

5. APPENDICE METODOLOGICA

Il dato di mortalità è soggetto da molti anni a leggi e regole che ne determinano la creazione, la raccolta e la codifica standardizzata e sistematica (Classificazione delle Malattie e Cause di Morte, nona revisione, *ICD9*)⁽⁴⁾, il flusso e l'elaborazione sia centrale (*ISTAT*) sia regionale (Registri di Mortalità). In altre parole, può essere considerato un sistema informativo affidabile ed esaustivo fondato, ormai da anni, su una buona concordanza tra diagnosi clinica e certificazione della causa di decesso, almeno per quanto riguarda la patologia neoplastica^(5,6).

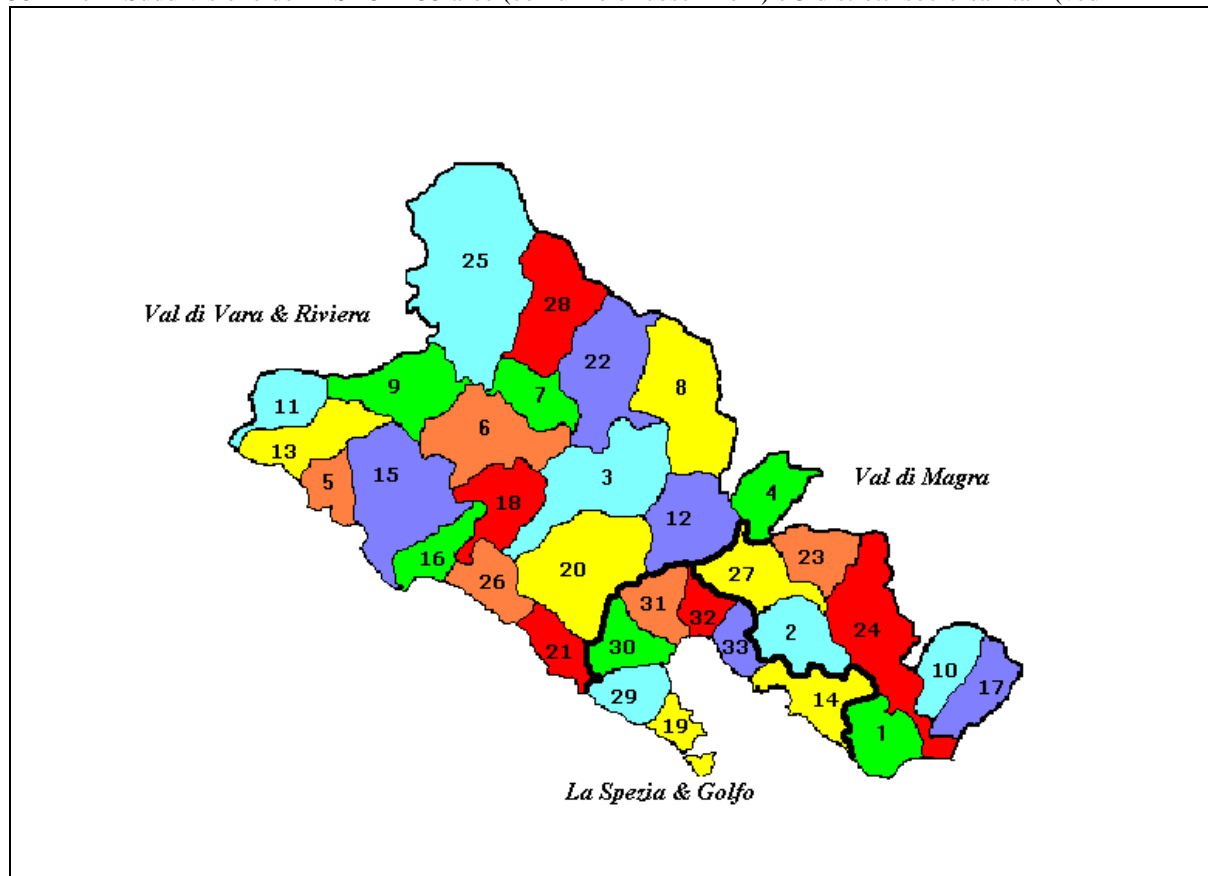
5.1 Fonte dei dati

I dati di mortalità per i Comuni dell'ASL 5 Spezzino sono stati forniti dal Registro di Mortalità Regionale per il periodo 1988-2006. Le popolazioni a rischio per sesso, età ed area sono state ottenute dalla banca dati dell'*ISTAT* e dal Servizio Anagrafe e Demografia del Comune della Spezia⁽⁷⁾.

