di un'istituzione (come è l'ente regione), molti problemi di natura tecnica possono essere affrontati e risolti.

Il risultato è che la ricchezza di informazioni che gli enti territoriali di varia natura raccolgono quotidianamente sull'individuo possono, una volta collegati correttamente, rispondere ad importanti quesiti di ricerca che si traducono in potenziali indicazioni di azione a favore del cittadino stesso.

Ovviamente, vanno ricordati alcuni limiti di questo lavoro.

Innanzitutto, benché la scelta della regione in cui è stato svolto lo studio sia stata guidata non solo da elementi di praticità empirica⁴⁶, ma anche dall'aver voluto cercare un'area che fosse sufficientemente omogenea con le caratteristiche medie del territorio italiano per garantire una più ampia confrontabilità dei risultati, il sistema sanitario umbro funziona particolarmente bene, tanto da far registrare saldi positivi in termini di migrazione sanitaria (le strutture regionali attraggono pazienti da alcune aree limitrofe extraregionali) (ISTAT – Health for All Italia; 2015).

Questo potrebbe spiegare, almeno in parte, perché il capitale culturale ed economico tradizionalmente associato alle disuguaglianze in salute e, in particolare, alle differenze di sopravvivenza per i due tumori considerati, perde di importanza a fronte di altre variabili che identificano i bisogni di una popolazione già ben assistita a livello di base. Se a livello nazionale questo risultato sembra piuttosto solido, anche a fronte del fatto che il sistema sanitario pubblico italiano è fra i migliori in Europa in termini di risultati di cura, in particolare per quanto riguarda i tumori (si veda in merito l'ultimo rapporto sulla qualità del servizio sanitario italiano ad opera dell'OECD, 2015), è probabile che in regioni dove la qualità è inferiore riemergano disuguaglianze legate al capitale economico e culturale del malato.

Un'altra considerazione è legata all'arco temporale considerato per l'incidenza della malattia, ossia il periodo 2001-2010 con fine del follow-up al 31 Dicembre 2012.

In tale periodo, infatti, gli effetti della grande crisi economica mondiale su individui, famiglie e sistema sanitario sono stati sicuramente ancora limitati e potrebbe essere interessante ampliare in futuro l'intervallo temporale ad anni più recenti per valutare meglio questo aspetto.

In realtà, esistono pochi studi su questo tema sviluppati in nazioni con sistemi sanitari pubblici universalistici simili a quello italiano, anche perché i dati epidemiologici di popolazione stanno solo recentemente raggiungendo il periodo che potrebbe essere interessato dalla recessione.

A dire il vero, i pochi studi in materia, peraltro su un arco temporale simile a quello considerato in questa sede, sembrano mostrare come in un contesto pubblico lo stato di

⁴⁶ In particolare, la buona disponibilità e qualità dei dati e la già avviata integrazione fra le diverse fonti informative territoriali.

salute della popolazione non sia peggiorato, se non per alcune patologie. Ad esempio, il lavoro di Regidor et al. (2014) ha mostrato come in Spagna, pur pesantemente colpita dalla crisi economica, la condizione di salute della popolazione ha continuato a migliorare nel periodo considerato (2004-2012), fatta eccezione a dire il vero per la mortalità da tumore (di cui abbiamo visto l'impatto estremamente rilevante sui costi per il sistema). Uno studio precedente, sempre spagnolo, aveva preso in considerazione proprio il tumore al colon-retto e le attività di prevenzione e cura, dimostrando come, attraverso un'opportuna attività di screening opportunamente mirata in termini di popolazione-bersaglio, si potevano ridurre i costi della cura grazie ad un'individuazione precoce della malattia e garantire adeguati livelli di qualità terapeutica alla popolazione anche in tempi di crisi (Carballo et al., 2012). Questa è, peraltro, anche la strada intrapresa dal sistema sanitario della regione umbra.

Malgrado questi limiti, i risultati ottenuti sembrano essere piuttosto rilevanti per capire alcuni aspetti emergenti del rapporto fra differenze socio-economiche, disuguaglianze sanitarie e invecchiamento delle popolazioni.

Tutto ciò permette di concludere questo lavoro con un'ultima raccomandazione: nessuno dovrebbe affrontare da solo il tumore, perché le relazioni sociali possono supplire a determinate mancanze individuali nell'affrontare la malattia e che, al contrario, l'isolamento sociale può peggiorare lo stato di salute. Ogni ammalato, che usufruisca di un determinato apporto sociale opportunamente gestito da parte degli operatori sanitari, stimola la lotta contro il tumore, rende più umana e socialmente condivisibile un'esperienza per molti aspetti alienante come quella della malattia neoplastica e facilita il corso dei trattamenti e della riabilitazione, aumentando il numero di vite salvate.

Bibliografia

Abadi A, Yavari P, Dehghani-Arani M, Alavi-Majd H, Ghasemi E, Amanpour F, Bajdik C. (2014) Cox models survival analysis based on breast cancer treatments. *Iranian Journal of Cancer Prevention*; 7(3): 124-129.

Adler NE, Boyce T, Chesney MA, Cohen S, Folkman S, Kahn RL, Syme SL. (1994) Socioeconomic status and health. The challenge of the gradient. *American Psychologist*; 49: 15–24.

Adler NE, Ostrove JM. (1999) Socioeconomic status and health: what we know and what we don't. *Annals of the New York Academy of Sciences*; 896: 3-15.

AIRTUM (2014) ITACAN – I Tumori in Italia v. 2.0. Firenze, AIRTUM-ISPO. <http://itacan.ispo.toscana.it/italian/itacan.htm>

AIRTUM Working Group. (2006) Italian Cancer Figures – Report 2006. *Epidemiologia e Prevenzione*; Supplemento 2.

AIRTUM Working Group. (2011) Italian cancer figures, report 2011: Survival of cancer patients in Italy. *Epidemiologia e Prevenzione*; 35(5-6 Supplemento 3): 1-200.

AIRTUM Working Group (2014) Prevalenza e guarigione da tumore in Italia. *Epidemiologia e Prevenzione*; 38(6, Supplemento 1): 1-122.

AIRTUM/AIOM (2014) I numeri del cancro in Italia – 2014. Brescia, Intermedia Editore.

Aizer AA, Chen MH, McCarthy EP, Mendu ML, Koo S, Wilhite TJ, Graham PL, Choueiri TK, Hoffman KE, Martin NE, Hu JC, Nguyen PL. (2013) Marital status and survival in patients with cancer. *Journal of Clinical Oncology*; 31(31): 3869-3876.

Akaike H. (1974) A new look at the statistical model identification. *IEEE Transactions on Automatic Control*; 19(6): 716–723.

Akaike H. (1985) Prediction and entropy. In: Atkinson AC, Fienberg SE. (1985) *A Celebration of Statistics*. New York, Springer.

Akinyemiju TF, Soliman AS, Johnson NJ, Altekruse SF, Welch K, Banerjee M, Schwartz K, Merajver S. (2013) Individual and neighborhood socioeconomic status and healthcare resources in relation to black-white breast cancer survival disparities. *Journal of Cancer Epidemiology*; 2013: 1-13.

Allemani C, Sant M, Weir HK, Richardson LC, Baili P, Storm H, Siesling S, Torrella-Ramos A, Voogd AC, Aareleid T, Ardanaz E, Berrino F, Bielska-Lasota M, Bolick S, Cirilli C, Colonna M, Contiero P, Cress R, Crocetti E, Fulton JP, Grosclaude P,

Hakulinen T, Izarzugaza MI, Malmström P, Peignaux K, Primic-Žakelj M, Rachtan J, Safaei Diba C, Sánchez MJ, Schymura MJ, Shen T, Traina A, Tryggvadottir L, Tumino R, Velten M, Vercelli M, Wolf HJ, Woronoff AS, Wu X, Coleman MP. (2013) Breast cancer survival in the US and Europe: a CONCORD high-resolution study. *International Journal of Cancer*; March, 1, 132(5): 1170-1181.

Allemani C, Weir HK, Carreira H, Harewood R, Spika D, Wang XS, Bannon F, Ahn JV, Johnson CJ, Bonaventure A, Marcos-Gragera R, Stiller C, Azevedo e Silva G, Chen WQ, Ogunbiyi OJ, Rachet B, Soeberg MJ, You H, Matsuda T, Bielska-Lasota M, Storm H, Tucker TC, Coleman MP; CONCORD Working Group. (2015) Global surveillance of cancer survival 1995-2009: analysis of individual data for 25,676,887 patients from 279 population-based registries in 67 countries (CONCORD-2). *The Lancet*; 385(9972): 977-1010.

Anderson NB, Armstead CA. (1995) Toward understanding the association of socioeconomic status and health: a new challenge for the biopsychosocial approach. *Psychosomatic Medicine*; 57: 213–225.

Annunziata MA, Muzzati B. (2011) Le dimensioni psicosociali del cancro. *Nóos*; 2: 73-86.

Antonovsky A. (1987) *Unravelling the mystery of health – How people manage stress and stay well*. London, Jossey-Bass.

Ardissone A, Carletti R (2013) Le teorie sulle diseguaglianze in sanità. In: (2013) *Disuguaglianze sociali in sanità*. Milano, Franco Angeli.

Ardissone A, Maturo A. (a cura di) (2013) *Disuguaglianze sociali e vissuto nei malati oncologici*. Milano, FrancoAngeli.

Arman M, Rehnsfeldt A, Lindholm L, Hamrin E. (2002) The face of suffering among women with breast cancer-being in a field of forces. *Cancer Nursing*; 25(2): 96-103.

Arndt V, Merx H, Sturmer T, Stegmaier C, Ziegler H, Brenner H. (2004). Age-specific detriments to quality of life among breast cancer patients one year after diagnosis. *European Journal of Cancer*; 40: 673-680.

Artazcoz L, Benach J, Borrel C, Cortes I. (2004) Unemployment and mental health: understanding the interactions among gender, family roles and social class. *American Journal of Public Health*; 94: 82–88.

Auvinen A, Karjalainen S. (1997) Possible explanations for social class differences in cancer patient survival. In: Kogevinas M, Pearce N, Susser M, Boffetta P (eds.) (1997) *Social Inequalities and Cancer*. IARC Scientific Publications No. 138. Lyon; IARC.

Ayanian JZ, Zaslavsky AM, Fuchs CS, Guadagnoli E, Creech CM, Cress RD, O'Connor LC, West DW, Allen ME, Wolf RE, Wright WE. (2003) Use of adjuvant chemotherapy

and radiation therapy for colorectal cancer in a population-based cohort. *Journal of Clinical Oncology*; 21(7): 1293-1300.

Aziz NM. (2007). Late effects of treatment. In: Ganz PA (ed.) (2007) *Cancer survivorship: Today and tomorrow*. New York, Springer.

Baider L, Ever-Hadani P, Goldzweig G. (2003) Is perceived family support a relevant variable in psychological distress? A sample of prostate and breast cancer couples. *Journal of Psychosomatic Research*; 55: 453–460.

Baili P, Vicentini M, Tumino R, Vercelli M, Lorenzo M, Foschi R, Guzzinati S, Dal Maso L, Minicozzi P, de Lorenzo F, Micheli A, di Salvo F; CAREMORE Group. (2013) A method for differentiating cancer prevalence according to health status, exemplified using a population-based sample of Italian colorectal cancer cases. *Acta Oncologica*; 52(2): 294-302.

Bal O, Yalcintas Arslan U, Durnali A, Uyetrk U, Demirci A, Tastekin D, Ekinici A, Esbah O, Turker I, Uysal Sonmez O, Oksuzoglu B. (2015) Progesterone receptor status in determining the prognosis of estrogen receptor positive/ HER2 negative breast carcinoma patients. *Journal of BUON*; 20(1): 28-34.

Baquet CR, Commiskey P. (2000) Socioeconomic factors and breast carcinoma in multicultural women. *Cancer*; 88 (5 Suppl.): 1256–1264.

Barry J, Breen N. (2005) The importance of place of residence in predicting late-stage diagnosis of breast or cervical cancer. *Health Place*; 11: 15–29.

Bassett MT, Krieger N. (1986) Social class and black-white differences in breast cancer survival. *American Journal of Public Health*; 76: 1400-1403.

Belin TR, Rubin DB. (1995) A method for calibrating false-match rates in record linkage. *Journal of the American Statistical Association*; 90: 694-707.

Bentley R, Kavanagh AM, Subramanian SV, Turrell G. (2009) Area disadvantage, individual socio-economic position, and premature cancer mortality in Australia 1998 to 2000: a multilevel analysis. *Cancer Causes & Control*; 19(2): 183-193.

Berg JW, Ross R, Latourette HB. (1977) Economic status and survival of cancer patients. *Cancer*; 39: 467-477.

Berkman L, Syme S. (1979) Social networks, host resistance and mortality: a nine-year follow-up of Alameda County residents. *American Journal of Epidemiology*; 109: 186–204.

Berrino F, Sant M, Verdecchia A, Capocaccia R, Hakulinen T, Esteve J. (eds). (1995) *Survival of cancer patients in Europe. The EURO CARE study*. Lyon, IARC.

- Berry DA, Cronin KA, Plevritis SK, Fryback DG, Clarke L, Zelen M, Mandelblatt JS, Yakovlev AY, Habbema JD, Feuer EJ. (2005) Effect of screening and adjuvant therapy on mortality from breast cancer. *New England Journal of Medicine*; 353: 1784–1792.
- Bianconi F, Brunori V, Valigi P, Stracci F, La Rosa F. (2010) Cancer Registry and information technology: a new management system for integrating Cancer Registry and Oncology Departments. *Health Care Management (WHCM) IEEE Workshop on*; Feb.: 1-5.
- Bianconi F, Brunori V, Valigi P, Stracci F, La Rosa F. (2012) Information Technology as Tools for Cancer Registry and Regional Cancer Network Integration. *Systems, Man and Cybernetics, Part A: Systems and Humans, IEEE Transactions on*; 42(6): 1410-1424.
- Biggeri A, Grisotto L. (2009) Fonti di distorsione nella misura delle disuguaglianze di salute: la validazione, il confronto temporale e spaziale, l'aggiustamento per altre covariate, il bias ecologico. In: Costa G, Cislighi C, Caranci N. (eds.) (2009) *Le disuguaglianze sociali di salute. Problemi di definizione e di misura*. Milano, Franco Angeli.
- Biolcati Rinaldi F, Sarti S, Vitalini A. (2013) Le disuguaglianze sociali nella salute in Lombardia e in Italia (2010-2012): un'analisi multilivello delle condizioni individuali e degli effetti di contesto dei comuni. *Working Paper del Dipartimento di Scienze Sociali e Politiche dell'Università degli Studi di Milano*; 3: 1-16.
- Biondi M, Costantini A, Grassi L. (1995) *La mente e il cancro. Insidie e risorse della psiche nelle patologie tumorali*. Roma, Il Pensiero Scientifico Editore.
- Black R. (2005) Ensuring high-quality epidemiological data on cancer. *European Journal of Cancer Prevention*; 14(4): 305-306.
- Blais S, De Jardin O, Boutreux S, Launoy G. (2006) Social determinants of access to reference care centres for patients with colorectal cancer. A multilevel analysis. *European Journal of Cancer*; 42(17): 3041-3048.
- Blalock HM. (1964) *Causal inference in non-experimental research*. Chapel Hill, University of North Carolina Press.
- Blaxter M. (1990) *Health and lifestyles*. London, Tavistock-Routledge.
- Bloom JR, Stewart SL, Johnston M, Banks P, Fobair P. (2001) Sources of support and the physical and mental well-being of young women with breast cancer. *Social Science and Medicine*; 53(11): 1513-1524.
- Bolard P, Quantin C, Estève J, Faivre J, Abrahamowicz M. (2001) Modelling time-dependent hazard ratios in relative survival: Application to colon cancer. *Journal of Clinical Epidemiology*; 54: 986–996.

