



# BOLLETTINO DELLA SOCIETA' GEOGRAFICA ITALIANA

Serie 14 – 2021

Vol. 4 – n. 2

ISSN 1121-7820



Poste Italiane spa - Tassa pagata - Piego di libro  
Aut. n. 072/DCB/ETWF del 31.03.2005



COMITATO SCIENTIFICO

COMITATO EDITORIALE

**Franco Adamo**, Università del Piemonte Orientale, Italia

**Vittorio Amato**, Università di Napoli Federico II, Italia

**Gianfranco Battisti**, Università degli Studi di Trieste, Italia

**Giuliano Bellezza**, Università di Roma La Sapienza, Italia

**Jean-Marc Besse**, Université Paris 1 - Panthéon-Sorbonne, Francia

**Edoardo Boria**, Sapienza Università di Roma, Italia

**Giuseppe Campione**, Università di Messina, Italia

**Laura Cassi**, Università degli Studi di Firenze, Italia

**Gabriele Ciampi**, Università degli Studi di Firenze, Italia

**Federico Cugurullo**, Trinity College Dublin, Irlanda

**Fiorella Dallari**, Alma Mater Studiorum - Università di Bologna, Italia

**Hélène Dang Vu**, École d'Urbanisme de Paris, Francia

**Giuseppe Dematteis**, Politecnico di Torino, Italia

**Gino De Vecchis**, Università di Roma La Sapienza, Italia

**Francesco Dramis**, Università degli Studi Roma Tre, Italia

**Paolo Roberto Federici**, Università di Pisa, Italia

**Carmen García Martínez**, Universidad de Castilla-La Mancha, Albacete, Spagna

**Tiago Luis Gil**, Universidade de Brasília, Brasile

**Alberto Giordano**, Texas State University, USA

**Guillaume Giroir**, Université d'Orléans, Francia

**Maria Gemma Grillotti**, Università Campus Bio-Medico di Roma, Italia

**Luc Gwiazdzinski**, Université Grenoble Alpes, Grenoble, Francia

**Piergiorgio Landini**, Università G. D'Annunzio Chieti Pescara, Italia

**Jean-Pierre Lozato-Giotart**, Université Paris Sorbonne Nouvelle, Francia

**Elio Manzi**, Università degli Studi di Palermo, Italia

**Glaucio José Marafon**, Universidade do Estado do Rio de Janeiro, Brasile

**Maria Mautone**, Università degli Studi di Napoli Federico II, Italia

**Jean Nogué**, Universitat de Girona, Spagna

**Elvidio Lupia Palmieri**, Università di Roma La Sapienza, Italia

**Bas Pedroli**, Wageningen University & Research, Paesi Bassi

**Maria Pigaki**, National Technical University of Athens, Grecia

**Hindenburg Pires**, Universidade de São Paulo, Brasile

**Mariagiovanna Riitano**, Università degli Studi di Salerno, Italia

**Joseph Salukvadze**, Tbilisi State University, Georgia

**Michael Samers**, University of Kentucky, USA

**Paola Sereno**, Università degli Studi di Torino, Italia

**Claudio Smiraglia**, Università degli Studi di Milano, Italia

**Olivier Soubeyran**, Institut Géographie Alpine, Grenoble, Francia

**Marcello Tanca**, Università degli Studi di Cagliari, Italia

**Maria Tinacci Mossello**, Università degli Studi di Firenze, Italia

**Maria Chiara Zerbi**, Università degli Studi di Milano, Italia

*Direttore scientifico*

**Margherita Azzari**, Università degli Studi di Firenze

*Vicedirettori*

**Marco Maggioli**, Università Iulm, Milano

**Marina Fuschi**, Università degli Studi "G. D'Annunzio" Chieti e Pescara

*Redazione*

**Valentina E. Albanese**, Università degli Studi dell'Insubria

**Arianna Antonielli**, Redattore, Università degli Studi di Firenze

**Camillo Berti**, Redattore, Università degli Studi di Firenze

**Pierluigi De Felice**, Università degli Studi di Salerno

**Pauline Deguy**, Redattore, Università degli Studi di Firenze

**Fabrizio Ferrari**, Redattore, Università degli Studi "G. D'Annunzio" Chieti e Pescara

**Luisa Spagnoli**, Istituto di Storia dell'Europa Mediterranea, Consiglio Nazionale delle Ricerche

**Paola Zamperlin**, Redattore, Università di Pisa

**Cecilia Maria Roberta Luschi e Laura Aiello**, Progetto grafico di copertina, Università degli Studi di Firenze

# BOLLETTINO DELLA SOCIETA' GEOGRAFICA ITALIANA

Fondato nel 1868



SERIE XIV - VOLUME 4

FASCICOLO 2

DICEMBRE 2021

Pubblicato dalla

**SOCIETA' GEOGRAFICA ITALIANA**  
**VILLA CELIMONTANA • ROMA<sup>ONLVS</sup>**

SEMESTRALE

***Bollettino della Società Geografica Italiana***

*Published by*

**Firenze University Press** – University of Florence, Italy

Via Cittadella, 7 - 50144 Florence - Italy

<http://www.fupress.com/bsgi>

Direttore Responsabile: **Margherita Azzari**, University of Florence, Italy

Immagine di copertina: Anonimo, *I fiori di glicine a Tokyo*, 1868-1908 (Società Geografica Italiana, Archivio fotografico, Campagna oceanica della Regia Nave Vesuvio in Estremo Oriente, inv. 2/1/313)

**Copyright** © 2021 **Authors**. The authors retain all rights to the original work without any restrictions.

**Open Access.** This issue is distributed under the terms of the [Creative Commons Attribution 4.0 International License \(CC-BY-4.0\)](https://creativecommons.org/licenses/by/4.0/) which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The Creative Commons Public Domain Dedication (CC0 1.0) waiver applies to the data made available in this issue, unless otherwise stated.



**Citation:** F. Bitonti, A. Mazza (2021). Determinanti geografiche della mortalità per tumore tiroideo nella Sicilia Orientale. *Bollettino della Società Geografica Italiana* serie 14, 4(2): 45-64. doi: 10.36253/bsgi-1290

**Copyright:** © 2021 F. Bitonti, A. Mazza. This is an open access, peer-reviewed article published by Firenze University Press (<http://www.fupress.com/bsgi>) and distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Data Availability Statement:** All relevant data are within the paper and its Supporting Information files.

**Competing Interests:** The Author(s) declare(s) no conflict of interest.

## Determinanti geografiche della mortalità per tumore tiroideo nella Sicilia orientale

### Geographical determinants of thyroid cancer mortality in Eastern Sicily

FRANCESCA BITONTI, ANGELO MAZZA

*Dipartimento di Economia e Impresa, Università di Catania, Italia*  
E-mail: francesca.bitonti@phd.unict.it; a.mazza@unict.it

**Abstract.** The study of geographic determinants in the processes of spreading infectious diseases has a long tradition. As a result of the increase in the incidence of chronic and non-infectious diseases, such as cancer and various types of heart disease, especially in the more advanced countries, the typical methods of spatial analysis have also been applied to these types of diseases. In fact, their heterogeneous aetiology often includes environmental risk factors that need to be investigated using geographic techniques. The object of this study is the possible influence of an environmental risk factor on the onset of thyroid cancer (TC): volcanic activity. Numerous studies conducted in different parts of the world show that in volcanic areas the incidence of this tumour is significantly higher than the national average. In the present work, we propose an exploratory study of the spatial distribution of TC in Eastern Sicily, where the Mt. Etna is located. Georeferencing the data provided by the Cancer Registry of Eastern Sicily, we will provide a visualization of the phenomenon through geographical maps, highlighting the areas with significantly higher incidence than expected. We will then apply the generalized Ripley's K-function and the local Moran's I to verify the presence of high-risk areas, taking into account both the proximity to the volcano and the non-homogeneous distribution of the population over the territory. Our results seem to confirm the consolidated opinion that the presence of a volcano can determine the increase in the incidence of TC.