5.2 Suddivisione del territorio

Al fine di avere aree geografiche con la massima omogeneità numerica possibile, per il Comune capoluogo si sono considerate le circoscrizioni, ottenendo in totale per l'ASL 33 aree (28 Comuni e 5 Circoscrizioni) (Figura 79), all'interno delle quali sono stati calcolati gli indici sintetici di mortalità.

FIGURA 79 – Suddivisione dell'ASL 5 in 33 aree (comuni e circoscrizioni) e 3 distretti socio-sanitari (vedi **TABELLA 8**).



<i>Distretto socio-sanitario</i>			<i>Aree</i>		
17	<i>Val di Vara & Riviera</i>	VVR	3	<i>Beverino</i>	<i>BEV</i>
			4	<i>Bolano</i>	<i>BOL</i>
			5	<i>Bonassola</i>	<i>BON</i>
			6	<i>Borghetto</i>	<i>BOR</i>
			7	<i>Brugnato</i>	<i>BRU</i>
			8	<i>Calice C</i>	<i>CAL</i>
			9	<i>Carro</i>	<i>CAR</i>
			11	<i>Deiva M</i>	<i>DEI</i>
			12	<i>Follo</i>	<i>FOL</i>
			13	<i>Framura</i>	<i>FRA</i>
			15	<i>Levanto</i>	<i>LEV</i>
			16	<i>Monterosso M</i>	<i>MON</i>
			18	<i>Pignone</i>	<i>PIG</i>
			20	<i>Riccò G</i>	<i>RIC</i>
21	<i>Riomaggiore</i>	<i>RIO</i>			
22	<i>Rocchetta V</i>	<i>ROC</i>			
25	<i>Sesta G</i>	<i>SES</i>			
26	<i>Vernazza</i>	<i>VER</i>			
28	<i>Zignago</i>	<i>ZIG</i>			
18	<i>La Spezia & Golfo</i>	LSG	14	<i>Lerici</i>	<i>LER</i>
			19	<i>Portovenere</i>	<i>POR</i>
			29	<i>La Spezia C1</i>	<i>SP1</i>
			30	<i>La Spezia C2</i>	<i>SP2</i>
			31	<i>La Spezia C3</i>	<i>SP3</i>
			32	<i>La Spezia C4</i>	<i>SP4</i>
33	<i>La Spezia C5</i>	<i>SP5</i>			
19	<i>Val di Magra</i>	VDM	1	<i>Ameglia</i>	<i>AME</i>
			2	<i>Arcola</i>	<i>ARC</i>
			10	<i>Castelnuovo M</i>	<i>CAS</i>
			17	<i>Ortonovo</i>	<i>ORT</i>
			23	<i>S. Stefano</i>	<i>SAN</i>
			24	<i>Sarzana</i>	<i>SAR</i>
27	<i>Vezzano L</i>	<i>VEZ</i>			

TABELLA ? – Suddivisione dell'ASL 5 in 33 aree (comuni e circoscrizioni) e 3 distretti socio-sanitari (**FIGURA**).

5.3 Metodi statistici

Sono stati utilizzati i seguenti indicatori:

- **Tasso di mortalità specifico per età (T_s)** – In generale, il tasso è una misura della frequenza (relativa) con cui una particolare malattia compare in una popolazione, in un dato intervallo di tempo. Deriva dalla divisione tra il numero di casi di malattia e la popolazione ritenuta a rischio di sviluppare la stessa malattia. Comunemente, il tasso è espresso come numero (medio) di casi per centomila abitanti. Il tasso specifico per età è l'indice di frequenza calcolato all'interno di date classi d'età (0-4, 5-9, ...) della popolazione in studio¹⁹⁻²². Se O_{esp} è il numero dei decessi osservati all'età e -esima ($e = 1, 2, \dots, E$), nel sesso s -esimo ($s = 1$ per i maschi, $s = 2$ per le femmine), nel periodo p -esimo ($p = 1, 2, \dots, P$) e per la causa c -esima ($c = 1, 2, \dots, C$), allora il tasso sarà:

$$T_{s_{esp}} = \frac{O_{esp}}{P_{esp}} \cdot k$$

dove P_{sea} è la popolazione a rischio e k è una costante moltiplicativa (in genere centomila) che riflette l'ammontare della popolazione a rischio.