Bosma H, Peter R, Siegrist J, Marmot M. (1998) Alternative job stress models and the risk of heart disease. *American Journal of Public Health*; 88: 68–74.

Bouchardy C, Rapiti E, Fioretta G, Laissue P, Neyroud-Caspar I, Schafer P, Kurtz J, Sappino AP, Vlastos G. (2003) Undertreatment strongly decreases prognosis of breast cancer in elderly women. *Journal of Clinical Oncology*; 21: 3580-3587.

Bouchardy C, Verkooijen HM, Fioretta G. (2006) Social class is an important and independent prognostic factor of breast cancer mortality. *International Journal of Cancer*; 119(5): 1145-1151.

Bourdieu P. (1984) *La Distinzione. Critica sociale del gusto*. Bologna, Il Mulino.

Bower JE, Ganz PA, Desmond KA, Bernaards C, Rowland JH, Meyerowitz BE, Belin TR. (2006) Fatigue in long-term breast carcinoma survivors: a longitudinal investigation. *Cancer*; 106(4): 751-758.

Boyd C, Zhang-Salomons JY, Groome PA, Mackillop WJ. (1999) Associations between community income and cancer survival in Ontario, Canada, and the United States. *Journal of Clinical Oncology*; 17: 2244-2255.

Braaten T, Weiderpass E, Lund E. (2009) Socioeconomic differences in cancer survival: the Norwegian Women and Cancer Study. *BMC Public Health*; 9: 178.

Braun MM, Caporaso NE, Page WF, Hoover RN. (1995) A cohort study of twins and cancer. *Cancer Epidemiology, Biomarkers and Prevention*; 4(5): 469-473.

Bray F, Parkin DM. (2009) Evaluation of data quality in the cancer registry: principles and methods. Part I: comparability, validity and timeliness. *European Journal of Cancer*; 45(5): 747-755.

Brenner M. (1979) Mortality and the national economy. *Lancet*; 2: 568–573.

Brown G, Harris T. (1978) *The social origins of depression*. London, Tavistock.

Brown ML, Lipscomb J, Snyder C. (2001) The burden of illness of cancer: economic cost and quality of life. *Annual Review of Public Health*; 22: 91–113.

Brunner E, Marmot M. (2006) Social organization, stress and health. In: Marmot M, Wilkinson R. (eds.) (2006) *Social determinants of health*. Oxford, Oxford University Press.

Bultz BD, Carlson LE. (2006) Emotional distress: the sixth vital sign. Future directions in cancer care. *Psychooncology*; 15(2): 93-95.

Byers TE, Wolf HJ, Bauer KR, Bolick-Aldrich S, Chen VW, Finch JL, Fulton JP, Schymura MJ, Shen T, Van Heest S, Yin X; Patterns of Care Study Group. (2008) The impact of socioeconomic status on survival after cancer in the United States: Findings

from the National Program of Cancer Registries Patterns of Care Study. *Cancer*; 113(3): 582-591.

Cadum E, Costa G, Biggeri A. (1999) Deprivazione e mortalità: un indice di deprivazione per l'analisi delle disuguaglianze su base geografica. *Epidemiologia e Prevenzione*; 23: 175-187.

Caiazzo A, Cardano M, Cois E, Costa G, Marinacci C, Spadea T, Vannoni F, Venturini L. (2004) Diseguaglianze di salute in Italia. *Epidemiologia e Prevenzione*; 28(Suppl. 3): 1-161.

Caldarella A, Crocetti E, Messerini L, Paci E. (2014) Trends in colorectal incidence by anatomic subsite from 1985 to 2005: a population-based study. *International Journal of Colorectal Disease*; 28(5): 637-641.

Canadian Agency for Drugs and Technologies in Health. (2013) Prostate Cancer Screening: A Review of the Guidelines. Ottawa, CADTH Rapid Response Reports.

Cancer Research UK Cancer Survival Group. (2009) Life tables for England and Wales by sex, calendar period, region and deprivation. <http://www.lshtm.ac.uk/ncdeu/cancersurvival/tools/>

Canestrini P. (2013) Il valore della relazione sociale per il malato oncologico. In: Ardisson A, Maturo M. (2013) *Disuguaglianze sociali e vissuto dei malati oncologici*. Milano, Franco Angeli.

Caranci N, Biggeri A, Grisotto L, Pacelli B, Spadea T, Costa G. (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.

Carnon AG, Ssemwogerere A, Lamont DW. (1994) Relation between socioeconomic deprivation and pathological prognostic factors in women with breast cancer. *BMJ*; 309: 1054-1057

Cash TF. (2004) Body image: Past, present and future. *Body Image*; 1: 1-5.

Cassileth BR, Lusk EJ, Miller DS, Brown LL, Miller C. (1985) Psychosocial correlates of survival in advanced malignant disease. *New England Journal of Medicine*; 312: 1551-1555.

Cavalli-Björkman N, Lambe M, Eaker S, Sandin F, Glimelius B. (2011) Differences according to educational level in the management and survival of colorectal cancer in Sweden. *European Journal of Cancer*; 47(9): 1398-1406.

Centro Nazionale di Epidemiologia, Sorveglianza e Promozione della Salute. (2015) PASSI. Il Portale dell'Epidemiologia per la Sanità Pubblica. Roma, Istituto Superiore di Sanità. <http://www.epicentro.iss.it/passi/default.asp>

Challier B, Meslans Y, Viel JF. (2000) Deprived areas and attendance to screening of cervix uteri cancer in a French region. *Cancer Causes*; 11(2): 157-162.

Chandola T, Brunner E, Marmot M. (2006) Chronic stress at work and the metabolic syndrome: Prospective study. *BMJ*; 332: 521–525.

Chandola T, Sigerist J, Marmot M. (2005) Do changes in effort-reward imbalance at work contribute to an explanation of the social gradient in angina? *Occupational and Environmental Medicine*; 62: 223–230.

Chang SM, Barker FG. (2005) Marital status, treatment, and survival in patients with glioblastoma multiform: a population based study. *Cancer*; 104(9): 1975-1984.

Chen L, Cnattingius S, Nyman Iliadou A, Oberg AS. (2016) Cancer risks in twins and singletons from twin and non-twin families. *International Journal of Cancer*; 138(5): 1102-1110.

Chieppa A, Panizon F. (2001) Data quality control system for the 2001 Italian population census. *International Conference on quality in official statistics, Q2001, Stoccolma*: 14-15 Maggio.

Chow, E., Tsao, M., & Harth, T. (2004) Does psychosocial intervention improve survival in cancer? A meta-analysis. *Palliative Medicine*; 18: 25-31.

Christensen K, Doblhammer G, Rau R, Vaupel JW. (2009) Ageing populations: The challenges ahead. *The Lancet*; 374(9696): 1196-1208.

Ciccolallo L, Capocaccia R, Coleman MP. (2005) Survival differences between European and US patients with colorectal cancer: role of stage at diagnosis and surgery. *Gut*; 54: 268–273.

Cicccone G, Prastaro C, Ivaldi C. (2000) Access to hospital care, clinical stage and survival from colorectal cancer according to socio-economic status. *Annals of Oncology*; 11: 1201–1204.

Cipolla C, Remuzzi G. (a cura di) (2008) *Dire, fare, curare. Le parole tra medico e malato*. Milano, Franco Angeli.

Cislaghi C. (2009) Disuguaglianze a livello individuale e a livello di contesto: significati e indicazioni per l'uso. In: Costa G, Cislaghi C, Caranci N. (eds.) (2009) *Le disuguaglianze sociali di salute. Problemi di definizione e di misura*. Milano, Franco Angeli.

Clemmesen J. (1965-1974) *Statistical studies in the aetiology of malignant neoplasms*. Vol. I-IV. Munksgaard Kobenhaven.

Cockerham WC. (2005) Health Lifestyle Theory and the convergence of Agency and Structure. *Journal of Health and Social Behaviour*; 43(3): 51-67.

- Cockerham WC. (2013) *Social Causes of Health and Disease*. Cambridge, Polity.
- Coleman MP, Babb P, Damiecki P et al. (1999) Cancer survival trends in England and Wales 1971–1995: deprivation and NHS Region. Series SMPS No. 61. London, The Stationery Office.
- Coleman MP, Rachet B, Woods LM, Mitry E, Riga M, Cooper N, Quinn MJ, Brenner H, Estève J. (2004) Trends and socioeconomic inequalities in cancer survival in England and Wales up to 2001. *British Journal of Cancer*; 90: 1367–1373.
- Collin SM, Martin RM, Metcalfe C, Gunnell D, Albertsen PC, Neal D, Hamdy F, Stephens P, Lane JA, Moore R, Donovan J. (2008) Prostate-cancer mortality in the USA and UK in 1975-2004: an ecological study. *Lancet Oncology*; May, 9(5): 445-452.
- Cornfield J. (1951) A method of estimating comparative rates from clinical data; applications to cancer of the lung, breast and cervix. *Journal of National Cancer Institute*; 11: 1269-75.
- Costanzo ES, Ryff CD, Singer BH. (2009) Psychosocial adjustment among cancer survivors: Findings from a national survey of health and well-being. *Health Psychology*, 28: 147-156.
- Coups EJ, Manne SL, Meropol NJ, Weinberg DS. (2007) Multiple Behavioral Risk Factors for Colorectal Cancer and Colorectal Cancer Screening Status. *Cancer Epidemiology, Biomarkers and Prevention*; 16(3): 510–516.
- Crocetti E, Buzzoni C, Quaglia A, Lillini R, Vercelli M, AIRTUM Working Group. (2012) Ageing and other factors behind recent cancer incidence and mortality trends in Italy. *Journal of Geriatric Oncology*; 3(2): 111-119.
- Crowther MJ, Look MP, Riley RD. (2014) Multilevel mixed effects parametric survival models using adaptive Gauss-Hermite quadrature with application to recurrent events and individual participant data meta-analysis. *Statistics in Medicine*; 33(22): 3844-3858.
- Cummins S, Stafford M, Macintyre S, Marmot M, Ellaway A. (2005) Neighborhood environment and its association with self-rated health: Evidence from Scotland and England. *Journal of Epidemiology & Community Health*; 59: 207–213.
- Dahlgren G, Whitehead M. (1991) *Policies and strategies to promote social equity*. Stockholm, Health Institute of Future Studies.
- Dal Maso L, Braga C, Franceschi S. (2001) Methodology used for “Software for Automated Linkage in Italy” (SALI). *Journal of Biomedical Informatics*; 34: 387-395.
- Dalton SO, Ross L, Düring M, Carlsen K, Mortensen PB, Lynch J, Johansen C. (2007) Influence of socioeconomic factors on survival after breast cancer. A nationwide cohort

study of women diagnosed with breast cancer in Denmark 1983-1999. *International Journal of Cancer*; 121(11): 2524-2531.

Dayal HH, Power RN, Chiu C. (1982) Race and socio-economic status in survival from breast cancer. *Journal of Chronic Diseases*; 35(8): 675-683.

De Angelis R, Sant M, Coleman MP, Francisci S, Baili P, Pierannunzio D, Trama A, Visser O, Brenner H, Ardanaz E, Bielska-Lasota M, Engholm G, Nencke A, Siesling S, Berrino F, Capocaccia R; EUROCARE-5 Working Group. (2014) Cancer survival in Europe 1999-2007 by country and age: results of EUROCARE-5-a population-based study. *Lancet Oncology*; January, 15(1): 23-34.

De Cuyper L, Van Hee R, Sterckx F. (2003) The effect of social factors on diagnosis and treatment of breast cancer. *Acta Chirurgica Belgica*; 103(6): 585-588.

Dejardin O, Bouvier AM, Faivre J, Boutreux S, De Pourville G, Launoy G. (2008) Access to care, socioeconomic deprivation and colon cancer survival. *Alimentary Pharmacology & Therapeutics*; 27(10): 940-949.

Denberg TD, Beaty BL, Kim FJ, Steiner JF. (2005) Marriage and ethnicity predict treatment in localized prostate carcinoma. *Cancer*; 103: 1819-1825.

Department of Health and Social Security UK. (1980) *The Black Report*. London, Penguin Books.

Dialla PO, Quipourt V, Gentil J, Marilier S, Poillot ML, Roignot P, Altwegg T, Darut-Jouve A, Guiu S, Arveux P, Dabakuyo-Yonli TS. (2015) In breast cancer, are treatments and survival the same whatever a patient's age? A population-based study over the period 1998-2009. *Geriatrics and Gerontology International*; 15(5): 617-626.

Dickerson SS, Kemeny ME. (2004) Acute stressors and cortisol responses: a theoretical integration and synthesis of laboratory research. *Psychological Bulletin*; 130(3): 355-391.

Dickman PW, Coviello E, Hills M. (2013) Estimating and modelling relative survival. *The Stata Journal*; 2: 1-25.

Dickman PW, Coviello E. (2015) Estimating and modelling relative survival. *The Stata Journal*; 15(1): 186-215.

Dickman PW, Adami HO. (2006). Interpreting trends in cancer patient survival. *Journal of Internal Medicine*; 260(2): 103-117.

Dickman PW, Sloggett A, Hills M, Hakulinen T. (2004). Regression models for relative survival. *Statistics in Medicine*; 23(1): 51-64.

Diez-Roux AV. (1998) Bringing context back into epidemiology: variables and fallacies in multilevel analysis. *American Journal of Public Health*; 88: 216-222.

Diez-Roux AV. (2000) Multilevel analysis in public health research. *Annual Review of Public Health*; 21: 171–192.

Dittner AJ, Wessely SC, Brown RG. (2004) The assessment of fatigue: a practical guide for clinicians and researchers. *Journal of Psychosomatic Research*; 56(2): 157-170.