**Keywords:** spatial autocorrelation, Local Moran's I, generalized Ripley's K-function, point process.

**Riassunto.** Lo studio delle determinanti geografiche nei processi di diffusione delle malattie infettive vanta una lunga tradizione. In conseguenza dell'aumento dell'incidenza di malattie croniche e non infettive, come il cancro e diversi tipi di patologie cardiache, soprattutto nei paesi maggiormente avanzati, i metodi tipici dell'analisi spaziale sono stati applicati anche a questi tipi di patologie. La loro eterogenea eziologia, infatti, spesso include fattori ambientali di rischio che necessitano di essere indagati mediante tecniche geografiche. Oggetto di questo studio è la possibile influenza di un fattore ambientale di rischio sull'insorgenza del tumore tiroideo (TT): l'attività vulcanica. Numerosi studi condotti in diverse parti del mondo rilevano che in prossimità di aree vulcaniche l'incidenza di tale tumore sia nettamente superiore alle medie nazionali. Nel presente lavoro, proponiamo uno studio esplorativo della distribuzione spaziale del TT

nella Sicilia orientale, in cui è situato il vulcano Etna. Georeferenziando i dati del Registro Tumori della Sicilia orientale, forniamo una visualizzazione del fenomeno attraverso carte geografiche, evidenziando le aree con incidenza significativamente maggiore rispetto a quella attesa. Successivamente applichiamo la funzione K di Ripley generalizzata e l'indice locale Moran I per verificare la presenza di *cluster* di zone a maggiore rischio, tenendo in considerazione sia la prossimità al vulcano che la distribuzione non omogenea della popolazione sul territorio. I risultati di questa ricerca sembrano confermare l'opinione consolidata che la presenza di un vulcano possa determinare l'aumento di incidenza del TT.

**Parole chiave:** autocorrelazione spaziale, Moran Locale, funzione K di Ripley generalizzata, processo puntuale.

## 1. Introduzione

Sul finire del XVIII secolo il dottor Valentine Seaman mappò i casi di febbre gialla a New York e in tal modo riuscì a evidenziare una possibile correlazione tra i siti di varie discariche e la localizzazione dei casi (Stevenson 1965). Circa sessant'anni dopo John Snow ebbe l'idea di creare una mappa dei casi di colera che in quel periodo stava affliggendo il quartiere di Soho a Londra e capì che la causa dell'epidemia era dovuta all'acqua contaminata dagli scarichi fognari riversati nel Tamigi e che poi sgorgava da una fontana pubblica. Facendo chiudere la fontana riuscì a contenere i contagi (Snow 1855; Walter 2000). Questi sono solo due dei primi tentativi di usare l'analisi geografica come strumento per fornire indicazioni in ambito epidemiologico. A partire da quegli studi iniziali, l'analisi geografica dei dati sanitari ha conosciuto una diffusione sempre crescente, divenendo uno strumento essenziale degli studi epidemiologici.

Con l'avvio, nei paesi più avanzati, della terza fase della transizione epidemiologica (Omran 1971), durante la quale si è assistito ad un'incidenza decrescente della mortalità derivante da malattie trasmissibili e ad una rilevanza crescente delle patologie di tipo degenerativo (e.g., le patologie cardiovascolari, o il cancro) e di quelle determinate da specifici modelli di comportamento umano (e.g., violenza, incidenti, o abuso di sostanze), l'impiego delle metodologie tipiche dell'analisi spaziale si è diffuso anche nell'ambito dello studio delle patologie non trasmissibili, nelle quali i fattori di rischio ambientali sono sovente annoverati tra le possibili concause (Palagiano 1991; Ghosh et al. 1999; Wakefield 2007; Bertazzon 2009; Geraghty et al. 2010; Palagiano, Pesaresi 2011; Najafabadi e Pourhassan 2011; Musa et al. 2013; Bristow et al. 2015; Mahdavifar et al.

2016; Tlacuilo-Parra et al. 2017; Ye et al. 2017; Anand et al. 2020). Come osservano Palagiano e Pesaresi (2011), "l'ambiente di per sé non genera malattie. Tuttavia, alcuni fattori dell'ambiente possono concorrere allo sviluppo di particolari patologie o quanto meno costituiscono i luoghi dove si riscontrano le affezioni". Già da tempo, infatti, la malattia è considerata come il prodotto dell'interazione tra fattori patologici (come vettori e cause genetiche) e fattori geografici che agiscono a livello fisico, biologico e sociale (May 1950). Ormai da decenni si pone l'attenzione verso l'ambiente, sia esso fisico che antropico, e si guarda alla salute umana come a un bene da perseguire all'interno di un complesso "sistema salute", come lo definisce Palagiano, e come alla principale finalità degli studi di geografia (Palagiano 1990; Palagiano et al. 1992).

Il crescente sviluppo dei sistemi di informazione geografica (GIS) ha facilitato l'integrazione di dati ambientali e sanitari, spesso tra loro eterogenei sia per diversa configurazione spaziale (puntuale, lineare, areale) che per livello di aggregazione spaziale. Inoltre, si è assistito ad una convergenza sempre maggiore tra GIS e piattaforme di calcolo statistico che ha reso agevole l'implementazione di metodologie di analisi spaziale anche molto complesse da un punto di vista analitico e computazionale. Queste tecnologie facilitano la comprensione dei problemi di salute pubblica e favoriscono la collaborazione tra medici, epidemiologi e geografi per mappare e prevedere il rischio di malattia (Croner et al. 1996; Craglia, Maheswaran 2004; Berke 2010; Skinner 2010; Palagiano, Pesaresi 2011, 315-318; Nykiforuk, Flaman 2011; Sanders, McKay 2013; Delaunay et al. 2015; Keddem et al. 2015; Zambrano et al. 2019; Dangermond et al. 2020).

Il tumore tiroideo (TT), oggetto di questo studio, è la neoplasia endocrina più diffusa, la cui incidenza è cresciuta costantemente in tutto il mondo negli ultimi decenni (Curado et al. 2007; Kilfoy et al. 2009; Fitzmaurice et al. 2015; Liu et al. 2017). Gli studi epidemiologici suggeriscono che l'eziologia del TT comprenda il consumo di iodio, la familiarità e l'esposizione a radiazioni ionizzanti (Fagin, Wells 2016; Kitahara, Sosa 2016; Seib, Sosa 2019). Le cause sottostanti al progressivo aumento dell'incidenza del TT sono attribuite al progressivo aumento dell'esposizione alle radiazioni ionizzanti e all'avanzamento delle tecniche diagnostiche.

Kung et al. (1981) hanno ravvisato gli elementi presenti nei gas vulcanici come possibili ulteriori agenti eziologici del TT. Altri studi hanno confermato un'incidenza del TT maggiore rispetto alle medie nazionali nelle isole Hawaii (Goodman et al. 1988; Kolonel et al. 1990), in Islanda (Arnbjörnsson et al. 1986; Hrafnkels-

son et al. 1989), nelle Filippine (Duntas, Doumas 2009; Caguioa et al. 2019), in Nuova Caledonia (Truong et al. 1985) e nella Polinesia francese (Curado et al. 2007; Bray et al. 2017), tutte regioni il cui comun denominatore è la presenza di aree vulcaniche attive. Nello studio di Goodman et al. (1988) sono stati confrontati i tassi standardizzati di incidenza di TT dei diversi gruppi etnici residenti nelle Hawaii con quelli degli stessi gruppi residenti in altre aree geografiche (Tabella 1). I risultati mostrano un'incidenza superiore per i residenti nelle isole Hawaii, con la sola eccezione per i "bianchi" residenti in Islanda (Waterhouse et al. 1982; Menck e Henderson 1985).

Dati aggiornati pubblicati dall'OMS (Bray et al. 2017) confermano per le Hawaii, la Nuova Caledonia e

l'Islanda, un tasso di incidenza standardizzato più elevato rispetto alla maggior parte degli altri paesi analizzati.

Anche in Sicilia, in prossimità delle aree vulcaniche attive, sono stati riscontrati tassi di incidenza del TT più alti rispetto alle medie regionali e nazionali (Pellegriti et al. 2009; Malandrino et al. 2013; Vigneri et al. 2015; Malandrino et al. 2016; Vigneri et al. 2017) e, in particolare, nella zona intorno alle pendici dell'Etna è stata osservata un'incidenza doppia rispetto a quella della regione nel suo complesso (Pellegriti et al. 2009).

In particolare, sembrerebbe essere ampia la letteratura medica riguardo le aree vulcaniche tipo basaltico, come quelle situate nei paesi precedentemente menzionati: Kilauea e Mauna Loa nelle Hawaii (Macdonald 1968; Moore e Trusdell 1993), Pinatubo, Taal e Majon nelle Filippine (Newhall 1979; Pallister et al. 1996; Alanis et al. 2013), Eyjafjallajökull in Islanda (Keiding, Sigmarsson 2012) ed Etna in Sicilia (Polacci et al. 2019).

Gli studi relativi alla Sicilia si fondano su confronti tra tassi di incidenza osservati in un'area specifica, in genere l'unità statistica impiegata è il comune, e tassi riferiti ad aggregati spaziali di controllo più ampi, quali la regione Sicilia o l'Italia. Tali confronti si fondano sull'impiego di test inferenziali che presuppongono l'indipendenza tra le unità statistiche. Tale presupposto, tuttavia, risulta spesso violato dalla presenza di autocorrelazione, i.e., i valori di una variabile osservati in luoghi vicini tra di loro tendono ad avere andamenti simili, mentre osservazioni di una stessa variabile effettuate in luoghi lontani tendono ad avere valori differenti. Questo postulato è stato espresso in maniera efficace da Waldo Tobler (1970) che, nella prima legge della geografia, afferma che: "Ogni cosa è correlata a qualsiasi altra, ma le cose vicine sono più correlate di quelle lontane". Il nostro intento, in questo lavoro, è quello di indagare aspetti più prettamente geografici della diffusione del TT, quali la presenza di forme di concentrazione spaziale nell'allocatione dei casi tenendo sotto controllo l'eterogeneità spaziale derivante, ad esempio, dalla diversa struttura per sesso ed età della popolazione esposta al rischio.