- **Tasso di mortalità standardizzato per età (TST)** – Si tratta di un indice di frequenza ottenuto applicando i tassi specifici per fasce d'età quinquennali ad una popolazione tipo (p.e., la popolazione ligure del 1988-2006). L'utilizzo dei tassi standardizzati per età è necessario se si vogliono confrontare differenti periodi e popolazioni con differente struttura per età perché consente di "rimuovere" l'effetto confondente di tale caratteristica¹⁹⁻²².
- **Rapporto Standardizzato di Mortalità (SMR)** – Mentre un tasso è una misura dell'intensità di un fenomeno sanitario, SMR è invece un indice di confronto dell'intensità del fenomeno in due aree e/o periodi differenti. E' infatti dato dal rapporto tra il numero di decessi osservati (O) e il numero dei decessi attesi (A) in base ad una popolazione di riferimento¹⁹⁻²²:

$$SMR_{spc} = \frac{\sum_{e=1}^E O_{esp}}{\sum_{e=1}^E T_{esp} \cdot P_{esp}} = \frac{O_{spc}}{A_{spc}}$$

dove T_{esp} è il tasso specifico per età nella popolazione standard e P_{esp} è la popolazione in studio. In pratica, gli attesi non sono altro che i decessi che si potrebbero osservare se la popolazione in studio morisse come la popolazione di riferimento. Il computo dell'intervallo di confidenza al 95% (IC 95%) è attuato assumendo gli O_{spc} come una realizzazione di una variabile casuale che segue la distribuzione di Poisson.

- **Differenza Media Assoluta (DMA)** – Esprime quanto in media il tasso è cambiato nel periodo in studio ogni anno rispetto all'anno precedente. Se TST_a è il tasso per centomila abitanti nell' a -esimo anno e TST_{a-1} è l'analogia misura nell'anno precedente, allora la DMA_a è data da:

$$DMA_a = TST_a - TST_{a-1}$$

L'aumento del numero di decessi è indicato dal segno positivo (+), mentre la diminuzione da un segno negativo (-).

- **Differenza Media Percentuale (DMP)**: rappresenta la variazione percentuale annuale di casi nel periodo in studio:

$$DMP_a = 100 \cdot (TST_a - TST_{a-1}) / TST_{a-1}$$

Anche in questo caso, l'aumento della percentuale di decessi è indicato dal segno positivo (+), mentre la diminuzione da un segno negativo (-).

Per quanto riguarda il SMR , si sottolinea che la sua struttura analitica può essere suddivisa in due componenti: un "segnale epidemiologico" e un "rumore statistico". La prima componente è il rischio reale di un'area, la seconda è invece da ricondurre a fluttuazioni casuali spazio-temporali. Quando si considerano piccole aree geografiche e/o piccole popolazioni e/o malattie rare, il SMR diciamo "grezzo", cioè così come

viene calcolato, è scarsamente affidabile e occorre applicare adeguate tecniche statistiche, dette bayesiane⁶⁵, in grado di amplificare il segnale e ridurre il rumore^{23,1.a.i.24}. In altri termini, oltre a prendere in esame i rischi relativi “grezzi”, è necessario considerare anche quelli “corretti” secondo la metodologia bayesiana.

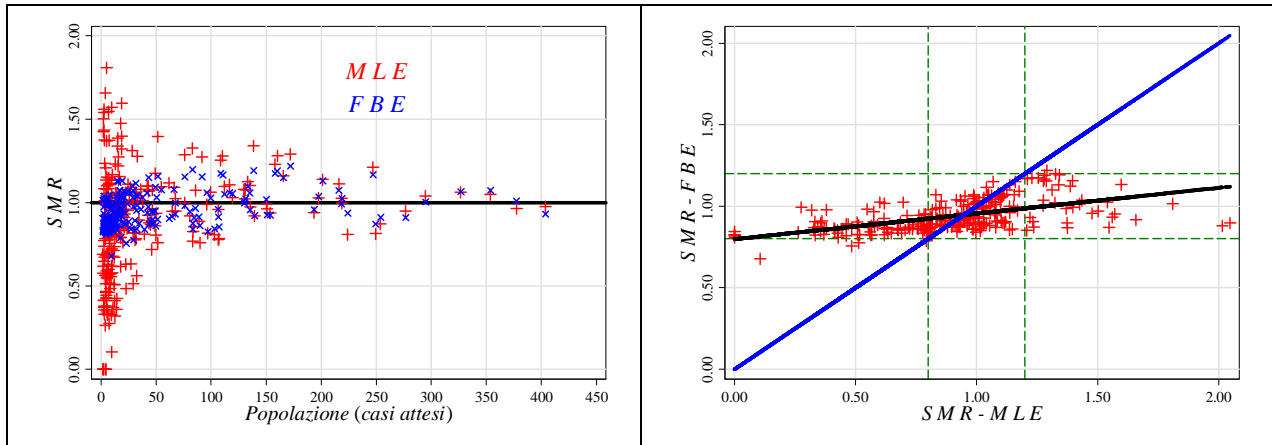


FIGURA 80 – A sinistra, comportamento del rischio grezzo ($MLE-SMR$) e di quello aggiustato con la tecnica bayesiana ($FBE-SMR$) in relazione alla dimensione della popolazione delle aree in studio. A destra, confronto tra $SMR-FBE$ e $SMR-MLE$.