Dobie SA, Baldwin LM, Dominitz JA, Matthews B, Billingsley K, Barlow W. (2006) Completion of therapy by Medicare patients with stage III colon cancer. *Journal of National Cancer Institute*; 98(9): 610-619.

Doll R, Peto R. (1981) Sources of bias in estimating trends in cancer mortality, incidence and curability. Quantitative estimates of avoidable risks of cancer in the United States today. Oxford, Oxford University Press.

Doll R. (2004) Cohort studies: history of the method. In: Morabia A. (ed). *A history of epidemiologic methods and concepts*. Basel, Birkhauser Verlag.

Dossus L, Boutron-Ruault MC, Kaaks R, Gram IT, Vilier A, Fervers B, Manjer J, Tjonneland A, Olsen A, Overvad K, Chang-Claude J, Boeing H, Steffen A, Trichopoulou A, Laggiou P, Sarantopoulou M, Palli D, Berrino F, Tumino R, Vineis P, Mattiello A, Bueno-de-Mesquita HB, van Duijnhoven FJ, Bakker MF, Peeters PH, Weiderpass E, Bjerkaas E, Braaten T, Menéndez V, Agudo A, Sanchez MJ, Amiano P, Tormo MJ, Barricarte A, Butt S, Khaw KT, Wareham N, Key TJ, Travis RC, Rinaldi S, McCormack V, Romieu I, Cox DG, Norat T, Riboli E, Clavel-Chapelon F. (2014) Active and passive cigarette smoking and breast cancer risk: results from the EPIC cohort. *International Journal of Cancer*; 134(8): 1871-1888.

Dunn HL. (1946) Record linkage. *American Journal of Public Health*; 36: 1412-1416.

Durkheim E. (1897) *Il suicidio*. *Studio di sociologia (Le Suicide, étude de sociologie)*.

Ederer F, Axtell LM, Cutler SJ. (1961) The relative survival rate: a statistical methodology. *National Cancer Institute Monograph*; 6: 101–121.

Esteban D, Whelan SL, Laudico A, Parkin DM et al. (1995) *Manual for cancer registry personnel*. Lyon, IARC Technical report n. 10.

Estève J, Benhamou E, Raymond L. (1994) *Statistical Methods in Cancer Research - Volume IV: Descriptive Epidemiology*. IARC Scientific Publications No. 128. Lyon, IARC.

European network of cancer registries. (1995) *EUROCIM User manual*. 2nd edition, Lambda plus software products.

Evans RG, Barer ML, Marmor TR. (eds.) (1994) *Why are some people healthy and others not? The determinants of health of populations*. New York, Aldine De Gruyter.

Fall K, Strömberg F, Rosell J, Andrèn O, Varenhorst E; South-East Region Prostate Cancer Group. (2008) Reliability of death certificates in prostate cancer patients. *Scandic Journal of Urology and Nephrology*; 42(4): 352-357.

Farr W. (1885) The mortality of cholera in England, 1848-49 and 17th Annual report. In: Humphreys NA. (ed). *Vital statistics. A memorial volume of selections from the reports and writings of William Farr*. London, Sanitary Institute of Great Britain.

Fellegi IP, Sunter A. (1969) A theory of record linkage. *Journal of American Statistics Association*; 64: 1183-1210.

Fentiman IS. (1996) Are the elderly receiving appropriate therapy for cancer? *Annals of Oncology*; 7: 657-658.

Ferrante JM, Gonzalez EC, Roetzheim RG, Pal N, Woodard L. (2000) Clinical and demographic predictors of late-stage cervical cancer. *Archives of Family Medicine*; 9: 439-445.

Ferretti S, Giacomini A, Gruppo di lavoro AIRTUM. (eds.) (2007) *Manuale di Tecniche di Registrazione dei Tumori*. Milano, Inferenze Scarl.

Feuerstein M. (2007) Cancer survivorship: A bird's eye view from an insider. In: Feuerstein M. (ed.) (2007) *Handbook of cancer survivorship*. New York, Springer.

Fisher RA. (1937) *The design of experiments*. London, Oliver and Boyd.

Fitzpatrick R, Newman S, Archer R, Shipley M. (1991) Social support, disability and depression: a longitudinal study of arthritis and depression. *Social Science & Medicine*; 33: 605-611.

Foley KL, Farmer DF, Petronis VM, Smith RG, McGraw S, Smith K, Carver CS, Avis N. (2006) A qualitative exploration of the cancer experience among long-term survivors: comparisons by cancer type, ethnicity, gender and age. *Psycho-oncology*; 15(3): 248-258.

Foster C, Fenlon D. (2011) Recovery and self-management support following primary cancer treatment. *British Journal of Cancer*; 105: 21-28.

Friedman GD. (1967) Cigarette smoking and geographic variation in coronary heart disease mortality in the United States. *Journal of Chronic Diseases*; 20(10): 769-779.

Gakidou E, Nordhagen S, Obermeyer Z. (2008) Coverage of cervical cancer screening in 57 countries: low average levels and large inequalities. *PLoS Medicine*; 5(6): e132.

Gallo V, Mackenbach JP, Ezzati M, Menvielle G, Kunst AE, Rohrmann S, Kaaks R, Teucher B, Boeing H, Bergmann MM, Tjønneland A, Dalton SO, Overvad K, Redondo ML, Agudo A, Daponte A, Arriola L, Navarro C, Gurrea AB, Khaw KT, Wareham N, Key T, Naska A, Trichopoulou A, Trichopoulos D, Masala G, Panico S, Contiero P, Tumino R, Bueno-de-Mesquita HB, Siersema PD, Peeters PP, Zackrisson S, Almquist

- M, Eriksson S, Hallmans G, Skeie G, Braaten T, Lund E, Illner AK, Mouw T, Riboli E, Vineis P. (2012) Social inequalities and mortality in Europe--results from a large multi-national cohort. *PLoS One*; 7(7): e39013.
- Galobardes B, Lynch J, Smith GD. (2007) Measuring socioeconomic position in health research. *British Medical Bulletin*; 81-82:21-37.
- Garofalo G, Cella P, Paggiaro A, Torelli N, Viviano C. (2001) Demografia d'impresa: utilizzo di tecniche di abbinamento per l'analisi della continuità. Roma, ISTAT.
- Garofalo G. (1998) The ASIA Project (setting up of the Italian business register) synthesis of the methodological manual. In: *Proceedings of the 12th Meeting of the International Roundtables on Business Survey Frames*, Helsinki, Settembre.
- Giusti A, Marliani G, Torelli N. (1991) Procedure per l'abbinamento dei dati individuali delle forze di lavoro. In: *Forze di Lavoro: Disegno dell'Indagine e Analisi Strutturali*, *Annali di Statistica*, Roma, ISTAT, serie IX, A 120: 121-148.
- Glanz K, Maskarinec G, Carlin L. (2005) Ethnicity, sense of coherence and tobacco use among adolescents. *Annals of Behavioral Medicine*; 29: 192-199.
- Glaser SL, Clarke CA, Gomez SL et al. (2005) Cancer surveillance research: a vital subdiscipline of cancer epidemiology. *Cancer Causes Control*; 16(9): 1009-1019.
- Goldman DP, Lakdawalla DN. (2005) A theory of health disparities and medical technology. *Contributions to Economic Analysis and Policy*; 4: 1-30.
- Goldstein H. (1995) *Multilevel statistical models*. New York, Halsted Press.
- Goodman MT, Hernandez BY, Hewitt S et al. (2005) Tissues from population-based cancer registries: a novel approach to increasing research potential. *Human Pathology*; 36(7): 812-820.
- Goodwin JS, Hunt C, Samet J. (1987) Relationship of marital status to stage at diagnosis, choice of treatment and survival in individuals with cancer. *Journal of American Medical Association*; 258: 3125-3130.
- Goodwin JS, Hunt WC, Samet JM. (1993) Determinants of cancer therapy in elderly patients. *Cancer*; 72: 594-601.
- Gorey KM, Holowaty EJ, Fehringer G, Laukkanen E, Moskowitz A, Webster DJ, Richter NL. (1997) An international comparison of cancer survival: Toronto, Ontario, and Detroit, Michigan, metropolitan areas. *American Journal of Public Health*; 87: 1156-1163.
- Gorey KM, Holowaty EJ, Fehringer G, Laukkanen E, Richter NL, Meyer CM. (2000) An international comparison of cancer survival: relatively poor areas of Toronto,

Ontario and three US metropolitan areas. *Journal of Public Health Medicine*; 22: 343-348.

Gove W. (1979) Sex, marital status and mortality. *American Journal of Sociology*; 79: 45–67.

Graunt J. (1662) *Natural and political observations made upon the Bills of mortality*.

Grunfeld EA, Ramirez AJ, Hunter MS, Richards MA. (2002) Women's knowledge and beliefs regarding breast cancer. *British Journal of Cancer*; 86: 1373–1378.

Guralnik JM, Land KC, Blazer D, Fillenbaun GG, Branch LG. (1993) Educational status and active life expectancy among older blacks and whites. *New England Journal of Medicine*; 329: 110-116.

Guy GP Jr, Ekwueme DU, Yabroff KR, Dowling EC, Li C, Rodriguez JL, de Moor JS, Virgo KS. (2013) Economic burden of cancer survivorship among adults in the United States. *Journal of Clinical Oncology*; 31(30): 3749-3757.

Hallqvist J, Diderichsen F, Theorell Y and the SHEEP study. (1998) Is the effect of job strain on myocardial infarction due to interaction between high psychological demands and low decision latitude. Results from the Sweden Heart Epidemiology Program. *Social Science & Medicine*; 46: 1405–1415.

Halpern D. (1995) *Mental health and the built environment. More than bricks and mortar?* London, Taylor and Francis.

Hansen RP, Olesen F, Sørensen HT, Sokolowski I, Søndergaard J. (2008) Socioeconomic patient characteristics predict delay in cancer diagnosis: a Danish cohort study. *BMC Health Services Research*; 8: 49-58.

Harteloh P, de Bruin K, Kardaun J. (2010) The reliability of cause-of-death coding in The Netherlands. *European Journal of Epidemiology*; 25(8): 531-538.

Harvei S, Kravdal O. (1997) The importance of marital and socioeconomic status in incidence and survival of prostate cancer. An analysis of complete Norwegian birth cohorts. *Preventive Medicine*; 26: 623–632.

Harwood J, Sparks L. (2003) Social identity and health: An intergroup communication approach to cancer. *Health Communication*; 15(2): 145-159.

Hastert TA, Beresford SA, Sheppard L, White E. (2015) Disparities in cancer incidence and mortality by area-level socioeconomic status: a multilevel analysis. *Journal of Epidemiology and Community Health*; 69(2): 168-176.

Hastert TA, Ruterbusch JJ, Beresford SA, Sheppard L, White E. (2016) Contribution of health behaviors to the association between area-level socioeconomic status and cancer mortality. *Social Science and Medicine*; 148: 52-58.

Helms RL, O'Hea EL, Corso M. (2008) Body image issues in women with breast cancer. *Psychology, Health and Medicine*; 13(3): 313-325.

Herbert C, Lefevre H, Gignoux M, Launoy G. (2002) Influence of social and occupational class and area of residence on management and survival in patients with digestive tract cancer: a population study in the Calvados area (France). *Revue d'Epidémiologie et de Santé Publique*; 50: 253–264.

Hewitt M, Rowland JH, Yancik R. (2003) Cancer survivors in the United States: Age, health, and disability. *Journals of Gerontology Series A: Biological Sciences & Medical Sciences*; 58: 82-91.

Hiatt RA. (2006) The future of cancer surveillance. *Cancer Causes Control*; 17: 639–646.

Hiatt RA, Breen N.(2008) The Social Determinants of Cancer: A Challenge for Transdisciplinary Science. *American Journal of Preventive Medicine*; 35(2S): S141-S150.

Hill AB. (1937) *Principles of medical statistics*. London, The Lancet.

Hiscock R, Bauld L, Amos A, Fidler JA, Munafò M. (2012) Socioeconomic status and smoking: a review. *Annals of the New York Academy of Sciences*; 1248: 107-123.

Hislop TG, Waxler NE, Coldman AJ, Elwood MJ, Kan L. (1987) The prognostic significance of psychosocial factors in women with breast cancer. *Journal of Chronic Diseases*; 40: 729-735.

Ho SC. (1991) Health and social predictors of mortality in an elderly Chinese cohort. *American Journal of Epidemiology*; 133: 907-920.

Hoffman MA, Lent RW, Raque-Bogdan TL. (2013) A Social Cognitive Perspective on Coping With Cancer: Theory, Research, and Intervention. *The Counseling Psychologist*; 41(2): 240-267.

Hofman M, Ryan JL, Figueroa-Moseley CD, Jean-Pierre P, Morrow GR. (2007) Cancer-related fatigue: the scale of the problem. *Oncologist*; 12 Suppl. 1: 4-10.

Hole DJ, McArdle CS. (2002) Impact of socioeconomic deprivation on outcome after surgery for colorectal cancer. *British Journal of Surgery*; 89(5): 586-590.

Holland JC, Alici Y. (2010) Management of distress in cancer patients. *Journal of Supportive Oncology*; 8(1): 4-12.

Holmes FF, Hearne E 3rd. (1981) Cancer stage-to-age relationship: Implications for cancer screening in the elderly. *Journal of the American Geriatrics Society*; 29(2): 55-57.

Hong TS, Clark JW, Haigis KM. (2012) Cancers of the colon and rectum: identical or fraternal twins? *Cancer Discovery*; 2(2):117-121.

Howe HL, Wu X, Ries LA, et al. (2006) Annual report to the nation on the status of cancer, 1975–2003, featuring cancer among U.S. Hispanic/Latino populations. *Cancer*; 107: 1711–1742.

(a) Hussain SK, Altieri A, Sundquist J, Hemminki K. (2008) Influence of education level on breast cancer risk and survival in Sweden between 1990 and 2004. *International Journal of Cancer*; 122: 165-169.

(b) Hussain SK, Lenner P, Sundquist J, Hemminki K. (2008) Influence of education level on cancer survival in Sweden. *Annals of Oncology*; 19(1): 156-162.

Iseppato I. (2009) Le relazioni con i medici. In: Cipolla C, Maturo T. (a cura di) (2009) *Con gli occhi del paziente. Esperienze di cura e vissuto dei malati oncologici. Una ricerca nazionale*. Milano, Franco Angeli.

ISTAT (2013) Health for All Database. Roma, ISTAT. (<http://www.istat.it/it/archivio/14562>).

ISTAT (2015) Health for All Database. Roma, ISTAT. (<http://www.istat.it/it/archivio/14562>).