Il presente lavoro è così organizzato: il secondo paragrafo descrive le caratteristiche principali dell'area oggetto di studio, le fonti di dati a disposizione e le strategie per la loro integrazione; il terzo paragrafo descrive la metodologia di analisi spaziale applicata; il quarto paragrafo illustra e discute la distribuzione del TT nella parte est della Sicilia e mostra la presenza di *cluster* di zone ad alto e basso rischio; il quinto e ultimo paragrafo riporta le conclusioni del lavoro.

**Tabella 1.** Confronto dei tassi di incidenza medi annui aggiustati per età (standard mondiale) per 100.000 abitanti per cancro alla tiroide di gruppi etnici simili che vivono alle Hawaii e in altre aree. Basato sui tassi di incidenza del lavoro (Waterhouse et al. 1982), ad eccezione dei tassi per i filippini, che si basano sul lavoro (Menck, Henderson 1985). Fonte: Goodman et al. 1988.

Ethnic group	Sex	Incidence rates	
		Hawaii	Other populations
Chinese	F	12.1	9.6 San Francisco
			5.3 Los Angeles
	M	7.8	3.7 Singapore
			3.7 San Francisco
Japanese	F	6.2	8.2 San Francisco
			6.6 Los Angeles
	M	4.8	1.7 Osaka
			2.4 Los Angeles
Polynesian	F	17.6	7.5 New Zealand Maoris
	M	6.0	1.9 New Zealand Maoris
White	F	6.7	15.0 Iceland
			6.5 San Francisco
	M	3.8	1.5 Norway
			5.0 Iceland
Filipino	F	16.6	2.9 San Francisco
			1.5 Norway
	M	6.6	5.0 Manila
			12.0 Los Angeles
			2.0 Manila
			4.0 Los Angeles

## 2. Dati epidemiologici e territorio

Il TT è la neoplasia endocrina più diffusa al mondo, e negli ultimi decenni sta registrando un incremento costante. Nel periodo 2008-2016, in Italia è stato rilevato un tasso annuale di incidenza del TT pari a 9.4 per i maschi e a 26 per le femmine (Fondazione AIOM 2020); i tassi riportati sono stati standardizzati con il metodo della standardizzazione diretta, utilizzando come popolazione tipo la “Revised European Standard Population” (RESP), al fine di eliminare l’effetto della diversa struttura demografica della popolazione residente nelle diverse aree geografiche considerate (Eurostat 2013). Per entrambi i sessi, si rileva un’incidenza maggiore nel centro e nel sud Italia rispetto al nord (Tabella 2). Confrontando invece i sessi, risulta evidente che il TT colpisca maggiormente le donne rispetto agli uomini.

Tassi d’incidenza superiori rispetto alle medie nazionali sono stati registrati in diverse aree vulcaniche che presentano caratteristiche simili alla Sicilia orientale; tale area comprende quattro province: Messina, Catania, Enna e Siracusa. La zona vulcanica, che fa riferimento all’Etna, il vulcano europeo attivo più alto d’Europa, è situata nella provincia di Catania, ma coinvolge anche alcune zone del Messinese meridionale. La Tabella 3 riporta il tasso di incidenza standardizzato di TT per le province della Sicilia orientale per il periodo 2003-2011 (Assessorato Regionale alla Salute 2016). Per entrambi i sessi, i tassi osservati nella provincia di Catania risultano superiori sia rispetto alla media regionale che a quella nazionale; tale dato trova

un riscontro per il periodo 2002-2004 nelle analisi statistiche di Pellegriti et al. (2009).

Molteplici studi hanno rilevato, nel corso del tempo, la presenza di livelli elevati di metalli pesanti nelle aree limitrofe all’Etna, in conseguenza delle continue emissioni di gas (composte in prevalenza da CO<sub>2</sub> e SO<sub>2</sub>), cenere e lava da parte del vulcano (Buat-Ménard, Arnold 1978; Cimino, Ziino 1983; Caltabiano et al. 2004; Andronico et al. 2009; D’Aleo et al. 2016). Tali metalli pesanti includono tra gli altri arsenico, cadmio, cromo, cobalto, mercurio, tungsteno e zinco che, in concentrazioni elevate, potrebbero contaminare il suolo, l’acqua e l’atmosfera, entrando infine nella catena alimentare (Vigneri et al. 2017). In particolare, un recente studio (Malandrino et al. 2016) ha rilevato una concentrazione di oligoelementi nell’acqua potabile, nei licheni e nelle urine dei residenti dell’area circumetnea maggiore (e significativamente superiore per quanto riguarda le urine) rispetto ad altre zone siciliane non vulcaniche. Sono stati registrati elevati contenuti di metalli pesanti anche nei prodotti ortofrutticoli coltivati nelle vicinanze del vulcano (Ferrante et al. 2013) e nei capelli dei bambini residenti nei pressi dell’Etna (Varrica et al. 2014). Questi studi indicano che la presenza di un vulcano attivo potrebbe contaminare, attraverso le ripetute emissioni di fumi, ceneri e lava, il territorio circostante portando a potenziali ripercussioni per la salute umana. Una discussione maggiormente estensiva riguardo le possibili relazioni tra la concentrazione di metalli pesanti nell’organismo e l’insorgenza del TT è presente nel lavoro di Vigneri et al. (2017).

Dal punto di vista geografico, il territorio delle province analizzate risulta essere molto variegato e comprende sia l’area vulcanica ma anche aree urbane, rurali e industriali (Istat, 2013). In particolare, negli ultimi cinquant’anni si è assistito a una progressiva concentrazione della popolazione lungo le coste settentrionali e orientali, soprattutto nelle aree metropolitane, che ha accentuato le differenze tra aree urbanizzate, congestionate ma con accesso ai servizi, e aree interne, isolate, scarsamente popolate e quasi completamente carenti nei servizi di base (Scrofani et al. 2020). La conseguenza maggiormente evidente del rafforzamento di tale gap è stata la forte perdita demografica registrata dalle aree interne negli ultimi quarant’anni circa (Novembre 2015). Secondo la classificazione introdotta dalla Strategia Nazionale per le Aree Interne (SNAI) per definire le aree territoriali in base al grado di spopolamento e degrado, e alla distanza dei servizi essenziali, come quelli riguardanti istruzione, salute e mobilità, è possibile individuare fra i 207 comuni costituenti la Sicilia orientale: 4 poli, corrispondenti ai quattro capoluoghi di provincia, 47

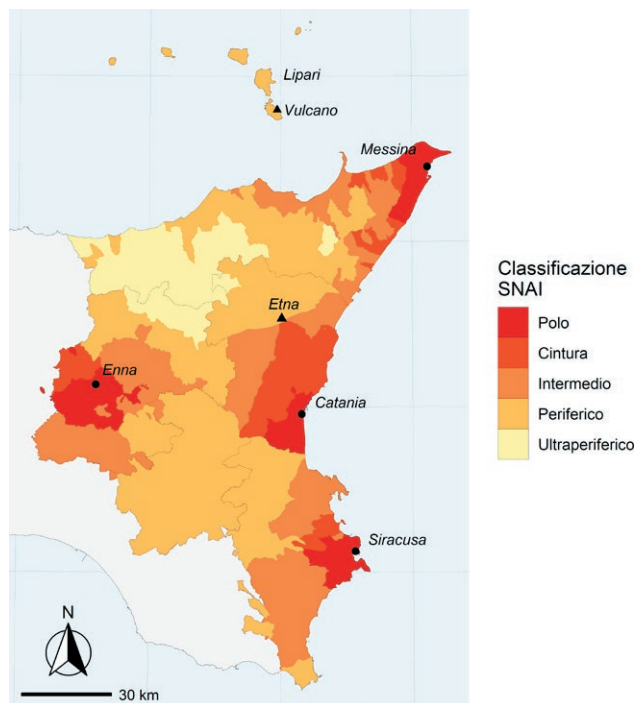
**Tabella 2.** Tasso medio annuale di incidenza del TT, per sesso e area geografica. Fonte: Fondazione AIOM, 2020.