Gli stimatori bayesiani completi del SMR , SMR_{FBE} ($FBE = full\ bayes\ estimates$)²⁴, adottati nel presente lavoro sono calcolati tenendo conto anche della correlazione spaziale tra gli eventi, ovvero della tendenza spesso osservata dei tassi di mortalità ad assumere valori simili in aree attigue. Gli SMR calcolati senza applicare tecniche di correzione (perequazione o lisciamento), cioè gli SMR “grezzi”, vengono qui indicati come SMR_{MLE} , ($MLE = maximum\ likelihood\ estimates$).

La Figura 80 (sinistra) illustra un esempio di comportamento dei due indicatori di rischio (MLE e FBE) in funzione dei casi attesi, che riflettono la numerosità della popolazione di ogni area. Qui, le aree sono i 267 comuni e circoscrizioni (per Genova, La Spezia e Savona) della Liguria e la patologia in esame è il tumore del polmone nei maschi, nel periodo 1988-1998. La popolazione ligure nello stesso periodo è utilizzata come standard. Si nota che il MLE ha una grande variabilità quando la popolazione è piccola, mentre tende a stabilizzarsi al crescere della stessa. Viceversa, il rischio FBE presenta una variabilità minore e praticamente indipendente dai casi attesi. Inoltre, mentre al di sotto dei 50 casi attesi (popolazione più piccola) il rischio è fortemente perequato o ristretto (*shrinkage*) verso 1.00, quasi indipendentemente dal valore del MLE , al di sopra di tale soglia (popolazione più grande) le differenze sono molto piccole o nulle. La parte destra della Figura 80 mostra infine la relazione tra i due indici. Ovviamente, in assenza di *shrinkage*, la diagonale principale (linea blu) rappresenterebbe il loro andamento comune. In realtà, l’andamento è quello riassunto dalla linea nera, che evidenzia una forte regressione dei valori estremi di rischio (molto minori e molto maggiori di 1.00), in quanto ricavati da piccole popolazioni. Valori di MLE compresi tra 0.80 e 1.20 (linee verdi tratteggiate) e basati su popolazioni più grandi, mostrano una stabilità maggiore. Infatti, in quell’intervallo di MLE ricadono praticamente tutti gli FBE .

Occorre infine ricordare che entrambi gli indici sono soggetti a fluttuazione casuale nello spazio e nel tempo, ed è quindi necessario conoscere l’intervallo di confidenza al 95% ($IC\ 95\%$), intervallo in cui, con una probabilità del 95%, cadrà il reale valore dell’ SMR .

Se l’ **$IC\ 95\%$ comprende il valore 1.00**, le differenze riscontrate, in termini di SMR , tra l’area in studio e la popolazione di riferimento non sono statisticamente differenti e quindi da attribuirsi al caso. Se, viceversa, **l’ $IC\ 95\%$ non comprende il valore 1.00** la differenza osservata in più o in meno sarà effettiva (sempre con un margine di errore del 5%).

Considerazioni analoghe valgono anche per la DMA e la DMP . Entrambi gli indici riflettono un andamento differente da zero (nessun incremento o decremento) quando l’ $IC\ 95\%$ esclude il valore nullo. In pratica, perché esista una variazione statisticamente significativa è necessario che entrambi gli estremi dell’ $IC\ 95\%$ abbiano lo stesso segno algebrico: negativo per la diminuzione, positivo per l’aumento.

Gli $SMR-FBE$ sono stati ottenuti mediante una specifica metodologia statistica⁶⁷ attraverso il programma *WinBUGS*⁶⁸. Le stime dei TST , dei $SMR-MLE$ e della DMA e della DMP sono state calcolate utilizzando il pacchetto statistico *STATA*⁶⁹.

⁶⁵ Il nome di questo settore della statistica deriva da Thomas Bayes (Londra 1702, Tunbridge 1761) matematico e reverendo metodista, che approfondì e sviluppò un importante settore del calcolo delle probabilità (probabilità condizionate), il cui valore scientifico fu evidenziato solo dopo la sua scomparsa.