ISTAT (2015). Tavole di mortalità della popolazione residente. Roma, ISTAT. <http://www.istat.it/it/archivio/114885>

Jarvis M, Wardle J. (2006) Social patterning of health behaviours: the case of cigarette smoking. In: Marmot M, Wilkinson R. (eds.) (2006) *Social determinants of health*. Oxford, Oxford University Press.

Jenkins SP, Lynn P, Jäckle A, Sala E. (2008) The Feasibility of Linking Household Survey and Administrative Record Data: New Evidence for Britain. *International Journal of Social Research Methodology*; 11(1): 29–43.

Jensen OM, Parkin DM, Mac Lennan R et al. (1991) *Cancer registration, principles and methods*. Lyon, IARC Scientific publications n. 95.

Jin RL, Shah CP, Sroboda TJ. (1996) The impact of unemployment on health: a review of the evidence. *Canadian Medical Association Journal*; 153: 529–540.

Jin X, Carlin BP. (2005) Multivariate parametric spatiotemporal models for county level breast cancer survival data. *Lifetime Data Analysis*; 11(1): 5-27.

Johansen C, Schou G, Soll-Johanning H, Mellempgaard A, Lyng E. (1996) Influence of marital status on survival from colon and rectal cancer in Denmark. *British Journal of Cancer*; 74(6): 985-988.

Kaffashian F, Godward S, Davies T, Solomon L, McCann J, Duffy SW. (2003) Socioeconomic effects on breast cancer survival: proportion attributable to stage and morphology. *British Journal of Cancer*; 89: 1693–1696.

Kaiser K. (2008) The meaning of the survivor identity for women with breast cancer. *Social Science & Medicine*; 67(1): 79-87.

Kanavos P, Mossialos E. (1999) International comparisons of health care expenditures: what we know and what we do not know. *Journal of Health Services Research & Policy*; 4: 122-126.

Kang SH, Bloom JR. (1993) Social support and cancer screening among older Black Americans. *Journal of National Cancer Institute*; 95: 737-742.

Kant AK, Glover C, Horm J, Schatzkin A, Harris TB. (1992) Does cancer survival differ for older patients? *Cancer*; 70: 2734-2740.

Kaplan GA, Wilson TW, Cohen RD, Kauhanen J, Wu M, Salonen JT. (1994) Social functioning and overall mortality: prospective evidence from the Kuopio ischemic heart disease risk factor study. *Epidemiology*; 5: 495–500.

Kaplan GA. (1996) People and places: contrasting perspectives on the association between social class and health. *International Journal of Health Services*; 26: 507–519.

Kaplan GA. (2001) Economic policy is health policy: findings from the study of income, socioeconomic status, and health. In: Auerbach JA, Krimgold BK. (eds.) (2001) *Income, socioeconomic status, and health: Exploring the relationships*. Washington DC, National Policy Association.

Kawachi I, Berkman L. (2000) Social cohesion, social capital and health. In: Berkman L, Kawachi I. (eds.) (2000) *Social epidemiology*. New York, Oxford University Press.

Kawachi I, Colditz G, Ascherio A. (1996) A prospective study of social networks in relation to total mortality and cardiovascular disease in men in the USA. *Journal of Epidemiology & Community Health*; 50: 245–251.

Kawachi I, Kennedy P, Glass R. (1999) Social capital and self-rated health: a contextual analysis. *American Journal of Public Health*; 89: 1187–1194.

Kawachi I, Kennedy P, Lochner K. (1997) Social capital, income inequality and mortality. *American Journal of Public Health*; 87: 1491–1499.

Kawachi I, Kennedy P. (1997) Health and social cohesion: why care about income inequality? *BMJ*; 314: 1037–1041.

Keirn W, Metter G. (1985) Survival of cancer patients by economic status in a free care setting. *Cancer*; 55: 1552-1555.

Kingsmore D, Hole D, Gillis C. (2004) Why does specialist treatment of breast cancer improve survival? The role of surgical management. *British Journal of Cancer*; 90: 1920–1925.

Kingsmore D, Ssemwogerere A, Hole D, Gillis C. (2003) Specialisation and breast cancer survival in the screening era. *British Journal of Cancer*; 88: 1708–1712.

Kissane D. (2013) Marriage is as protective as chemotherapy in cancer care. *Journal of Clinical Oncology*; 31(31): 3852–3853.

Kitagawa EM, Hauser PM. (1973) *Differential mortality in the United States: A study in socioeconomic epidemiology*. Cambridge MA, Harvard University Press.

Klassen AC, Smith KC. (2011) The enduring and evolving relationship between social class and breast cancer burden: a review of the literature. *Cancer Epidemiology*; 35(3): 217–234.

Knowles J. (1977) *Doing better and feeling worse*. New York, Norton.

Kogevinas M, Pearce N, Susser M, Boffetta P. (eds.) (1997) *Social inequalities and cancer*. Lyon, IARC Scientific Publications No.138.

Kravdal Ø. (2000) Social inequalities in cancer survival. *Population Studies*; 54: 1–18.

Krieger N, Rowley D, Hermann AA, Avery B, Philips MT. (1993) Racism, sexism and social class: implications for studies of health, disease and well-being. *American Journal of Preventive Medicine*; 9(6): 82–122.

Krieger N. (1993) Analyzing socioeconomic and racial/ethnic patterns in health and health care. *American Journal of Public Health*; 83: 1086–1087.

Krieger N. (1994) Epidemiology and the web of causation: has anyone seen the spider? *Social Science and Medicine*; 39: 887–903.

Krieger N. (2001) Socioeconomic data in cancer registries. *American Journal of Public Health*; January; 91(1): 156–157.

Krieger N. (2005) Defining and investigating social disparities in cancer: critical issues. *Cancer Causes Control*; 16: 5–14.

Krieger N. (2008) Proximal, distal, and the politics of causation: what's level got to do with it? *American Journal of Public Health*; 98: 221–230.

Kroenke CH, Kubzansky LD, Schernhammer ES, Holmes MD, Kawachi I. (2006) Social networks, social support and survival after breast cancer diagnosis. *Journal of Clinical Oncology*; 24: 1105–1111.

Kuper H, Marmot M. (2003) Job strain, job demands, decision latitude and risk of coronary heart disease within the Whitehall II study. *Journal of Epidemiology and Community Medicine*; 57: 147–153.

Laaksonen M, Rahkonen O, Karvonen S, Lahti E. (2005) Socioeconomic status and smoking. Analysing inequalities with multiple indicators. *European Journal of Public Health*; 15(3): 262-269.

Labonte R. (1993) *Health Promotion and Empowerment: Practice Frameworks*. Toronto, Centre for Health Promotion, University of Toronto. *Issues in Health Promotion* no. 3.

Lagerlund M, Bellocco R, Karlsson P, Tejler G, Lambe M. (2005) Socio-economic factors and breast cancer survival. A population-based cohort study (Sweden). *Cancer Causes and Control*; 16(4): 419-430.

Lambert P, Royston P. (2009) Further development of flexible parametric models for survival analysis. *The Stata Journal*; 9(2): 265-290.

Lambert PC, Smith LK, Jones DR, Botha JL. (2005) Additive and multiplicative covariate regression models for relative survival incorporating fractional polynomials for time-dependent effects. *Statistics in Medicine*; 24(24): 3871-3885.

Lamont DW, Symonds RP, Brodie MM, Nwabine NJ, Gillis CR. (1980) Age, socio-economic status and survival from cancer of cervix in the West of Scotland 1980-87. *British Journal of Cancer*; 67: 351-357.

Langenbach MR, Schmidt J, Neumann J, Zirngibl H. (2003) Delay in treatment of colorectal cancer: multifactorial problem. *World Journal of Surgery*; 27(3): 304-308.

Lavelle K, Todd C, Moran A, Howell A, Bundred N, Campbell M. (2007) Non-standard management of breast cancer increases with age in the UK: A population based cohort of women > 65 years. *British Journal of Cancer*; 96: 1197-1203.

Lazarus RS, Folkman S. (1991) The concept of coping. In: Monat A, Lazarus RS. (eds.) (1991) *Stress and coping: An anthology* (3rd ed.). New York, Columbia University Press.

Le H, Ziogas A, Lipkin SM, Zell JA. (2008) Effects of socioeconomic status and treatment disparities in colorectal cancer survival. *Cancer, Epidemiology, Biomarkers and Prevention*; 17(8): 1950-1962.

Lemmens VEPP, van Halteren AH, Janssen-Heijnen MLG, Vreugdenhil G, van Driel OJR, Coebergh JWW. (2005) Adjuvant treatment for elderly patients with stage III colon cancer in the southern Netherlands is affected by socioeconomic status, gender, and comorbidity. *Annals of Oncology*; 16: 767-772.

Levitz NR, Haji-Jama S, Munro T, Gorey KM, Luginaah IN, Bartfay E, Zou G, Wright FC, Kanjeekal SM, Hamm C, Balagurusamy MK, Holowaty EJ. (2015) Multiplicative

disadvantage of being an unmarried and inadequately insured woman living in poverty with colon cancer: historical cohort exploration in California. *BMC Women's Health*; 15: 8.

Li B, Quan H, Fong A, Lu M. (2006) Assessing record linkage between health care and Vital Statistics databases using deterministic methods. *BMC Health Services Research*; 6: 48.

Li Q, Gan L, Liang L, Li X, Cai S. (2015) The influence of marital status on stage at diagnosis and survival of patients with colorectal cancer. *Oncotarget*; 6(9): 7339-7347.

Liberatos P, Link BG, Kensey JL (1988) The measurements of social class in epidemiology. *Epidemiologic Reviews*; 10: 87-121.

Lieberman S. (1985) *Making it count: The improvement of social research and theory*. Berkeley, University of California Press.

Lillini R, Vercelli M, Quaglia A, Micheli A, Capocaccia R. (2011) Use of socio-economic factors and healthcare resources to estimate cancer survival in European countries with partial national cancer registration. *Tumori*; 97(3): 265-274.

Linden G. (1969) The influence of social class in the survival of cancer patients. *American Journal of Public Health*; 59: 267-274.

Lindop E, Cannon S. (2001) Evaluating the self-assessed support needs of women with breast cancer. *Journal of Advanced Nursing*; 34(6): 760-771.

Lindwall L, Bergbom I. (2009) The altered body after breast cancer surgery. *International Journal of Qualitative Studies on Health and Well-being*; 4: 280-287.

Link BG, Northridge ME, Phelan JC, Ganz ML. (1998) Social epidemiology and the fundamental cause concept: On the structuring of effective cancer screens by socioeconomic status. *Milbank Quarterly*; 76: 375.

Liu L, Cozen W, Bernstein L, Ross RK, Deapen D. (2001) Changing relationship between socioeconomic status and prostate cancer incidence. *Journal of National Cancer Institute*; 93(9): 705-709.

Liu L, Huang X. (2008) The use of gaussian quadrature for estimation in frailty proportional hazards models. *Statistics in Medicine*; 27(14): 2665-2683.

Locker D. (2008) Social determinant of health and disease. In: Scambler G (2008) *Sociology as applied to medicine - 6th Edition*. London, Elsevier.

Lucchini M, Sarti S, Tognetti M. (2009) I welfare regionali e le differenze territoriali nelle disuguaglianze di salute. In: Brandolini A, Saraceno C, Schizzerotto A. (eds.) (2009) *Dimensioni della disuguaglianza in Italia: povertà, salute, abitazione*. Bologna, Il Mulino.

- Lutz W, Sanderson W, Scherbov S. (2008) The coming acceleration of global population ageing. *Nature*; 451: 716-719.
- Lynch JW, Kaplan GA, Salonen GT. (1997) Why do poor people behave poorly? Variation in adult health behaviours and psychosocial characteristics by stages of the socioeconomic life course. *Social Science & Medicine*; 44(6): 809-819.
- Lyyra TM, Heikkinen RL. (2006) Perceived social support and mortality in older people. *Journal of Gerontology*; 61: S147–S152.
- Machado CJ. (2004) A literature review of record linkage procedures focusing on infant health outcomes. *Cad. Saúde Pública*; 20(2): 362-371.
- MacIntyre S, MacIver S, Soomans A. (1993) Area, class and health: should we be focusing on places or people? *Journal of Social Policy*; 22: 213–234.
- Mackenbach JP, Looman CWN. (1994) Living standards and mortality in the European Community. *Journal of Epidemiology & Community Health*; 48: 140-145.
- MacLeod MC, Bray CA, Kendrick SW, Cobbe SM. (1998) Enhancing the power of record linkage involving low quality personal identifiers: use of the best link principle and cause of death prior likelihoods. *Computing Biomedical Research*; 31: 257-270.
- Macleod U, Ross S, Fallowfield L, Watt GC. (2004) Anxiety and support in breast cancer: Is this different for affluent and deprived women? A questionnaire study. *British Journal of Cancer*; 91: 879–883.
- Macleod U, Ross S, Gillis C, McConnachie A, Twelves C, Watt GC. (2000) Socio-economic deprivation and stage of disease at presentation in women with breast cancer. *Annals of Oncology*; 11: 105–107.
- Marchi M. (2003) Protezione dei Dati Personali e Ricerca Biomedica. *Statistica*; anno LXIII, n. 4: 685-692.
- Marinacci C, Ferracin E, Landriscina T, Cislighi C, Gargiulo L, Costa G. (2010) Differenze geografiche o differenze sociali. *Rapporto OsservaSalute 2010*: 473-484.
- Marinacci C, Grippo F, Pappagallo M, Sebastiani G, Demaria M, Vittori P, Caranci N, Costa G. (2013) Social inequalities in total and cause-specific mortality of a sample of the Italian population, from 1999 to 2007. *European Journal of Public Health*; 23(4): 582-587.
- Marmot M, Kogevinas M, Elston MA. (1987) Social/economic status and disease. *Annual Review of Public Health*; 8: 111-135.