Area	Maschi	Femmine
Nord	9.2	24.9
Centro	11	28.6
Sud e Isole	9.5	27.2
Italia	9.4	26

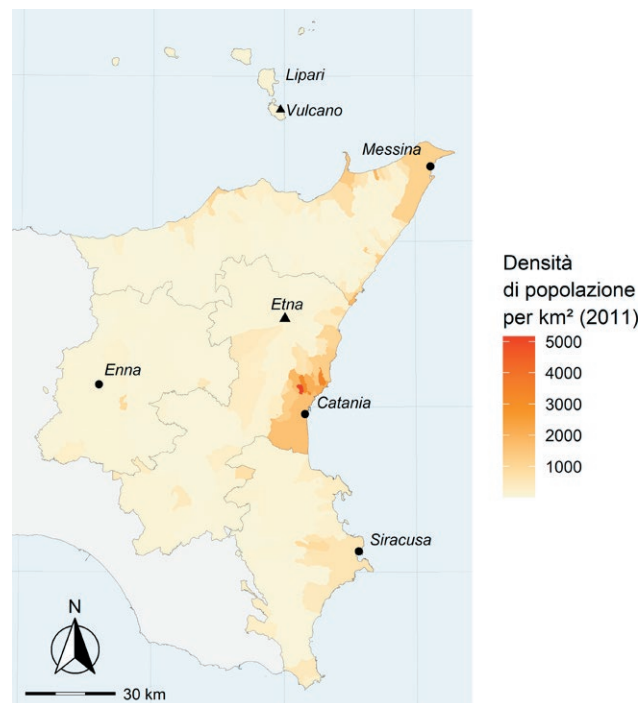
**Tabella 3.** Tasso standardizzato di incidenza del TT, per sesso e area geografica. Fonte: Atlante Sanitario della Sicilia 2016.

Area	Maschi	Femmine
Catania	10.6	35.5
Enna	7.7	25.5
Messina	9.6	29.1
Siracusa	5.2	18.9
Sicilia	7.7	25.6





**Figura 1.** Classificazione delle aree SNAI nella Sicilia orientale. Fonte: nostra elaborazione su dati della Agenzia per la Coesione Territoriale disponibili al seguente link: <https://www.agenziacoesione.gov.it/strategia-nazionale-aree-interne>.



**Figura 2.** Densità di popolazione per km<sup>2</sup> al 2011 su base comunale in Sicilia orientale. Fonte: nostra elaborazione su dati Istat.

aree di cintura, 60 aree intermedie, 76 aree periferiche e 20 aree ultraperiferiche (Fig. 1) (Accordo di Partenariato 2014-2020, 2013).

Risulta chiaro, dunque, che la disomogeneità della morfologia urbana, del territorio e della copertura di servizi e trasporti comporti una distribuzione eterogenea della popolazione residente e di conseguenza anche dei casi di TT nella regione considerata (Fig. 2).

I dati sanitari a disposizione provengono dal Registro Tumori della Sicilia Orientale e si riferiscono ai casi di TT diagnosticati nel periodo 2003-2016 a individui residenti nelle quattro province di interesse con un'età compresa tra 5-95 anni. In totale, i casi ammontano a 7.085.

La geocodifica degli indirizzi di residenza degli individui è stata effettuata tramite Google Maps API (<https://cloud.google.com/maps-platform/>). Il servizio di geocodifica offerto da Google è stato preferito ad ArcGIS\* ([www.esri.com](http://www.esri.com)) e ad OpenStreetMap (<https://www.openstreetmap.org/>). Il primo è stato escluso poiché non fruibile gratuitamente. Il secondo, OpenStreetMap, è completamente gratuito e molto promettente; tuttavia, allo stato attuale, presenta una limitazione nel numero di richieste al secondo che possono essere effettuate, precisamente una al secondo. Google Maps garantisce una

velocità di risposta notevolmente maggiore (50 richieste al secondo) rispetto a quella di OpenStreetMap ed offre dei piani tariffari gratuiti ed a pagamento a seconda del numero di richieste di geocodifica inviate (<https://developers.google.com/maps/documentation/geocoding/usage-and-billing>). Per questa applicazione, il numero di richieste di geocodifica necessario rientra ampiamente nella fascia gratuita. Sul piano qualitativo, il suo costante aggiornamento e l'ampia copertura, soprattutto nelle aree interne ed isolate, lo rendono comparabile e talvolta preferibile ad OpenStreetMap (Karmacharya 2018; Plantin 2018; Franzini et al. 2020). La geocodifica dei dati è stata effettuata tramite la piattaforma statistica R (R Core Team 2020), che include pacchetti dedicati, quali *ggmap* e *googleway*, che consentono un accesso *user-friendly* alle API Google (Kahle, Wickham 2013; Cooley 2020). La geocodifica degli indirizzi dei soggetti colpiti da TT ha permesso, successivamente, di riferire ogni individuo a una delle 7318 sezioni censuarie costituenti la Sicilia orientale.

Gli indirizzi dei casi di TT forniti dal Registro Tumori non presentavano una struttura normalizzata mentre l'API di Google restituisce, ad ogni interrogazione, la geocodifica e l'ubicazione completa in formato

standard (la struttura del formato standard utilizzato da Google Maps come output della geocodifica è consultabile alla seguente pagina web: <https://developers.google.com/maps/documentation/geocoding/overview>). La mancata corrispondenza tra il database a disposizione e l'API di Google ha comportato delle difficoltà nel processo di geocodifica e ha inciso, anche se in maniera marginale, sull'accuratezza della stessa poiché il servizio cartografico opera una ricerca più precisa se l'indirizzo fornito è già normalizzato. Per correggere eventuali errori nella geocodifica e misurarne l'accuratezza è stato verificato anzitutto se le coordinate spaziali restituite da Google ricadessero all'interno dell'area oggetto di studio, e poi nel comune di residenza. Successivamente, è stato possibile accertare se ci fosse una corrispondenza esatta tra i toponimi risultanti nel Registro Tumori e quelli restituiti da Google Maps. I casi che presentavano delle discrepanze sono stati corretti mediante procedura manuale. I dati riguardanti la popolazione residente nelle sezioni censuarie delle quattro province di interesse provengono dal 15° Censimento Generale della Popolazione effettuato dall'Istat nell'anno 2011 (<https://www.istat.it/it/archivio/104317>).

### 3. Metodologia

#### 3.1. Analisi della distribuzione spaziale

La distribuzione spaziale dei tumori può essere rappresentata mediante un processo puntuale, visualizzato come una serie di punti su una mappa in cui i punti, chiamati "eventi", rappresentano i casi di tumore. La probabilità di riscontrare un caso di tumore in dato punto nello spazio è influenzata da un lato dalla distribuzione geografica della popolazione, con le sue caratteristiche demografiche, e dall'altro dalla presenza di specifici fattori ambientali di rischio. È infatti noto che la popolazione non si dispone uniformemente sul territorio, ma si concentra in aree urbane densamente abitate, lasciando per lo più deserte ampie zone rurali e montuose. Inoltre, le caratteristiche demografiche della popolazione stessa potrebbero non essere uniformi, basti pensare ai comuni interni in cui sono in atto dinamiche di spopolamento che rendono preponderante la percentuale di donne anziane rispetto alle altre fasce di età e al sesso maschile. L'eterogeneità nella composizione demografica della popolazione, e quindi dell'esposizione al rischio di tumore, richiede il ricorso a tecniche di standardizzazione. In aggiunta, anche la morfologia del territorio può presentare notevoli difformità persino tra aree limitrofe, come la presenza di zone vulcaniche adia-

centi a zone costiere e pianeggianti. Pertanto, il rischio che si manifesti un evento dipenderà dalla popolazione a rischio, e quindi dalla presenza di caratteristiche demografiche particolari, e dalla prossimità a determinati fattori ambientali. Se da un lato l'aspetto legato alla distribuzione geografica e alle caratteristiche demografiche della popolazione possono essere tenuti sotto controllo, come accennato, ricorrendo ai dati disponibili sulla popolazione residente e a metodologie matematiche, non disponiamo, di dati sufficientemente dettagliati per tenere conto in maniera esplicita degli altri fattori ambientali che influenzano il rischio di TT.

La distribuzione spaziale dei casi di TT può essere descritta tramite il processo puntuale di Poisson non omogeneo, poiché la probabilità che si verifichi un evento, nel caso del TT, è molto bassa e la popolazione a rischio è invece molto ampia. In questo modello, il numero di eventi  $N(U)$  atteso per una generica sotto area  $U$  della regione oggetto di studio  $R$  segue una distribuzione di Poisson la cui intensità  $\lambda$  è variabile in funzione delle caratteristiche del territorio. Per un generico punto  $u$  avremo che  $\lambda = \lambda(u)$  e il numero atteso di eventi  $E[N(U)]$  è

$$E[N(U)] = \int_U \lambda(u) du. \quad (1)$$

In tal caso è possibile che zone limitrofe con simile densità e struttura di popolazione o in presenza (o meno) di altri fattori di rischio, diano vita a veri e propri gruppi (o *cluster*) di aree ad alto, medio e basso rischio di TT. L'analisi della similarità degli attributi di aree geografiche vicine rientra in generale nello studio dell'autocorrelazione spaziale, che valuta la distribuzione spaziale di un particolare processo in termini di relazioni, influenze reciproche e distanza (Cressie 1991; Anselin, Rey 2010; Borruso, Murgante 2012).