- Marmot M, Madge N. (1987) An epidemiological perspective on stress and health. In: Kasl S, Cooper C. (eds.) (1987) *Stress and health: issues in research methodology*. Winchester, Wiley.
- Marmot M, Rose G, Shipley M, Hamilton PJS. (1978) Employment grade and coronary heart disease in British civil servants. *Journal of Epidemiology and Community Health*; 32: 244-249.
- Marmot M, Siegrist J, Theorell T, Feeny A. (2006) Health and the psychosocial environment at work. In: Marmot R, Wilkinson R. (eds.) (2006) *Social determinants of health*. Oxford, Oxford University Press.
- Marmot M, Syme L, Kagan A. (1975) Epidemiological studies of heart disease and stroke in Japanese men living in Japan, Hawaii and California. Prevalence of coronary and hypertensive disease and associated risk factors. *American Journal of Epidemiology*; 102: 514–525.
- Marmot M, Theorell T. (1988) Social class and cardiovascular disease: the contribution of work. *International Journal of Health Services*; 18: 37–45.
- Marmot M, Wilkinson RG. (eds.) (2006) *Social Determinants of Health*. New York, Oxford University Press.
- Marmot M. (1991) Health inequalities among British civil servants in the Whitehall II Study. *Lancet*; 337: 1387–1393.
- Marmot M. (1998) Improvement of social environment to improve health. *The Lancet*; 351: 57–60.
- Martin-Moreno JM, Anttila A, von Karsa L, Alfonso-Sanchez JL, Gorgojo L. (2012) Cancer screening and health system resilience: keys to protecting and bolstering preventive services during a financial crisis. *European Journal of Cancer*; 48(14): 2212-2218.
- Masi CM, Olopade OI. (2010) Racial and ethnic disparities in breast cancer: a multilevel perspective. *Medical Clinics of North America*; 89: 753–770.
- Maturo A, Agnoletti V. (2006) Il rapporto col paziente competente (e confuso). In: Cipolla C, Corposanto C, Tousijn W. (eds.) (2006) *I medici di medicina generale in Italia*. Milano, Franco Angeli.
- Maturo A. (2007) *Sociologia della malattia. Un'introduzione*. Milano, Franco Angeli.
- Maturo A. (2009) L'informazione sulla malattia. In: Cipolla C, Maturo A. (a cura di) (2009) *Con gli occhi del paziente. Una ricerca nazionale sui vissuti di cura dei malati oncologici*. Milano, Franco Angeli.

- McGinnis JM, Foege WH. (1993) Actual causes of death in the United States. *Journal of American Medical Association*; 270(18): 2207-2212.
- McGinnis LS, Menck HR, Eyre HJ, Bland KI, Scott-Conner CE, Morrow M, Winchester DP. (2000) National Cancer Data Base survey of breast cancer management for patients from low income zip codes. *Cancer*; 88(4): 933-945.
- McKee-Ryan F, Song Z, Wanberg C, Kinicki A. (2005) Psychological and physical well-being during unemployment: A meta-analytic study. *Journal of Applied Psychology*; 90: 53–76.
- McKenzie F, Ellison-Loschmann L, Jeffreys M. (2011) Investigating reasons for ethnic inequalities in breast cancer survival in New Zealand. *Ethnicity and Health*; 16(6): 535-549.
- McKinlay JB. (1979) A case for refocusing upstream: the political economy of illness. In: Jaco EG. (ed.) (1979) *Patients, physicians, and illness: a sourcebook in behavioral science and health*. 3rd ed. New York, Free Press.
- McLeroy KR, Bibeau D, Steckler A, Glanz K. (1988) An ecological perspective on health promotion programs. *Health Education Quarterly*; 15: 351–377.
- McTiernan A. (2003) Behavioral risk factors in breast cancer: can risk be modified? *Oncologist*; 8(4): 326-334.
- Micheli A, Coebergh JW, Mugno E, Massimiliani E, Sant M, Oberaigner W, Holub J, Storm HH, Forman D, Quinn M, Aareleid T, Sankila R, Hakulinen T, Faivre J, Ziegler H, Tryggvadóttir L, Zanetti R, Dalmas M, Visser O, Langmark F, Bielska-Lasota M, Wronkowski Z, Pinheiro PS, Brewster DH, Plesko I, Pompe-Kirn V, Martinez-Garcia C, Barlow L, Möller T, Lutz JM, André M, Steward JA. (2003) European health systems and cancer care. *Annals of Oncology*; 5: 41–60.
- Micheli A, Mariotto A, Giorgi Rossi A, Gatta G, Muti P, the EUROCORE Working Group. (1998) The prognostic role of gender in survival of adult cancer patients. *European Journal of Cancer*; 34: 2271–2278.
- Micheli A, Mariotto A, Giorgi Rossi A, Gatta G, Muti P, the EUROCORE Working Group. (1998) The prognostic role of gender in survival of adult cancer patients. *European Journal of Cancer*; 34: 2271–2278.
- Micheli A, Mugno E, Krogh V, Quinn MJ, Coleman M, Hakulinen T, Gatta G, Berrino F, Capocaccia R, EUROPREVAL Working Group. (2002). Cancer prevalence in European registry areas. *Annals of Oncology*; 13: 840-865.
- Midthune D, Fay M, Clegg L, Feuer E. (2005) Modeling reporting delay and reporting corrections in cancer registry data. *Journal of American Statistical Association*; 100: 61–70.

- Miller BA, Hankey BF, Thomas TL. (2002) Impact of sociodemographic factors, hormone receptor status, and tumor grade on ethnic differences in tumor stage and size for breast cancer in US women. *American Journal of Epidemiology*; 155: 534–545.
- Milner PC, Watts M. (1987) Effect of socioeconomic status on survival from cervical cancer in Sheffield. *Journal of Epidemiology & Community Health*; 41: 200-203.
- Ministero della Sanità. (2001) ICD-10: classificazione statistica internazionale delle malattie e dei problemi sanitari correlati. 10^a revisione. Roma, Istituto poligrafico e Zecca dello Stato.
- Mobley LR, Kuo TM. (2015) United States Health Policies and Late-stage Breast and colorectal cancer diagnosis: Why such disparities by age? *Health Economics Review*; 5(1): 58.
- Montella M, Crispo A, D'Aiuto G, De Marco M, de Bellis G, Fabbrocini G, Pizzorusso M, Tamburini M, Silvestra P. (2001) Determinant factors for diagnostic delay in operable breast cancer patients. *European Journal of Cancer Prevention*; 10(1): 53-59.
- Montgomery S, Cook D, Bartley M, Wadsworth M. (1999) Unemployment in young men predates symptoms of depression and anxiety resulting in medical consultation. *International Journal of Epidemiology*; 28: 95–100.
- Mor V, Masterson-Allen S, Houts P, Siegel K. (1992) The changing needs of patients with cancer at home: a longitudinal view. *Cancer*; 69: 829-838.
- Moran A, Sowerbutts AM, Collins S, Clarke N, Cowan R. (2004) Bladder cancer: Worse survival in women from deprived areas. *British Journal of Cancer*; 90: 2142–2144.
- Morgan M. (1980) Marital status, health, illness and service use. *Social Science & Medicine*; 14: 633–643.
- Moussa MA, Shafie MZ, Khogali MM, el-Sayed AM, Sugathan TN, Cherian G, Abdel-Khalik AZ, Garada MT, Verma D. (1990) Reliability of death certificate diagnoses. *Journal of Clinical Epidemiology*; 43(12): 1285-1295.
- Munro AJ, Bentley AH. (2004) Deprivation, comorbidity and survival in a cohort of patients with colorectal cancer. *European Journal of Cancer Care*; 13(3): 254-262.
- Musgrove P, Zeramdini R, Carrin G. (2002) Basic patterns in national health expenditure. *Bull World Health Organization*; 80: 134–142.
- Najman J. (1980) Theories of disease causation and the concept of general susceptibility: a review. *Social Science & Medicine*; 14(3): 231-237.
- National Cancer Institute, Division of Cancer Control and Population Sciences. (2007) Cancer control continuum. (www.cancercontrol.cancer.gov/od/continuum.html)

- Nelson CP, Lambert PC, Squire IB, Jones DR. (2007). Relative survival: What can cardiovascular disease learn from cancer? *Statistics in Medicine*; 26: 5486-5498.
- Nettleton S, Burrows B. (1998) Mortgage debt, insecure home ownership and health: An exploratory analysis. *Social Science & Illness*; 20: 731–753.
- Newcombe HB, Kennedy JM, Axford SJ, James AP (1959) Automatic linkage of vital records. *Science*, 954-959.
- Newcombe HB. (1988) *Handbook of record linkage: methods for health and statistical studies, administration and business*. Oxford, Oxford University Press.
- Nicholson A, Fuhrer R, Marmot M. (2005) Psychological distress as a predictor of CHD events in men: the effect of persistence and components of risk. *Psychosomatic Medicine*; 67: 522–530.
- Nielsen GP, Björnsson J, Jonasson JG. (1991) The accuracy of death certificates. Implications for health statistics. *Virchows Archive A*; 419(2): 143-146.
- OECD. (2015), *OECD Reviews of Health Care Quality: Italy 2014: Raising Standards*. Parigi, OECD Publishing.
- Ogden J, Lindridge L. (2008) The impact of breast scarring on perceptions of attractiveness: an experimental study. *The impact of breast scarring on perceptions of attractiveness: an experimental study. Journal of Health Psychology*; 13(3): 303-310.
- O'Malley CD, Le GM, Glaser SL, Shema SJ, West DW. (2003) Socioeconomic status and breast carcinoma survival in four racial/ethnic groups: A population-based study. *Cancer*; 97: 1303–1311.
- Osborne C, Ostir GV, Du X, Peek MK, Goodwin JS. (2005) The influence of marital status on the stage at diagnosis, treatment, and survival of older women with breast cancer. *Breast Cancer Research and Treatment*; 93(1): 41-47.
- Osservatorio Nazionale sulla Salute della Donna (ONDA) (2014) *Donne e tumori. Aspetti clinici, psicopatologici e terapeutico-assistenziali*. Milano, Franco Angeli.
- Osservatorio sulla condizione assistenziale dei malati oncologici. (2010) *Secondo Rapporto sulla condizione assistenziale dei malati oncologici*. Roma, Ministero della salute.
- Ostberg V, Lennartson C. (2007) Getting by with a little help: the importance of various types of social support for health problems. *Scandinavian Journal of Public Health*; 35: 197–204.
- Paci E, Puliti D. (eds.) (2011) *Come cambia l'epidemiologia del tumore della mammella in Italia. I risultati del progetto IMPATTO dei programmi di screening mammografico*. Pisa, Pacini Editore Medicina.

Page WF, Kuntz AJ. (1980) Racial and socioeconomic factors in cancer survival. *Cancer*; 45: 1029-1040.

Panizon F. (ed.) (2009) *La qualità dei dati. Conoscere il Censimento. 14° Censimento generale della popolazione e delle abitazioni.* Roma, ISTAT.

Parikh-Patel A, Bates JH, Campleman S. (2006) Colorectal cancer stage at diagnosis by socioeconomic and urban/rural status in California, 1988-2000. *Cancer*; 107(5 Suppl): 1189-1195.

Parkin DM, Bray F. (2009) Evaluation of data quality in the cancer registry: principles and methods Part II. Completeness. *European Journal of Cancer*; 45(5): 756-764.

Parkin DM, Chen VW, Ferlay J et al. (1995) Comparability and quality control in cancer registration. Lyon, IARC Technical report n. 19.

Parkin DM, Whelan SL, Ferlay J, Storm H. (eds). (2005) *Cancer incidence in five continents Vol. I-VIII.* Lyon, IARC.

Parkin DM. (2006) The evolution of the population-based cancer registry. *Nature review. Cancer*; 6: 603-612.

Pasick RJ, Hiatt RA, Paskett ED. (2004) Lessons learned from community-based cancer screening intervention research. *Cancer*; 101: 1146–1164.

Paterson IC, John G, Adams JD. (2002) Effect of deprivation on survival of patients with head and neck cancer: A study of 20,131 cases. *Clinical Oncology*; 14: 455–458.

Patrick D, Morgan M, Charlton J. (1986) Psychosocial support and change in the health status of physically disabled people. *Social Science & Medicine*; 22: 1347–1354.

Patterson J, Eberly L, Ding Y, Hargreaves M. (2004) Associations of smoking prevalence with individual and area level social cohesion. *Journal of Epidemiology and Community Health*; 58: 692–697.

Percy CL, Miller BA, Gloeckler Ries LA. (1990) Effect of changes in cancer classification and the accuracy of cancer death certificates on trends in cancer mortality. *Annals of the New York Academy of Sciences*; 609: 87-97.

Percy CL, Stanek E, Gloeckler L. (1981) Accuracy of cancer death certificates and its effect on cancer mortality statistics. *American Journal of Public Health*; 71: 242–250.

Perreault A, Bourbonnais FF. (2005) The experience of suffering as lived by women with breast cancer. *International Journal of Palliative Nursing*; 11: 512-519.

Phelan JC, Link BG, Diez-Roux A, Kawachi I, Levin B. (2004) “Fundamental causes” of social inequalities in mortality: A test of the theory. *Journal of Health Social Behavior*; 45: 265–285.

Phelan JC, Link BG, Tehranifar P. (2010) Social conditions as fundamental causes of health inequalities: Theories, evidences and policy implications. *Journal of Health and Social Behaviour*; 51(5): 528-540.

Pinelli A. (1984) The record linkage in the study of infant mortality: some aspects concerning data quality. *Statistica*, 44: 675-686.

Pinquart M, Hoffken K, Silbereisen R, Wedding U. (2007) Social support and survival in patients with acute myeloid leukaemia. *Care in Cancer*; 15: 81–87.

Polednak AP. (2001) Poverty, comorbidity, and survival of colorectal cancer patients diagnosed in Connecticut. *Journal of Health Care for the Poor and Underserved*; 12(3): 302-310.

Polednak AP. (2002) Survival of breast cancer patients in Connecticut in relation to socio-economic and health care access indicators. *Journal of Urban Health*; 79(2): 211-218.

Pollock AM, Vickers N. (1998) Deprivation and emergency admissions for cancers of colorectum, lung, and breast in south east England: Ecological study. *BMJ*; 317: 245–252.

Potosky AL, Merrill RM, Riley GF, Taplin SH, Barlow W, Fireman BH, Ballard-Barbash R. (1997) Breast cancer survival and treatment in health maintenance organization and fee-for-service settings. *Journal of National Cancer Institute*; 89: 1683–1691.

Quaglia A, Capocaccia R, Micheli A, Carrani E, Vercelli M, and the EURO CARE-3 Working Group. (2007) A wide difference in cancer survival between middle aged and elderly patients in Europe. *International Journal of Cancer*; 120: 2196-2201.

Quaglia A, Lillini R, Casella C, Giachero G, Izzotti A, Vercelli M; Liguria Region Tumour Registry. (2011) The combined effect of age and socio-economic status on breast cancer survival. *Critical Reviews in Oncology/Hematology*; 77(3): 210-220.

Quaglia A, Vercelli M, Lillini R, Mugno E, Coebergh JW, Quinn M, Martinez-Garcia C, Capocaccia R, Micheli A, and the ELDCARE Working Group. (2005) Socio-economic factors and health care system characteristics related to cancer survival in the elderly. A population-based analysis in sixteen European countries. *Critical Reviews in Oncology/Hematology*; 54: 117-128.

Rabe-Hesketh S, Skrondal A, Pickles A. (2002) Reliable estimation of generalized linear mixed models using adaptive quadrature. *The Stata Journal*; 2: 1–21.