#### 3.2. Rapporto standardizzato di incidenza (RSI)

Il rischio di TT è stato rappresentato attraverso la produzione di mappe che riportano la distribuzione spaziale, per ogni sezione censuaria, del rapporto standardizzato di incidenza (RSI). Gli RSI sono stati calcolati per le singole sezioni censuarie abitate mediante standardizzazione indiretta (Waller e Gotway 2004, 12-15), utilizzando il tasso di incidenza di TT osservato nello stesso periodo (2003-2016) nell'intera Sicilia orientale. L'RSI è il rapporto tra casi di TT osservati e casi di TT attesi in ogni sezione censuaria  $i$

$$RSI = \frac{O_i}{E_i} \quad (2)$$

dove  $O_i$  è il numero dei casi osservati per la sezione censuaria  $i$  e  $E_i$  è il numero atteso dei casi nella medesima sezione censuaria  $i$ . Il numero atteso dei casi è calcolato come prodotto tra la popolazione a rischio (e quindi tutta la popolazione residente) nella data sezione censuaria  $i$  e il tasso generale di incidenza relativo all'intera area investigata

$$E_i = P_i r_+ \quad (3)$$

dove per l'appunto  $P_i$  è la popolazione a rischio nella specifica sezione  $i$  e  $r_+$  è il tasso generale di incidenza di TT, calcolato per le quattro province di interesse nel loro insieme, come

$$r_+ = \frac{O_+}{P_+} \quad (4)$$

dove  $O_+$  corrisponde al numero di casi di TT osservati e  $P_+$  è la popolazione residente nell'intera Sicilia orientale. Il pedice + indica che le variabili sono calcolate per la totalità dell'area di studio. Da qui deriva che l'RSI di una singola sezione censuaria viene così calcolato

$$RSI = \frac{O_i}{P_i r_+} \quad (5)$$

Quando le caratteristiche della popolazione ne determinano una suddivisione in strati con livelli di rischio differenti (ad esempio, per sesso e fascia d'età), è bene tenerne conto e dare un peso diverso ad ogni strato proprio in base al rischio di ciascuno di essi. In tal caso, invece di calcolare un tasso generale di incidenza  $r_+$  per tutta l'area di riferimento, si calcola un tasso diverso per ogni strato  $j$  come  $r_j = \frac{\sum_i O_{ij}}{\sum_i P_{ij}}$ . Da qui il numero atteso di casi nella sezione  $i$  è dato da  $E_i = \sum_j P_{ij} r_j$ .

È interessante confrontare la distribuzione osservata dell'RSI con quella teorica che si otterrebbe qualora l'allocazione dei casi di TT nel territorio fosse casuale, non dipendesse cioè da altre variabili oltre a quelle relative alla struttura per età e sesso della popolazione residente. Tale confronto consente di verificare, per ciascuna area, se il potenziale ruolo di altre variabili ambientali determini variazioni significative del rischio. Il test inferenziale che permette di stabilire la significatività statistica di tale differenza è basato sulla distribuzione di Poisson e, in particolare, ammette che sotto l'ipotesi nulla l'RSI abbia un valore pari a 1 e che un valore osservato di RSI significativamente maggiore di 1 implichi un aumento del rischio (Bivand et al. 2008, 320-323). Poiché i casi di TT si distribuiscono secondo una Poisson (Banerjee et al. 2004, 150-152) è possibile costruire gli intervalli di confidenza al 95% del RSI con metodo esatto, usando la funzione "pois.

exact" del pacchetto "epitools" contenuto nel software R (R Core Team 2020; Aragon 2020).

L'indice RSI presenta dei limiti in termini di variabilità: in presenza di eventi rari, come l'insorgenza del TT, molto spesso per le sotto-aree più piccole il valore atteso dei casi è sempre minore di 1 e quando si verifica anche un solo caso in una sotto-area piccola, essa risulta avere un RSI significativamente superiore a 1. Per cui nello studio delle patologie rare, ci si attende a priori che anche nel caso di una distribuzione dell'RSI completamente casuale ci siano molte aree che risulteranno essere significative, anche se la maggior parte di esse non lo sono. Per superare questa problematica saranno considerate solamente le sezioni censuarie con più di 30 residenti per il calcolo degli RSI. Al contrario, nel computo del numero atteso globale dei casi per singolo strato, sarà tenuta in considerazione la totalità dei casi di TT e della popolazione residente nelle province di interesse.

Occorre, infine, evidenziare che l'RSI considera ogni unità spaziale, nel presente caso le sezioni censuarie, in maniera indipendente dalla loro collocazione geografica, pertanto a corredo della mappatura di tale indice, vengono implementati la funzione K di Ripley non omogenea e l'indice locale di Moran I. La prima è utile per identificare la presenza di concentrazione residuale rispetto alla disposizione casuale a diverse scale geografiche, mentre il secondo permette di visualizzare le aree di maggior rischio.

### 3.3. Funzione K di Ripley non omogenea

La funzione K di Ripley descrive la configurazione spaziale di un processo puntuale su una scala spaziale variabile (Ripley 1977) e viene definita come

$$K(d) = \lambda^{-1} E[N_0(d)] \quad (6)$$

in cui  $\lambda$  è l'intensità, assunta come costante, del processo e  $E[N_0(d)]$  è il numero atteso di eventi aggiuntivi in un intorno di raggio  $d$  centrato su un evento arbitrario. In particolare, la funzione K ha lo scopo di evidenziare quando la distribuzione spaziale si distacca da quella attesa sotto condizioni di casualità. Un processo puntuale può distribuirsi casualmente, in maniera concentrata o dispersa e la funzione K è in grado di rilevarlo come illustrato nella Figura 3, in cui la funzione K fuoriesce dalle bande di confidenza alle distanze in cui il processo puntuale diparte dalla disposizione casuale.

Nella maggior parte delle applicazioni, tuttavia, l'ipotesi di una intensità costante risulta essere troppo stringente. La posizione dei centri abitati, l'organizzazio-

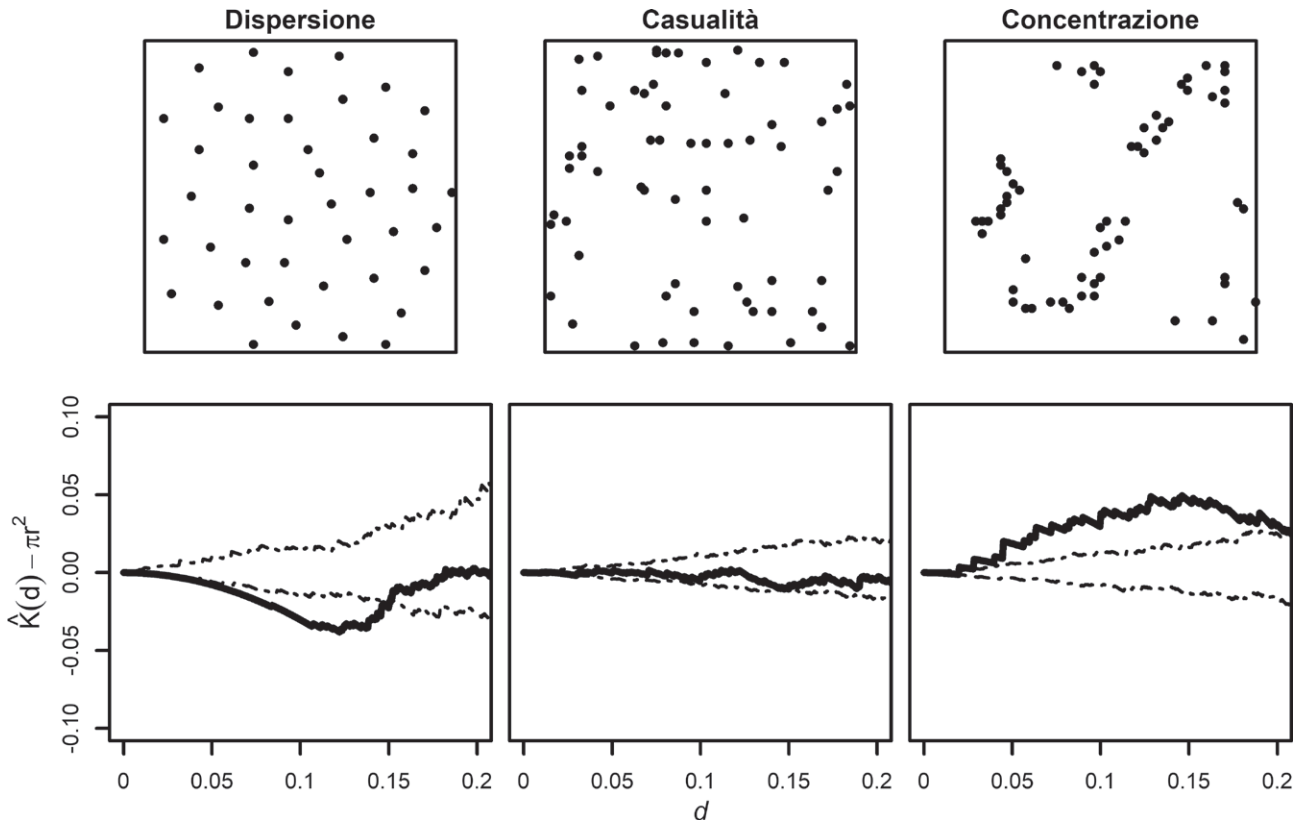


Figura 3. Analisi di tre differenti processi puntuali mediante funzione K di Ripley. Fonte: elaborazione degli autori.

ne delle reti di trasporto e dei siti industriali, la destinazione di campi ad uso agricolo e la morfologia del territorio sono tutti fattori sociali, economici e geografici che fungono da punti di attrazione o dispersione della popolazione sul territorio (Small 2004). In questi casi, la funzione K segnalerebbe la presenza di una deviazione da un'allocazione casuale degli eventi nello spazio derivante dal fatto che la popolazione tende a distribuirsi in maniera eterogenea e non per la presenza di specifici fattori di rischio o di relazioni di dipendenza tra gli eventi (ad esempio, il contagio per le malattie infettive o l'imitazione per i comportamenti e gli stili di vita). Per ovviare a tale problematica, Baddeley, Møller e Waagepetersen (2000) hanno proposto una generalizzazione della funzione K per processi non omogenei, cioè quelli nei quali l'intensità non è costante all'interno dello spazio. La funzione K di Ripley non omogenea è così definita

$$K_{inhom}(d) = E\left[\sum_{x \in X} \frac{1}{\lambda(x)} \mathbf{1}(0 < \|u-x\| \leq d) \mid u \in X\right] \quad (7)$$

dove  $\lambda(x)$  è l'intensità nel punto  $x$ ,  $\|u-x\|$  è la distanza euclidea tra gli eventi  $u$  e  $x$ , e  $\mathbf{1}(\cdot)$  è la funzione indicatrice.

Baddeley e Turner (2000) hanno proposto il seguente stimatore per la  $K_{inhom}(d)$

$$\hat{K}_{inhom}(d) = \frac{1}{|R|} \sum_i \sum_{j \neq i} \frac{1(\|x_i - x_j\| \leq d)}{\hat{\lambda}(x_i)\hat{\lambda}(x_j)} e(x_i, x_j; d) \quad (8)$$

in cui  $e(x_i, x_j; d)$  è un fattore di correzione per i bordi (Ripley 1991, 22) e  $\hat{\lambda}(x)$  è la stima della funzione di intensità  $\lambda(x)$ .

Poiché la distribuzione di  $\hat{K}_{inhom}(d)$  non è nota, per implementare procedure inferenziali è possibile ricorrere a simulazioni Monte Carlo con cui testare l'ipotesi nulla di distribuzione casuale nello spazio. Replicando  $n$  volte un processo puntuale di Poisson non omogeneo, si calcolano altrettanti valori di  $\hat{K}_{inhom}^j(d)$ ,  $j = 1, \dots, n$ , uno per ogni replica, e si ottengono le conseguenti bande puntuali di confidenza

$$L(d) = \hat{K}_{inhom}^{(k)}(d) \quad (9)$$

$$U(d) = \hat{K}_{inhom}^{(n-k)}(d) \quad (10)$$

di ampiezza  $\alpha = 2k/(n+1)$ , in cui  $\hat{K}_{inhom}^{(k)}(d)$  denota la  $k$ -esima statistica ordinata. I valori osservati di  $\hat{K}_{inhom}(d)$

che cadono al di fuori delle bande di confidenza ad una particolare distanza  $d$ , indicano un allontanamento dall'ipotesi nulla e quindi la presenza di concentrazione o dispersione, con un livello di significatività  $\alpha$ .

La stima dell'intensità  $\lambda(u)$  del processo è basata sulla popolazione residente al censimento del 2011. I dati censuari sono disponibili per sezione censuaria; abbiamo preferito ricalcolare la popolazione residente per una griglia composta da pixel quadrati di lato 50 km. La popolazione per ciascun quadrato è stata stimata per mezzo dell'algoritmo di interpolazione pycnophylactica (dal greco  $\pi\upsilon\kappa\nu\acute{o}\varsigma$  = massa e  $\phi\upsilon\lambda\acute{\alpha}\tau\tau\omega$  = che preserva) proposto da Tobler (1979) e appositamente concepito per questo tipo di applicazioni, che presenta il vantaggio di preservare la popolazione complessiva di ciascuna area (*ibi-dem*). Infine, abbiamo calcolato il numero di casi attesi per singolo pixel sulla base del tasso generale di incidenza osservato, applicando l'equazione (3). L'intero codice R elaborato dagli autori per la stima dell'intensità e della funzione K di Ripley è riportato in Tabella 4.

### 3.4. Indicatore locale Moran I

L'indicatore locale Moran I appartiene ai cosiddetti LISA (*Local Indicators of Spatial Association*), o indicatori locali di autocorrelazione spaziale, proposti da Anselin (1995). È calcolato con la seguente formula

$$I_i = \frac{(y_i - \bar{y})}{S_i^2} \sum_{j=1, j \neq i}^n (w_{ij} (y_j - \bar{y})) \quad (11)$$

in cui  $n$  è il numero delle unità geografiche,  $y_i$  è il valore della variabile  $y$  nella regione  $i$ ,  $\bar{y}$  è la media campionaria della variabile,  $y_j$  è il valore della variabile  $y$  nella regione  $j$ ,  $S_i^2$  è la varianza campionaria della variabile  $y$  e  $w_{ij}$  è un peso che può essere definito come l'inverso della distanza tra le varie unità geografiche. Esistono altri modi per definire  $w_{ij}$ , alcuni contemplano la scelta di una distanza limite per definire il vicinato di una data regione: le regioni che ricadono entro la distanza limite assumono un peso positivo, mentre quelle esterne assumono peso pari a zero.

Valori positivi dell'indice locale di Moran per una data regione indicano che quest'ultima è circondata da regioni che presentano perlopiù valori simili della variabile oggetto di studio (più precisamente, scarti dalla media generale concordi). In tal caso si evidenzia la presenza di gruppi spaziali, definiti come "alto-alto" (regione con un valore elevato circondata da regioni con valori elevati) o "basso-basso" (regione con valore basso circondata da regioni con valore basso). Parlando in termini di rischio di tumore, un *cluster* "alto-alto" indiche-

**Tabella 4.** Codice R per la stima della funzione K di Ripley non omogenea. Fonte: elaborazione degli autori.

---

```
#Caricamento delle librerie necessarie e generazione dei numeri
pseudo casuali:
library(spatstat)
library(pycno)
library(raster)
library(maptools)
set.seed(1234)

#Creazione di una mappa della popolazione residente per singolo
quadrato:
m <- matrix(rbinom(100, 1000, 0.3), nr=10)

m[1,1] = NA
m[5,2] = NA
m[5,3] = NA

r <- raster(m, xmn=0, xmx=10, ymn=0, ymx=10)
plot(r)
p <- rasterToPolygons(r)

plot(p)

#Algoritmo di interpolazione picnoflattica di Tobler per la stima
della densità della popolazione:
pycnocm = pycno::pycno(p, pops=p$layer, 0.05, celldim=0.01,
converge=1)

#Conversione al formato raster:
rpycnocm = raster(pycnocm)

#Conversione al formato im:
impycnocm = maptools::as.im.RasterLayer(rpycnocm)

#Creazione dataset casi di TT:
casiTT = data.frame(ID = c(1:50),
long = runif(50, 0, 10),
lat = runif(50, 0, 10))

#Calcolo del numero di casi attesi per ogni pixel:
rho = length(casiTT$ID)/sum(as.matrix.im(impycnocm),
na.rm=TRUE)
s = summary(impycnocm)
dens_impycnocm = eval.im(impycnocm*rho/(s$Xstep*s$Ystep))
summary(dens_impycnocm)$integral

#Conversione del dataset casi di TT in processo puntuale planare
(ppp):
m = unionSpatialPolygons(p, rep(0, length(p@polygons)))
W = maptools::as.owin.SpatialPolygons(W=m)
ppp = as.ppp(cbind(casiTT$long, casiTT$lat), W=W)

#Calcolo dell'intensità del processo nei punti in cui ricadono le
coordinate degli indirizzi dei casi di TT:
lambda = safelookup(Z=dens_impycnocm, x=ppp)
Stima della funzione K di Ripley non omogenea:
K = Kinhom(X=ppp, lambda=lambda, renormalise=TRUE,
correction="trans", normpower=1)
```

---

rebbe una zona ad alto rischio, mentre un *cluster* “basso-basso” denoterebbe un’area di basso rischio. Valori negativi elevati del Moran I locale rilevano che la regione in esame è un *outlier* spaziale. Un *outlier* spaziale è un’area che riporta un valore marcatamente diverso da quello dei suoi vicini (Cerioli, Riani 1999). Gli *outlier* spaziali determinano la formazione di due tipi di *cluster*: “alto-basso”, in presenza di un valore elevato circondato da vicini con valore basso, oppure “basso-alto” in presenza di un valore basso circondato da vicini con valori elevati.