- Randal J. (1976) The social origins of cancer. *Change*; 8(10): 54-55.
- Raphael D. (2006) Social determinants of health: present status, unanswered questions and future directions. *International Journal of Health Services*; 36: 651–677.
- Raque-Bogdan TL, Torrey CL, Lewis BL, Borges NJ. (2013) Counseling health psychology: Assessing health psychology training within counseling psychology doctoral programs. *The Counseling Psychologist*; 41(3): 428-452.
- Regidor E, Barrio G, Bravo MJ, de la Fuente L. (2014) Has health in Spain been declining since the economic crisis? *Journal of Epidemiology and Community Health*; 68(3): 280-282.
- Reine I, Novo M, Mammarstrom A. (2004) Does the association between health and unemployment differ between young people and adults? Results from a 14 year follow-up study with a focus on psychological health and smoking. *Public Health*; 118: 337–345.
- Remontet L, Bossard N, Belot A, Estève J, French Network of Cancer Registries (FRANCIM). (2007) An overall strategy based on regression models to estimate relative survival and model the effects of prognostic factors in cancer survival studies. *Statistics in Medicine* 26: 2214–2228.
- Ren XS. (1997) Marital status and quality of relationships: the impact on health perception. *Social Science and Medicine*; 44(2): 241-249.
- Reyes-Ortiz CA, Goodwin JS, Freeman JL, Kuo YF. (2006) Socioeconomic status and survival in older patients with melanoma. *Journal of American Geriatric Society*; 54(11): 1758-1764.
- Reynolds P, Kaplan GA (1990) Social connections and risk for cancer: Prospective evidence from the Alameda County study. *Behavioral Medicine*; 16: 101-110.
- Robert SA, Strombom I, Trentham-Dietz A, Hampton JM, McElroy JA, Newcomb PA, Remington PL. (2004) Socioeconomic risk factors for breast cancer: distinguishing individual- and community-level effects. *Epidemiology*; 15(4): 442-450.
- Roberts CS. (1998) Social support and the elderly cancer patient. In: Balducci L, Lyman GH, Ershler WB. (eds.) (1998) *Comprehensive Geriatric Oncology*. Amsterdam, Harwood Academic Publishers.
- Roetzheim RG, Pal N, Gonzalez EC, Ferrante JM, Van Durme DJ, Krischer JP. (2000) Effects of health insurance and race on colorectal cancer treatments and outcomes. *American Journal of Public Health*; 90: 1746–1754.
- Rosenfield PL. (1992) The potential of transdisciplinary research for sustaining and extending linkages between the health and social sciences. *Social Science & Medicine*; 35: 1343–1357.

Ross C, Mirowsky J. (2010) Why Education is the Key to Socioeconomic Differentials in Health. In: Bird C, Conrad P, Fremont A, Timmermans S. (a cura di) Handbook of Medical Sociology. Sixth Edition. Nashville, Vanderbilt U.P.

Rossi S, Crocetti E, Capocaccia R, Gatta G; AIRTUM Working Group. (2013) Estimates of cancer burden in Italy. *Tumori*; 99(3): 416-424.

Rowland JH. (2008) What are cancer survivors telling us? *The Cancer Journal*; 14; 361-368.

Rutqvist LE, Bern A, Stockholm Breast Cancer Study Group. (2006) Socioeconomic gradients in clinical stage at presentation and survival among breast cancer patients in the Stockholm area 1977-1997. *International Journal of Cancer*; 119(6): 1433-1439.

Ryan MC, Austin AG. (1989) Social support and social networks in the aged. *Image*; 21: 176-180.

Sala E, Knies G, Burton J. (2014) Propensity to consent to data linkage: experimental evidence on the role of three survey design features in a UK longitudinal panel. *International Journal of Social Research Methodology*; 17(5): 455-473.

Sala E. (2013) Le indagini campionarie del nuovo millennio. Quali novità all'orizzonte? *Quaderni di Sociologia*; vol. LVII, 62: 77-90.

Sant M, Francisci S, Capocaccia R, Verdecchia A, Allemani C, Berrino F. (2006) Time trends of breast cancer survival in Europe in relation to incidence and mortality. *International Journal of Cancer*; 119: 2417-2422.

Scabini E, Donati P. (a cura di) (2003) Tempo e transizioni familiari. *Studi Interdisciplinari sulla Famiglia*; 13.

Scabini E, Rossi G (a cura di) (2006) Le parole della famiglia. *Studi Interdisciplinari sulla Famiglia*; 21.

Scanu M. (2003) *Metodi Statistici per il Record Linkage*. Roma, ISTAT.

Scheuren F. (1999) Linking health records: human rights concerns. In: *Proceedings of an International Workshop and Exposition: Record Linkage Techniques*; 1997 March 20-21; Arlington, United States. Washington DC, National Academy Press.

Schillinger JA1, Grosclaude PC, Honjo S, Quinn MJ, Sloggett A, Coleman MP. (1999) Survival after acute lymphocytic leukaemia: effects of socioeconomic status and geographic region. *Archives of Disease in Childhood*; 80: 311-317.

Schootman M, Jeffe DB, Baker EA, Walker MS. (2006) Effect of area poverty rate on cancer screening across US communities. *Journal of Epidemiology & Community Health*; 60(3): 202-207.

Schrijvers CT, Coebergh JW, Mackenbach JP. (1997) Socioeconomic status and comorbidity among newly diagnosed cancer patients. *Cancer*; 80: 1482–1488.

Schrijvers CT, Mackenbach JP, Lutz JM, Quinn MJ, Coleman MP. (1995) Deprivation and survival from breast cancer. *British Journal of Cancer*; 72: 738–743.

Schrijvers CTM, Mackenbach JP. (1994) Cancer patient survival by socioeconomic status in the Netherlands: A review for six common cancer sites. *Journal of Epidemiology & Community Health*; 48: 441-446.

Schwartz KL, Crossley-May H, Vigneau FD, Brown K, Banerjee M. (2003) Race, socioeconomic status and stage at diagnosis for five common malignancies. *Cancer Causes Control* ; 14(8): 761-766.

Siegrist J, Peter R, Junge A, Cremer P, Seidel D. (1990) Low status control, high effort at work and ischemic heart disease: Prospective evidence from blue collar men. *Social Science & Medicine*; 31: 1127–1134.

Simnet I, Perkins E, Wright L. (1999) Evidence-based health promotion. New York, Chichester Wiley.

Singh GK, Miller BA, Hankey BF, Edwards BK. (2003) Area socioeconomic variations in U.S. cancer incidence, mortality, stage, treatment, and survival, 1975–1999. Bethesda, National Cancer Institute.

Singh GK, Siahpush M. (2001) All-cause and cause-specific mortality of immigrants and native born in the U.S. *American Journal of Public Health*; 91: 392–399.

Singh SM, Paszat LF, Li C, He J, Vinden C, Rabeneck L. (2004) Association of socioeconomic status and receipt of colorectal cancer investigations: A population-based retrospective cohort study. *CMAJ*; 171(5): 461-465.

Smedley BD, Syme SL, Committee on Capitalizing on Social Science and Behavioral Research to Improve the Public's Health. (2001) Promoting health: intervention strategies from social and behavioral research. *American Journal of Health Promotion*; 15: 149–166.

Smith GD, Neaton JD, Wentworth D, Stamler R, Stamler J. (1996) Socioeconomic differentials in mortality risk among men screened for the Multiple Risk Factor Intervention Trial: I. White men. *American Journal of Public Health*; 86: 486–496.

Smith ME. (1984) Record linkage: present status and methodology. *Journal of Clinical Monitoring and Computing*; 13: 52-71.

Smith-Gagen J, Cress RD, Drake CM, Felter MC, Beaumont JJ. (2005) Factors associated with time to availability for cases reported to population-based cancer registries. *Cancer Causes Control*; 16: 449–454.

Snow J. (1849) On the pathology and modes of communication of cholera. *London Medical Gazette*; 44: 745-752.

Stafford M, Marmot M. (2003) Neighborhood deprivation and health: Does it affect us all equally? *International Journal of Epidemiology*; 32: 357–366.

Stansfield S. (2006) Social support and social cohesion. In: Marmot M, Wilkinson R. (eds.) (2006) *Social determinants of health*. Oxford, Oxford University Press.

STATA. (2014) Package `stmixed` v1.1.0. Boston, Stata Corp. <http://www.stata.com/support/ssc-installation/>

StataCorp (2014). *Statistical Software: Release 13.0*. College Station, TX: Stata Corporation.

Stavraky KM, Kincade JE, Stewart MA, Donner AP. (1987) The effect of socioeconomic factors on the early prognosis of cancer. *Journal of Chronic Diseases*; 40: 237-244.

Stephoe A, Sigerist J, Kirschbaum C, Marmot M. (2004) Effort-reward imbalance, over commitment and measures of cortisol and blood pressure over the working day. *Psychosomatic Medicine*; 66: 323–329.

Stracci F, Bianconi F, La Rosa F, Tonato M. (2010) Cancer registry integration in the Umbria Regional Cancer Network: breast cancer as a model to build a supportive information system. *Annals of Oncology*; 21 (Suppl. 8): 339-340.

Suarez L, Lloyd L, Weiss N, Rainbolt T, Pulley L. (1994) Effect of social networks on cancer screening behavior of older Mexican-American women. *Journal of National Cancer Institute*; 86: 775-779.

Sugiura N. (1978) Further analysis of the data by Akaike's information criterion and the finite corrections. *Communications in Statistics - Theory and Methods*; A7: 13–26.

Supramaniam R, Smith D, Coates M. (1998) Breast cancer survival in NSW in 1973 to 1995. Sydney, Cancer Epidemiology Research Unit and NSW Central Cancer Registry Cancer Control Information Centre, NSW Cancer Council.

Syme S, Hyman M, Enterline P. (1964) Some social and cultural factors associated with the occurrence of coronary heart disease. *Journal of Chronic Diseases*; 17: 277–289.

Syme S. (1986) Social determinants of health and disease. In: Last J. (ed.) (1986) *Public health and preventative medicine*. Norwalk CT, Appleton-Century-Crofts.

Tamas K, Walenkamp AM, de Vries EG, van Vugt MA, Beets-Tan RG, van Etten B, de Groot DJ, Hospers GA. (2015) Rectal and colon cancer: Not just a different anatomic site. *Cancer Treatment Reviews*; 41(8): 671-679.

Teppo L. (1984) Cancer incidence by living area, social class and occupation, *Scandinavian Journal of Work and Environmental Health*; 10: 361-366.

Terracini B, Zanetti R. (2004) A short history of pathology registries, with emphasis on cancer registries. In Morabia A. (ed.) (2004) *A history of epidemiologic methods and concepts*. Basel, Birkhauser Verlag.

Thomas C, Benzeval M, Stansfeld S. (2005) Employment transitions and mental health: an analysis from the British household panel survey. *Journal of Epidemiology & Community Health*; 59: 243–249.

Thong MS, Kaptein AA, Krediet RT, Boeschoten EW, Dekker FW. (2007) Social support predicts survival in dialysis patients. *Nephrology Dialysis Transplantation*; 22: 845–850.

Timmons A, Gooberman-Hill R, Sharp L. (2013) The multidimensional nature of the financial and economic burden of a cancer diagnosis on patients and their families: qualitative findings from a country with a mixed public-private healthcare system. *Supportive Care in Cancer*; 21(1): 107-117.

Tomatis L. (2001) Inequalities in cancer risks. *Seminars in Oncology*; 28: 207–209.

Tresserras R, Canela J, Alvarez J, Sentis J, Salleras L. (1992) Infant mortality, per capita income and adult illiteracy: An ecological approach. *American Journal of Public Health*; 82: 435–438.

Turner JB. (1995) Economic context and the impact of unemployment. *Journal of Health and Social Behaviour*; 35: 213–219.

Van Den Brandt PA, Schouten LJ, Goldbohm RA, Dorant E, Hunen PMH. (1990) Development of a record linkage protocol for use in the Dutch Cancer Registry for epidemiological research. *International Journal of Epidemiology*; 19: 553-558.

Vercelli M, Capocaccia R, Quaglia A, Casella C, Puppo A, Coebergh JWW, The EURO CARE Working Group. (2000) Relative survival in elderly European cancer patients: Evidence for health care inequalities. *Critical Reviews in Oncology/Hematology*; 35: 161-179.

Vercelli M, Lillini R, Capocaccia R, Micheli A, Coebergh JW, Quinn M, Martinez-Garcia C, Quaglia A, and the ELDCARE Working group. (2006) Cancer survival in the elderly: Effects of socio-economic factors and health care system features. *European Journal of Cancer*; 42: 234-242.

Vercelli M, Lillini R, Capocaccia R, Quaglia A; Socio-Economic Indicators and Health Working Group; AIRTUM Contributors. (2012) Use of SERTS (Socio-Economic, health Resources and Technologic Supplies) models to estimate cancer survival at provincial geographical level. *Cancer Epidemiology* 36(6): 566-574.

- Vercelli M, Lillini R, Quaglia A, Capocaccia R; SEIH (Socio-Economic Indicators and Health) Working Group & AIRTUM contributors. (2014) Italian regional health system structure and expected cancer survival. *Tumori*; 100(4): 386-398.
- Vercelli M, Lillini R, Quaglia A, Micale RT, La Maestra S, De Flora S. (2014) Age-related mortality trends in Italy from 1901 to 2008. *PLoS One*; 9(12): e114027.
- Vercelli M, Quaglia A, Casella C, Parodi S, Capocaccia R, Martinez Garcia C, The EURO CARE Working Group. (1998) Relative survival in elderly cancer patients in Europe. *European Journal of Cancer*; 34: 2264-2270.
- Vercelli M, Quaglia A, Lillini R. (2013) Useful indicators to interpret the cancer burden in Italy. *Tumori*; 99(3): 425-438.
- Verdecchia A, Baili P, Quaglia A, Kunkler I, Ciampichini R, Berrino F, Micheli A. (2008) Patient survival for all cancers combined as indicator of cancer control in Europe. *European Journal of Public Health*; 18(5): 527-532.
- Villingshøj M, Ross L, Thomsen BL, Johansen C. (2006) Does marital status and altered contact with the social network predict colorectal cancer survival? *European Journal of Cancer*; 42(17): 3022-3027.
- Vineis P. (2006) Misuse of genetic data in environmental epidemiology. *Annals of the New York Academy of Sciences*; 1076: 163–167.
- Vismara Fugini A, Antonelli A, Giovanessi L, Gardini VC, Abuhilal M, Zambolin T, Tardanico R, Simeone C, Cunico SC. (2011) Insignificant prostate cancer: characteristics and predictive factors. *Urologia*; 78(3): 184-186.
- Vogt TM, Mullooly JP, Ernst D, Pope CR, Hollis JF. (1992) Social networks as predictors of ischemic heart disease, cancer, stroke and hypertension: incidence, survival and mortality. *Journal of Clinical Epidemiology*; 45: 659–666.
- Waller HT, Starky C. (1984) What is the best indicator of health care? 596th World Health Forum; 5: 276–279.
- Walker GV, Grant SR, Guadagnolo BA, Hoffman KE, Smith BD, Koshy M, Allen PK, Mahmood U. (2014) Disparities in stage at diagnosis, treatment, and survival in nonelderly adult patients with cancer according to insurance status. *Journal of Clinical Oncology*; 32(28): 3118-3125.
- Wang L, Wilson SE, Stewart DB, Hollenbeak CS. (2011) Marital status and colon cancer outcomes in US Surveillance, Epidemiology and End Results registries: does marriage affect cancer survival by gender and stage? *Cancer Epidemiology*; 35(5): 417-422.