Per testare la significatività statistica dell’indice locale di Moran I sono possibili due approcci. Il più immediato si basa su un’assunzione di normalità che non sempre può essere verificata, soprattutto in presenza di dati fortemente asimmetrici (Anselin 1995). Il secondo approccio è basato su tecniche di ricampionamento, come la permutazione condizionata, che non presuppongono assunzioni sui dati (*ibidem*). Secondo questo approccio, quando viene valutato il valore di un attributo di una data regione, il suo valore è tenuto fisso e tutti gli altri valori (delle altre regioni) vengono permutati in modo casuale su tutte le altre posizioni, senza ripetizione. Ogni volta che gli altri valori vengono permutati, l’indice locale Moran I viene calcolato per formare una distribuzione empirica di riferimento. Il pvalue (denominato in questo caso “pseudo pvalue”) può essere stimato confrontando l’indice effettivamente osservato sui dati con la distribuzione empirica creata tramite le permutazioni (Anselin 2005). Il singolo pseudo pvalue è calcolato come  $(M + 1) / (R + 1)$  dove R è il numero di permutazioni e M è il numero di istanze in cui una statistica calcolata tramite permutazioni è uguale o maggiore del valore osservato (per valori positivi dell’indice) o minore o uguale al valore osservato (per valori negativi dell’indice). In questo studio, tutti gli indici locali Moran I sono stati testati utilizzando 999 permutazioni e il livello di significatività è stato scelto pari a 0.05. Il calcolo del Moran I locale e la relativa analisi della distribuzione spaziale sono stati implementati tramite l’apposita funzione “local\_moran” contenuta nella libreria “rgeoda” per il software R (Li, Anselin 2021).

Le analisi statistiche sono state condotte a livello di singola sezione censuaria e quindi utilizzando dati areali per l’RSI e per il Moran locale poiché i dati a disposizione riguardanti la popolazione residente sono riferiti appunto alle sezioni censuarie. Al contrario, per quanto riguarda la funzione K di Ripley sono stati usati dati puntuali, e quindi a livello di singolo caso di TT.

Per utilizzare in modo appropriato i dati riguardanti la popolazione, aggregati secondo sezioni amministrative irregolari, è importante soffermarsi sulla specifica-

zione del vicinato (come definito dalla matrice dei pesi spaziali). Ad oggi non esiste in letteratura un metodo empirico o un chiaro fondamento teorico che siano da guida nella scelta della “corretta” matrice dei pesi spaziali che possono essere di tipo binario (peso uguale a 1 per la sezione che fa parte del vicinato e pari a 0 altrimenti) oppure continuo, ad esempio inversamente proporzionali alla distanza tra le sezioni, basati su kernel (Anselin, Bera 1998); per tale motivo si suggerisce di sperimentare diverse tipologie della matrice. In questo studio il sistema di pesi spaziali applicato appartiene alla famiglia dei kernel gaussiani. In particolare, sono stati fatti molteplici tentativi di mappatura dell’indice locale Moran I facendo variare il *bandwidth* in modo da tener conto del cosiddetto “problema delle aree modificabili” (*Modifiable Areal Unit Problem* – MAUP). Il MAUP, ampiamente discusso in Openshaw (1981), risulta dalla combinazione dell’effetto scala e dell’effetto aggregazione per cui statistiche basate su scale geografiche differenti o su diverse aggregazioni di unità areali alla medesima scala potrebbero generare risultati differenti e quindi difficilmente interpretabili. Pertanto, le analisi dell’indice locale Moran I sono state effettuate facendo variare il *bandwidth* e considerandone le lunghezze pari a 5 km, 10 km e 20 km.

#### 4. Risultati e discussione

Nella Sicilia orientale dal 2003 al 2016 sono stati 7.085 gli individui colpiti da TT. L’eziologia di questo tumore è complessa e variegata, comprende concause di natura genetica ma anche alimentare come il consumo di iodio, la familiarità e l’esposizione a radiazioni ionizzanti già menzionati.

La distribuzione dei casi di TT è influenzata dalla disposizione spaziale e dalla struttura per età e sesso della popolazione residente. La popolazione femminile risulta essere notoriamente più colpita dal TT (Parkin et al. 2005); Inoltre, l’invecchiamento della popolazione tende ad accentuare lo squilibrio tra i sessi, data la maggiore aspettativa di vita delle donne (*World Population Ageing* 2017). In Sicilia orientale, le donne costituiscono il 58% della popolazione sopra i 70 anni, pertanto nelle aree in cui i processi di spopolamento hanno un’intensità maggiore la popolazione anziana risulta essere preponderante e di conseguenza lo squilibrio fra i sessi emerge in maniera accentuata in favore delle donne perché mediamente più longeve rispetto agli uomini. La presenza di fattori ambientali correlati alla natura vulcanica di una parte del territorio può costituire un rilevante fattore di rischio. I fumi emessi da un vulcano attivo,

ad esempio, sono in grado di trasportare metalli pesanti e sostanze radioattive capaci di contaminare l'aria, l'acqua e il suolo delle aree circostanti (Fiore et al. 2019).

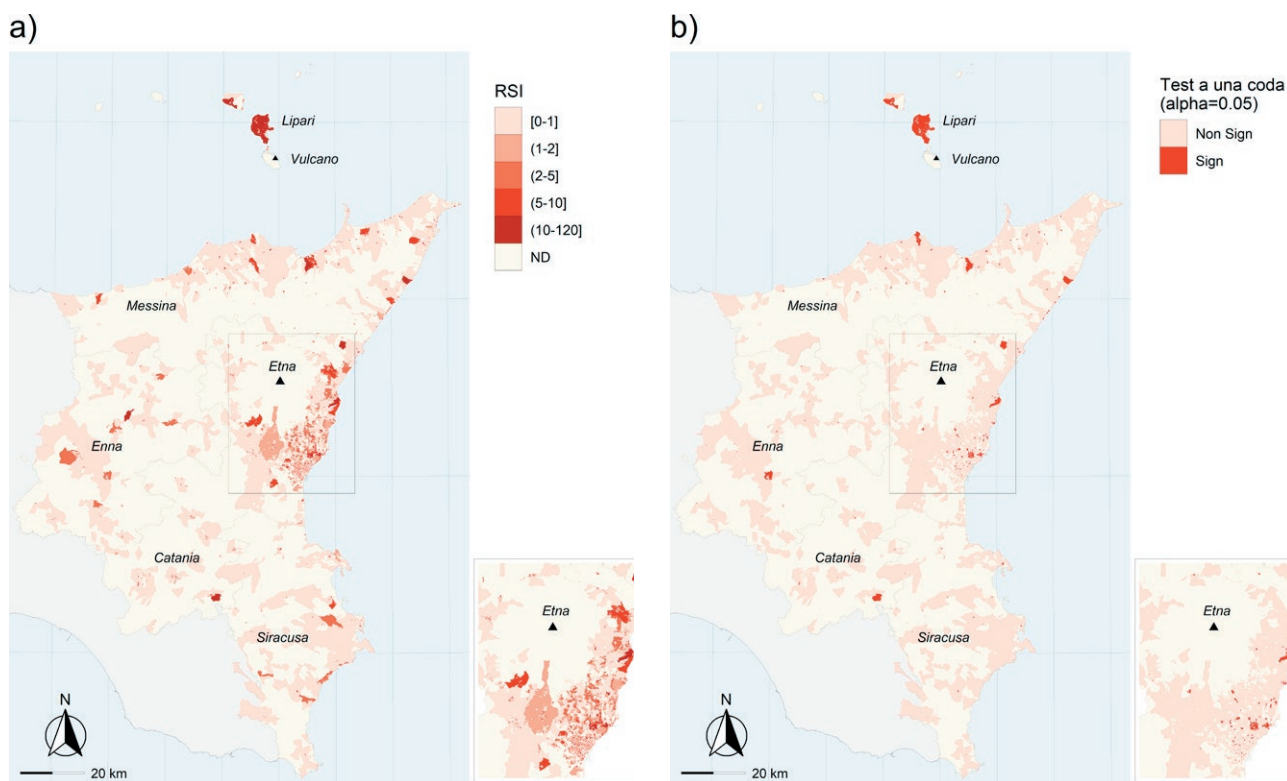
Nel tentativo di distinguere gli effetti delle due componenti, geografica e demografica, sulla distribuzione spaziale dei casi di TT, proponiamo le mappe dell'RSI per sezione censuaria e quelle della loro significatività statistica. Gli RSI, le cui statistiche descrittive sono riportate in Tabella 5, sono stati costruiti suddividendo la popolazione in strati sulla base dell'età e del sesso, per riflettere la variazione del rischio di TT dovuta a queste due variabili demografiche. Pertanto, è stato calcolato un diverso tasso di rischio globale per ogni strato (si veda sezione 3.1). Dalla tabella 5 appare evidente la notevole variabilità dell'indice tra le varie sezioni, evidenziata dal range tra i valori massimo e minimo, e dalla varianza. Inoltre, la media maggiore della mediana e l'asimmetria positiva evidenziano la presenza di numerose sezioni censuarie con un valore osservato di casi superiore rispetto a quello atteso.