- Wang X, Yin L. (2015) Point and interval estimation of baseline risk and treatment effect based on logistic model for observational studies. *Biometrical Journal*; 57(3): 441-452.
- Watson M, Haviland JS, Greer S, Davidson J, Bliss JM. (1999) Influence of psychological response on survival in breast cancer: a population-based cohort study. *Lancet*; 354: 1331–1336.
- Waxler NE, Hislop TG, Mears B, Kan L. (1991) Effects of social relationship on survival for women with breast cancer: A prospective study. *Social Science & Medicine*; 33: 177-183.
- Waxler-Morrison N, Hislop TG, Mears B, Kan L. (1991) Effects of social relationships on survival for women with breast cancer: a prospective study. *Social Science and medicine*; 33(2): 177-183.
- Wegner EL, Kolonel LN, Nomura AMY, Lee J. (1982) Racial and socioeconomic status differences in survival of colorectal cancer patients in Hawaii. *Cancer*; 49: 2208-2216.
- Weiss B. (2008) Chemobrain: a translational challenge for neurotoxicology. *Neurotoxicology*; 29(5): 891-898.
- Whitaker KL, Scott SE, Wardle J. (2015) Applying symptom appraisal models to understand sociodemographic differences in responses to possible cancer symptoms: a research agenda. *British Journal of Cancer*; 112 Suppl. 1: S27-34.
- WHO (2015) GLOBOCAN 2012. Estimated cancer incidence, mortality and prevalence worldwide in 2012. Lyon, IARC. http://globocan.iarc.fr/Pages/fact_sheets_cancer.aspx
- Whynes DK, Frew EJ, Manghan CM, Scholefield JH, Hardcastle JD. (2003) Colorectal cancer, screening and survival: The influence of socio-economic deprivation. *Public Health*; 117: 389–395.
- Wilkinson R, Marmot M. (eds.) (1998) *Social determinants of health: The solid facts*. Copenhagen, WHO Regional Office for Europe.
- Wilkinson R, Marmot M. (eds.) (2003) *Social determinants of health: The solid facts*. 2nd Edition. Copenhagen, WHO Regional Office for Europe.
- Wilkinson R, Pickett K. (2010) *The Spirit Level: why greater equality makes societies stronger*. New York, Bloomsbury.
- Wilkinson R. (1996) *Unhealthy societies: From inequality to well-being*. London, Routledge.
- Wilkinson R. (2005) *The Impact of Inequality: How to Make Sick Societies Healthier*. New York, The New Press.

Williams R, Barefoot J, Califf R. (1992) Prognostic importance of social and economic resources among medically treated patients with angiographically documented coronary heart disease. *Journal of American Medical Association*; 267: 520–524.

Wolff, S. N. (2007). The burden of cancer survivorship. In M. Feuerstein (Ed.), *Handbook of cancer survivorship* (pp. 7-18). New York: Springer.

Woods LM, Rachet B, Coleman MP. (2005) Choice of geographic unit influences socioeconomic inequalities in breast cancer survival. *British Journal of Cancer*; 92(7): 1279-1282.

Woods LM, Rachet B, Coleman MP. (2006) Origins of socio-economic inequalities in cancer survival: a review. *Annals of Oncology*; 17(1): 5-19.

Woods LM, Rachet B, O'Connell D, Lawrence G, Coleman MP. (2016) Impact of deprivation on breast cancer survival among women eligible for mammographic screening in the West Midlands (UK) and New South Wales (Australia): Women diagnosed 1997-2006. *International Journal of Cancer*; [Epub ahead of print].

Wrigley H, Roderick P, George S, Smith J, Mullee M, Goddard J. (2003) Inequalities in survival from colorectal cancer: A comparison of the impact of deprivation, treatment, and host factors on observed and cause specific survival. *Journal of Epidemiology & Community Health*; 57: 301–309.

Yabroff KR, Gordis L. (2003) Does stage at diagnosis influence the observed relationship between socioeconomic status and breast cancer incidence, case fatality, and mortality? *Social Science & Medicine*; 57: 2265–2279.

Young JL, Gloeckler-Ries L, Pollack ES. (1984) Cancer Patient Survival among ethnic groups in the United States. *Journal of National Cancer Institute*; 73: 341-352.

Zackrisson S, Andersson I, Manjer J, Janzon L. (2004) Non-attendance in breast cancer screening is associated with unfavourable socio-economic circumstances and advanced carcinoma. *International Journal of Cancer*; 108(5): 754-760.

Zanetti R. (2007) I Registri Tumori tra passato e future. In: Ferretti S, Giacomini A, Gruppo di lavoro AIRTUM (eds). *Manuale di Tecniche di Registrazione dei Tumori*. Milano, Inferenze Scarl.

Zhang-Salomons J, Qian H, Holowaty E, Mackillop WJ. (2006) Associations between socioeconomic status and cancer survival: choice of SES indicator may affect results. *Annals of Epidemiology*; 16(7): 521-538.

Appendice

Tab. A.1 – Elenco delle variabili presenti nel database del Registro Tumori, comuni a tutte le neoplasie.

| Variabili comuni a tutti i tumori |
|--|
| Tabella:Paziente |
| Nome |
| Cognome |
| Sesso |
| Data nascita |
| Comune di nascita |
| Codice fiscale |
| Id anagrafica |
| Tabella:Anagrafica |
| Comune di residenza |
| Codice sanitario |
| Indirizzo |
| Medico di base |
| Tipo di assistito |
| Tabella:Tumore |
| Fonte cartella |
| Fonte anatomia patologica |
| Fonte specialista |
| Fonte terapia |
| Fonte DCO |
| Fonte SDO |
| Fonte screening |
| Fonte RT esterno |
| Fonte medico di base |
| Fonte ospedale fuori regione |
| Fonte rimborsi |
| Fonte certificato morte |
| Data prima diagnosi |
| Età di prima diagnosi |
| Lateralità |
| Stato in vita |
| Lastdate |
| Ultimo comune di residenza |
| ICD |
| ICDOT |
| ICDOM |
| ICCC |

| Variabili comuni a tutti i tumori |
|--|
| Tabella:Tumore |
| ICCC esteso |
| Y |
| T |
| Lettera T |
| N |
| Lettera N |
| M |
| Lettera M |
| Grado |
| Valore stadio |
| Lettera stadio |
| TNM ricostruito |
| Stadio ricostruito |
| TNM senza controlli |
| Modalità diagnosi citologia |
| Modalità diagnosi biochimica immunologica |
| Modalità diagnosi ecografia |
| Modalità diagnosi ematologia |
| Modalità diagnosi istologia su metastasi o autoptica |
| Modalità diagnosi clinica |
| Modalità diagnosi DCO |
| Modalità diagnosi istologica |
| Modalità diagnosi radiologia |
| Modalità diagnosi RMN |
| Modalità diagnosi endoscopica |
| Modalità diagnosi autoptica senza istologia |
| Modalità diagnosi sconosciuta |
| Tipologia scheda |
| Commenti |
| Follow-up |
| Tabella:Deceduti |
| Data morte |
| Prima causa di morte |
| Seconda causa di morte |
| Terza causa di morte |
| Comune di morte |
| Giorni prima malattia |
| Mesi prima malattia |
| Anni prima malattia |
| Giorni seconda malattia |
| Mesi seconda malattia |
| Anni seconda malattia |
| Giorni terza malattia |
| Mesi terza malattia |

| Variabili comuni a tutti i tumori |
|--|
| Tabella:Deceduti |
| Anni terza malattia |
| Tabella:Metastasi |
| ICD |
| ICDOT |
| ICDOM |
| Data diagnosi |
| Modalità diagnosi citologica |
| Modalità diagnosi istologica |
| Modalità diagnosi radiologica |
| Modalità diagnosi ecografica |
| Modalità diagnosi TCRMN |
| Modalità diagnosi endoscopica |
| Modalità diagnosi chirurgica |
| Modalità diagnosi autoptica |
| Modalità diagnosi clinica |
| Modalità diagnosi anamnestica |
| Modalità diagnosi SDO |
| Modalità diagnosi medico di base |
| Modalità diagnosi certificato di morte |
| Altra modalità diagnosi |
| Tabella:SedeNodale |
| ICD |
| ICDOT |
| ICDOM |
| Data diagnosi |
| Modalità diagnosi citologica |
| Modalità diagnosi istologica |
| Modalità diagnosi radiologica |
| Modalità diagnosi ecografica |
| Modalità diagnosi TCRMN |
| Modalità diagnosi endoscopica |
| Modalità diagnosi chirurgica |
| Modalità diagnosi autoptica |
| Modalità diagnosi clinica |
| Modalità diagnosi anamnestica |
| Modalità diagnosi SDO |
| Modalità diagnosi medico di base |
| Modalità diagnosi certificato di morte |
| Altra modalità diagnosi |
| Tabella:AnatomiaPatologica |
| Data diagnosi |
| Numero referto |
| Ospedale |
| Reparto |

| Variabili comuni a tutti i tumori |
|--|
| Tabella:AnatomiaPatologica |
| ICD |
| ICDOT |
| ICDOM |
| Descrizione macroscopica |
| Descrizione microscopica |
| Y |
| T |
| Lettera T |
| N |
| Lettera N |
| M |
| Lettera M |
| Grado |
| Valore stadio |
| Lettera stadio |
| TNM ricostruito |
| Stadio ricostruito |
| Inizialmente codificato come stadio 0 |
| Tabella:Cartella |
| Data diagnosi |
| Numero |
| Ospedale |
| reparto |
| Data ricovero |
| Data dimissione |
| Stato dimissione |
| Codice dimissione |
| Informazione da rilevatore |
| ICD |
| ICDOT |
| ICDOM |
| T |
| Lettera T |
| N |
| Lettera N |
| M |
| Lettera M |
| Valore stadio |
| Lettera stadio |
| TNM ricostruito |
| Stadio ricostruito |
| Modalità diagnosi ecografia mammaria |
| Modalità diagnosi ecografia addome |
| Modalità diagnosi Rx torace |

| Variabili comuni a tutti i tumori |
|--|
| Tabella:Cartella |
| Modalità diagnosi Tc |
| Modalità diagnosi RMN |
| Modalità diagnosi PET |
| Modalità diagnosi scintigrafia ossea |
| Modalità diagnosi mammografia |
| Altra modalità diagnosi |
| Tabella:Recidiva |
| Data |
| Ospedale |
| Reparto |
| ICDOM |

Tab. A.2.1 – Elenco delle variabili relative allo studio ad alta risoluzione sul tumore alla mammella.

| Variabili relative allo studio ad alta risoluzione sul tumore alla mammella |
|--|
| Tabella:Mammella |
| Eccezione date |
| Modalità diagnosi mammografia |
| Data modalità diagnosi mammografia |
| Classe modalità diagnosi mammografia |
| Modalità diagnosi ecografia |
| Data modalità diagnosi ecografia |
| Modalità diagnosi citologia |
| Modalità diagnosi biopsia |
| Classe modalità diagnosi biopsia |
| Modalità diagnosi solo clinica |
| Data modalità diagnosi solo clinica |
| Modalità diagnosi RMN |
| Data modalità diagnosi RMN |
| Modalità diagnosi anamnestica |
| Data modalità diagnosi anamnestica |
| Modalità diagnosi autopsia |
| Data modalità diagnosi autopsia |
| Modalità diagnosi screening |
| Data modalità diagnosi screening |
| Tabella:PatologieConcomitantiMammella |
| IMA |
| Ipertensione arteriosa trattata |
| Arteriopatie periferiche |
| Demenza |
| Malattia cerebrovascolare |
| Connettivite |
| Insufficienza respiratoria cronica |

| Variabili relative allo studio ad alta risoluzione sul tumore alla mammella |
|--|
| Tabella:PatologieConcomitantiMammella |
| Diabete |
| Malattia ulcerosa |
| Tumore |
| Emiplegia |
| Leucemia |
| Insufficienza renale moderata/severa |
| Insufficienza epatica moderata/severa |
| Linfoma |
| AIDS |
| Insufficienza cardiaca congestizia |
| Menopausa |
| Classificazione ASA |
| Altro 1 |
| Altro 2 |
| Tabella:CaratteristicheTumoreMammella |
| Diametro (mm) |
| Focalità |
| Grading |
| Infiltrazione linfovaskolare |
| Infiltrazione neurovascolare |
| Recettori estrogeni |
| Recettori estrogeni segno |
| Recettori progesterone |
| Recettori progesterone segno |
| Recettori Ki67/Mib1 |
| Cerb B2 ICC pos |
| Cerb B2 ICC marcatura |
| Cerb B2 ICC punteggio |
| Cerb B2 segno |
| Cerb B2 Dako pos |
| Cerb B2 Dako marcatura |
| Cerb B2 Dako punteggio |
| Cerb B2 Fish |
| Cerb B2 polisomia cromosoma 17 |
| Tabella:CitologiaMammella |
| Non reperibile |
| Data |
| Ospedale |
| Reparto |
| Numero referto |
| Classe |
| Descrizione |
| Tabella:BiopsiaMammella |
| Non reperibile |

| Variabili relative allo studio ad alta risoluzione sul tumore alla mammella |
|--|
| Tabella:BiopsiaMammella |
| Data |
| Ospedale |
| Reparto |
| Numero referto |
| Classe |
| Tipo |
| Descrizione |
| Tabella:LinfonodiMammella |
| Data |
| Ospedale |
| Reparto |
| Numero referto |
| Sede |
| Modalità diagnosi |
| Esaminati |
| Con metastasi |
| Con micrometastasi |
| Con cellule neoplastiche in transito |
| Tabella:LinfonodiSentinellaMammella |
| Data |
| Ospedale |
| Reparto |
| Numero referto |
| Modalità diagnosi |
| Esaminati |
| Con metastasi |
| Con micrometastasi |
| Con cellule neoplastiche in transito |
| Tabella:ChirurgiaMammella |
| Data |
| Ospedale |
| Reparto |
| Tipo intervento |
| Classe intervento |
| Specifica intervento |
| Pezzo operatorio non disponibile |
| Data accettazione pezzo operatorio |
| Data referto pezzo operatorio |
| Ospedale pezzo operatorio |
| Reparto pezzo operatorio |
| Numero referto pezzo operatorio |
| Estemporanea pezzo operatorio |
| Macroscopica pezzo operatorio |
| Microscopica pezzo operatorio |

| Variabili relative allo studio ad alta risoluzione sul tumore alla mammella |
|--|
| Tabella:ChirurgiaMammella |
| Diametro (mm) pezzo operatorio |
| Margini pezzo operatorio |
| Distanza dal margine (mm) pezzo operatorio |
| Linfonodo sentinella |
| Data linfonodo sentinella |
| Linfonodi ascellari |
| Data linfonodi ascellari |
| Tabella:ChemioterapiaMammella |
| Tipo |
| Ospedale |
| Reparto |
| Protocollo |
| Numero cicli |
| Data inizio |
| Data fine |
| Note |
| Tabella:ImmunoterapiaMammella |
| Tipo |
| Ospedale |
| Reparto |
| Protocollo |
| Numero cicli |
| Data inizio |
| Data fine |
| Note |
| Tabella:OrmonoterapiaMammella |
| Tipo |
| Ospedale |
| Reparto |
| Protocollo |
| Classe farmacologica |
| Data inizio |
| Data fine |
| Note |
| Tabella:RadioterapiaMammella |
| Tipo |
| Ospedale |
| Reparto |
| Sede |
| Data inizio |
| Data fine |
| Note |

Tab. A.2.2 – Elenco delle variabili relative allo studio ad alta risoluzione sui tumori al colon-retto.

| Variabili relative allo studio ad alta risoluzione sui tumore al colon-retto |
|---|
| Tabella:ColonRetto |
| Modalità diagnosi solo clinica |
| Data modalità diagnosi solo clinica |
| Modalità diagnosi Rx (clisma opaco) |
| Data modalità diagnosi Rx (clisma opaco) |
| Modalità diagnosi ecografia transrettale |
| Data modalità diagnosi ecografia transrettale |
| Modalità diagnosi endoscopia |
| Tipo modalità diagnosi endoscopia |
| Data modalità diagnosi endoscopia |
| Modalità diagnosi RMN |
| Data modalità diagnosi RMN |
| Modalità diagnosi Tc spirale |
| Data modalità diagnosi Tc spirale |
| Modalità diagnosi SOF |
| Risultato modalità diagnosi SOF |
| Data modalità diagnosi SOF |
| Modalità diagnosi screening |
| Data ultimo invito programmato screening |
| Data ultimo SOF screening |
| Data richiamo dopo SOF screening |
| Adesione al richiamo screening |
| Classificazione caso secondo screening |
| Tabella:PatologieConcomitantiColonRetto |
| Ipertensione arteriosa trattata |
| Infarto miocardio |
| Diabete |
| Scompenso cardiaco |
| Emiplegia |
| Insufficienza epatica |
| Insufficienza renale |
| AIDS |
| Connettivite |
| Demenza |
| Malattia infiammatoria cronica |
| Malattia cerebrovascolare |
| BPCO |
| Arteriopatie periferiche |
| Linfoma |
| FAP |
| RCU |
| Leucemia |
| Altra neoplasia |

| Variabili relative allo studio ad alta risoluzione sui tumore al colon-retto |
|---|
| Tabella:PatologieConcomitantiColonRetto |
| Classificazione ASA |
| Altro 1 |
| Altro 2 |
| Tabella:ComplicanzeColonRetto |
| Tipologia |
| Fonte |
| Giorni degenza |
| Tabella:CitologiaColonRetto |
| Data paracentesi |
| Risultato paracentesi |
| Tabella:BiopsiaColonRetto |
| Data |
| Ospedale |
| Reparto |
| Ospedale diagnosi istologica |
| Data accettazione |
| Data referto |
| Numero referto |
| ICD |
| ICDOT |
| ICDOM |
| Microscopica |
| Tabella:ProceduraEndoscopicaColonRetto |
| Ospedale |
| Reparto |
| Data ricovero |
| Intervento |
| Tipologia intervento |
| Data intervento |
| Tabella:IstologiaColonRetto |
| Ospedale diagnosi istologica |
| Data accettazione |
| Data referto |
| Numero referto |
| ICD |
| ICDOT |
| ICDOM |
| Dimensione polipo (mm) |
| Distanza da orifizio anale (mm) |
| Aspetto macroscopico |
| Asse non valutabile |
| Descrizione macroscopica |
| Descrizione microscopica |
| Descrizione margine |

| Variabili relative allo studio ad alta risoluzione sui tumore al colon-retto |
|---|
| Tabella:IstologiaColonRetto |
| Infiltrazione perineurale |
| Dimensione maggiore di 1 cm |
| Componente villosa |
| Displasia di alto grado |
| Adenocarcinoma intramucoso |
| Tabella:ChirurgiaColonRetto |
| Ospedale |
| Reparto |
| Data prenotazione |
| Data ricovero |
| Intervento |
| Modalità intervento |
| Causa mancata esecuzione |
| Tipologia intervento |
| Data intervento |
| Urgenza |
| Esordio |
| Distanza da ano (mm) |
| Preservazione sfintere anale |
| Stomia |
| Ricanalizzazione dopo stomia |
| Data ricanalizzazione |
| Tabella:PezzoOperatorioColonRetto |
| Ospedale diagnosi istologica |
| Data accettazione |
| Data referto |
| Numero referto |
| Margine prossimale |
| Margine distale |
| Margine radiale |
| Distanza minima neof-margine |
| Infiltrazione linfovaskolare |
| Infiltrazione perineurale |
| Macroscopica |
| Microscopica |
| Lunghezza (mm) |
| Lunghezza non nota |
| TRG |
| Contorni noduli tumorali |
| Numero noduli tumorali |
| Tabella:LinfonodiColonRetto |
| Sede |
| Esaminati |
| Positivi |

| Variabili relative allo studio ad alta risoluzione sui tumore al colon-retto |
|---|
| Tabella: Marcatori Colon Retto |
| Data EGFR |
| Sede EGFR |
| Segno EGFR |
| Valore EGFR |
| Data K-ras |
| Sede K-ras |
| Segno K-ras |
| Tabella: Chemioterapia Colon Retto |
| Tipo |
| Causa mancata esecuzione |
| Ospedale |
| Reparto |
| Protocollo |
| Dose (mg) |
| Numero cicli |
| Completamento protocollo |
| Data inizio |
| Data fine |
| Tabella: Immunoterapia Colon Retto |
| Tipo |
| Causa mancata esecuzione |
| Ospedale |
| Reparto |
| Protocollo |
| Dose (mg) |
| Numero cicli |
| Completamento protocollo |
| Data inizio |
| Data fine |
| Tabella: Radioterapia Colon Retto |
| Tipo |
| Causa mancata esecuzione |
| Ospedale |
| Reparto |
| Dose totale programmata |
| Dose totale erogata |
| Numero cicli |
| Completamento protocollo |
| Data prenotazione |
| Data inizio |
| Data fine |

Doc. A.1 – Lettera di richiesta dati inviata a tutti i comuni umbri.



Perugia. ...

Ill.mo Signor Sindaco
del Comune di

...

e alla c.a. del
Responsabile Ufficio Anagrafe
FAX ...

Ill.mo Signor Sindaco,
il Registro Tumori Umbro di Popolazione, istituito con Legge regionale n. 18 del 12/11/2012, sta proseguendo il progetto di ricerca sulla sopravvivenza dei malati di tumore, che rappresenta il parametro fondamentale per la valutazione dell'efficacia e dell'efficienza dei servizi sanitari regionali nel controllo della patologia tumorale. Per la realizzazione di tale progetto risulta indispensabile, come già proficuamente avvenuto in passato, la collaborazione degli Uffici anagrafe dei Comuni per l'acquisizione di alcune informazioni relative ad alcuni pazienti registrati nel database ad alta risoluzione disponibile presso il nostro Registro. Al fine di aggiornare e completare le informazioni dei singoli pazienti, per finalità di ricerca orientata a definire come le disuguaglianze socio-economiche condizionano lo stato di salute, tematica quest'ultima particolarmente rilevante per finalità di valutazione dell'attività di cura e dell'utilizzo ottimale delle risorse a disposizione del sistema sanitario, si chiede di voler integrare le seguenti informazioni:

**stato civile,
titolo di studio,
condizione professionale (se possibile)**

per alcune persone di cui allegheremo le informazioni necessarie per il reperimento nei Vostri archivi.

Tale procedura ci permette di aggiornare l'archivio di incidenza ad alta risoluzione per le patologie di interesse, già presente presso il nostro Registro, che integrato con i dati delle ASL, può costituire una fonte completa e omogenea di dati, a livello della Regione dell'Umbria. Tale ricerca è inserita in un progetto nazionale che riguarda Registri tumori di diverse Regioni.

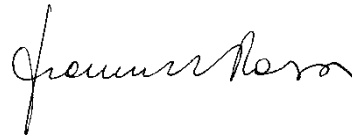
A tale proposito La informo che i dati saranno trattati a norma dell'art. 94 del "Codice in materia di protezione dei dati personali" (D. Leg.vo n. 196 dle 30.06.2003) e che

tutto il personale addetto alla rilevazione ed all'elaborazione dei dati stessi è soggetto al vincolo del segreto professionale.

Per ulteriori chiarimenti e per concordare le modalità di rilevazione è possibile contattare il dott. Roberto Lillini, coordinatore della ricerca, al numero di cellulare 347.3039225 (e-mail: r.lillini@campus.unimib.it), a cui potrete inviare le informazioni al seguente numero di fax: 075.5857317.

In attesa di conoscere la Vostra disponibilità e di un cortese riscontro, si porgono cordiali saluti

Il Direttore del RTUP
Prof. Francesco La Rosa



Mittente:

RTUP, Dipartimento di Medicina Sperimentale dell'Università di Perugia

Sezione di Sanità Pubblica, via del Giochetto, 06126 PERUGIA

Fax: 075.5857317; e-mail: francesco.larosa@unipg.it

Tab. A.3 – Stadio alla diagnosi e titolo di studio per pazienti affette da tumore alla mammella: modelli di regressione logistica.

| Stadio alla diagnosi | | OR | Sig. |
|----------------------|-----------------------------------|------|----------------------------|
| Stadio I | Licenza elementare (o analfabeta) | 1,13 | n.s. |
| | Licenza media inferiore | 1,32 | n.s. |
| | Diploma media superiore | 1,15 | n.s. |
| | Laurea | 1 | (categoria di riferimento) |
| Stadio II | Licenza elementare (o analfabeta) | 0,96 | n.s. |
| | Licenza media inferiore | 0,85 | n.s. |
| | Diploma media superiore | 0,97 | n.s. |
| | Laurea | 1 | (categoria di riferimento) |
| Stadio III | Licenza elementare (o analfabeta) | 0,76 | n.s. |
| | Licenza media inferiore | 0,80 | n.s. |
| | Diploma media superiore | 0,77 | n.s. |
| | Laurea | 1 | (categoria di riferimento) |
| Stadio IV | Licenza elementare (o analfabeta) | 1,18 | n.s. |
| | Licenza media inferiore | 0,86 | n.s. |
| | Diploma media superiore | 0,98 | n.s. |
| | Laurea | 1 | (categoria di riferimento) |

Nota: OR = Odds ratio.

Tab. A.4 – Terapie e variabili socio-economiche per pazienti affette da tumore alla mammella: modelli di regressione logistica.

| Terapie | | OR | Sig. |
|---------------|-----------------------------------|------|----------------------------|
| Chemioterapia | Licenza elementare (o analfabeta) | 0,45 | p<0,05 |
| | Licenza media inferiore | 1,08 | n.s. |
| | Diploma media superiore | 0,86 | n.s. |
| | Laurea | 1 | (categoria di riferimento) |
| Immunoterapia | Licenza elementare (o analfabeta) | 0,77 | p<0,05 |
| | Licenza media inferiore | 0,90 | n.s. |
| | Diploma media superiore | 1,11 | n.s. |
| | Laurea | 1 | (categoria di riferimento) |
| Radioterapia | Licenza elementare (o analfabeta) | 0,56 | p<0,05 |
| | Licenza media inferiore | 0,91 | n.s. |
| | Diploma media superiore | 0,77 | n.s. |
| | Laurea | 1 | (categoria di riferimento) |
| Chirurgia | Licenza elementare (o analfabeta) | 1,02 | n.s. |
| | Licenza media inferiore | 1,14 | n.s. |
| | Diploma media superiore | 1,09 | n.s. |
| | Laurea | 1 | (categoria di riferimento) |

Nota: OR = Odds ratio.

Ringraziamenti

Al termine di questi tre anni di dottorato desidero ringraziare tutte le persone che a vario titolo mi hanno accompagnato in questo percorso e senza le quali non sarebbe stato possibile realizzare questo lavoro.

Innanzitutto voglio ringraziare i miei tutor, i Professori Mario Lucchini ed Emanuela Sala, che mi hanno dato il loro contributo teorico e metodologico durante tutte le fasi della ricerca e hanno seguito con grande attenzione ogni fase del lavoro di scrittura di questa tesi.

Ringrazio poi tutte le persone che mi hanno aiutato nel lungo lavoro di raccolta e organizzazione dei dati presso il Registro Tumori dell'Umbria e il Dipartimento di Specialità Medico-Chirurgiche e Sanità Pubblica dell'Università degli Studi di Perugia, in particolare i Professori Francesco La Rosa e Fabrizio Stracci, che hanno messo a disposizione tutte le informazioni sui pazienti di tumore e il cui indispensabile lavoro di raccordo e contatto con gli enti locali ha permesso la costruzione dei *dataset* su cui questo studio si basa.

Un sentito ringraziamento va ai colleghi del Dipartimento di Epidemiologia Medica e Biostatistica del Karolinska Institutet, specialmente i Professori Paul Dickman, Paul Lambert, Michael Crowther e Flaminia Chiesa, non solo per il contributo di conoscenza metodologica e statistica che costituisce il punto di forza di questa tesi, ma anche per tutti i piacevoli momenti di tempo libero passato assieme in quella bellissima città che è Stoccolma.

Un grosso grazie va, ovviamente, a tutti i docenti e colleghi che hanno partecipato a questi tre anni di Scuola di Dottorato, con un pensiero particolare alla nostra coordinatrice, Professoressa Carmen Leccardi, che ci ha sempre seguito e stimolato a sfruttare il più possibile tutte le opportunità di conoscenza offerte dalla Scuola.

E, soprattutto, il ringraziamento più importante a Carla, mia compagna di vita, che in questi anni non mi ha mai fatto mancare presenza e supporto e, da sempre, mi stimola a continuare il mio percorso di crescita scientifica.