Le Figure 4a-b riportano rispettivamente gli RSI e il risultato del test a una coda con livello di significatività  $\alpha=0.05$ , condotto per evidenziare le sezioni censuarie

**Tabella 5.** Statistiche descrittive dell'RSI.

Statistica	Valore
Minimo	0.102
Massimo	225.114
Mediana	1.754
Media	2.856
Varianza	62.771
Indice di asimmetria	18.624
Curtosi	429.520

con RSI significativamente maggiore di 1. Si può notare che dal mero computo degli RSI (Fig. 4a) emergono diverse aree di rischio. Tali zone sono localizzate nell'area circumetnea ma anche nelle province non vulcaniche, soprattutto in quelle di Enna e Messina. Il test a una coda (Fig. 4b) mette, invece, in evidenza l'area a sud-est dell'Etna e diverse sezioni appartenenti soprattutto alla provincia messinese. In entrambe le mappe, ciò che appare evidente è che se nelle province non vulcaniche le sezioni censuarie con RSI maggiore di 1 si dispongono in ordine sparso sul territorio, nella provincia catanese le sezioni di rischio sembrerebbero trovare

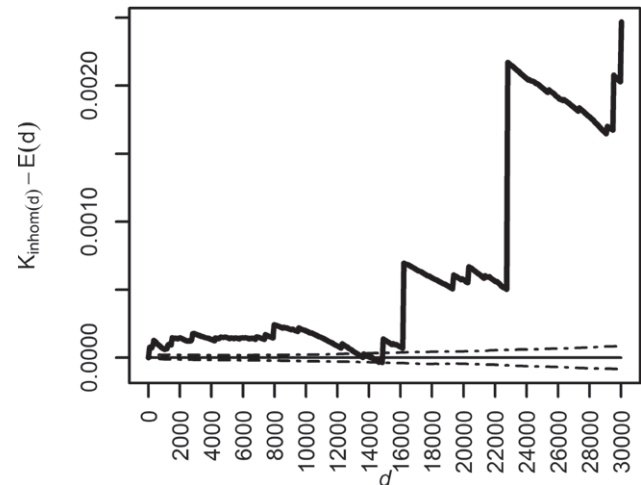


**Figura 4.** Distribuzione dell'RSI per sezione censuaria (a) e rappresentazione della sua significatività statistica (b). Fonte: elaborazione originale degli autori.

quasi tutte collocazione in una zona prossima all'Etna, lasciando quasi libero il resto della provincia. Inoltre, la collocazione delle zone a rischio lungo l'asse NO-SE potrebbe far supporre che la presenza di venti persistenti in direzione SE, documentata da Boffetta et al. (2020), possa trasportare le sostanze tossiche emesse dal vulcano e quindi inquinare l'atmosfera dei territori posizionati lungo l'asse stesso. È interessante notare, inoltre, che le sezioni censuarie sull'isola di Lipari mostrano un RSI significativamente elevato. Anche questa zona infatti è di tipo vulcanico e si trova nelle immediate vicinanze di Vulcano che è ad oggi attivo, sebbene con un'attività esigua rispetto a quella etnea. L'isola di Vulcano con il suo cratere chiamato La Fossa è sede di numerose fumarole solforose nonché campo di frequenti emissioni vulcaniche sottomarine di  $\text{CO}_2$ , la cui distribuzione spaziale segue la direzione conferita da venti persistenti che spirano da NO (Vizzini et al. 2020). Sempre in quest'area inoltre Vizzini et al. (2013) hanno rinvenuto la presenza di contaminazione, seppur definita "bassa", da elementi come Bario, Ferro, Arsenico e Cadmio. È documentato che i metalli pesanti siano tossici e possano indurre effetti avversi per l'organismo umano e che, in particolare, arsenico e cadmio svolgano un ruolo nella riduzione dei livelli di ormoni tiroidei (Vigneri et al. 2017).

Come già menzionato, un limite dell'analisi degli RSI proposta è che questa è riferita alle singole sezioni censuarie in maniera indipendente dalla loro posizione geografica. Ad esempio, le carte in Figura 4 rivelano la presenza di sezioni censuarie con valori significativamente elevati di RSI anche lontano dalle aree vulcaniche. Tali sezioni appaiono disposte in maniera sparsa nel territorio, tanto da sembrare compatibili con un processo di collocazione casuale. Viceversa, nelle aree vicine a quelle vulcaniche, le sezioni con RSI elevato mostrano dei pattern concentrati. L'RSI però, non è in grado di valutare la presenza di zone di alto o basso rischio in funzione della loro collocazione geografica. Le due analisi seguenti, quella basata sulla funzione K e quella sull'indice locale di Moran considerano invece in maniera esplicita la distribuzione spaziale dei casi.

La Figura 5 riporta l'andamento della funzione  $\hat{K}_{inhom}(d)$  e le bande di confidenza riferite all'ipotesi nulla di disposizione spaziale casuale con un livello di significatività pari ad  $\alpha = 0.05$  e basate su 999 simulazioni Monte Carlo. Dal punto di vista ecologico, molti processi puntuali mostrano andamenti diversi al variare della scala: concentrazione a grandi scale e regolarità a piccole scale. La combinazione di simili trend può essere messa in evidenza dal grafico della funzione K di Ripley (Dixon 2002). I casi di TT mostrano uno scostamento statisticamente significativo dalla casualità fino a circa



**Figura 5.** Stima della funzione K generalizzata. La funzione K è rappresentata dalla curva continua nera. Le corrispondenti bande di confidenza al 95% sotto l'ipotesi nulla di assenza di interazione, sono basate su 999 simulazioni Monte Carlo. I valori sono centrati rispetto alla media  $E(d)$  delle simulazioni. Le distanze sono espresse in metri. Fonte: elaborazione degli autori.

14 km. Oltre questa distanza la concentrazione cresce notevolmente indicando la presenza di raggruppamenti di casi di TT sul territorio. La funzione K non omogenea tiene conto della distribuzione eterogenea della popolazione sul territorio attraverso l'intensità  $\lambda(x)$ . Pertanto, la disposizione dei casi di TT in gruppi rilevata dalla funzione K non è dovuta all'eterogeneità della popolazione residente ma deve essere attribuita all'azione di ulteriori fattori ambientali e/o comportamentali.

Per visualizzare la localizzazione di eventuali raggruppamenti di casi di TT sull'area in esame e per valutarne visivamente la disposizione in relazione alla vicinanza dell'Etna, è stato calcolato l'indice locale di Moran I sugli RSI. Nelle Figure 6a-c-d è rappresentata la statistica per la totalità dei casi di TT con un *bandwidth* rispettivamente pari a 5 km, 10 km e 20 km. Nelle Figure 6b-d-f sono riportati gli pseudo pvalue ottenuti dalle permutazioni condizionate. Dalle mappe del Moran locale appaiono due zone di alto rischio: a sud-est rispetto all'Etna e a nord, tra le isole Eolie e nei pressi di Barcellona Pozzo di Gotto. Le altre aree risultano essere di rischio basso o non significativo. Aumentando il *bandwidth*, e quindi ingrandendo i vicinati considerati, si osserva un progressivo ampliamento dell'area di alto rischio ai piedi dell'Etna e di quella di basso rischio a sud, nel siracusano. Al contrario, la zona di alto rischio a nord, osservata con un *bandwidth* di 5 e 10 km (Figg. 6a-c), si riduce notevolmente aumentando la grandezza del vicinato (Fig. 6e). Dalle figure che riportano i pvalue (Figg. 6b-d-f) risulta che le

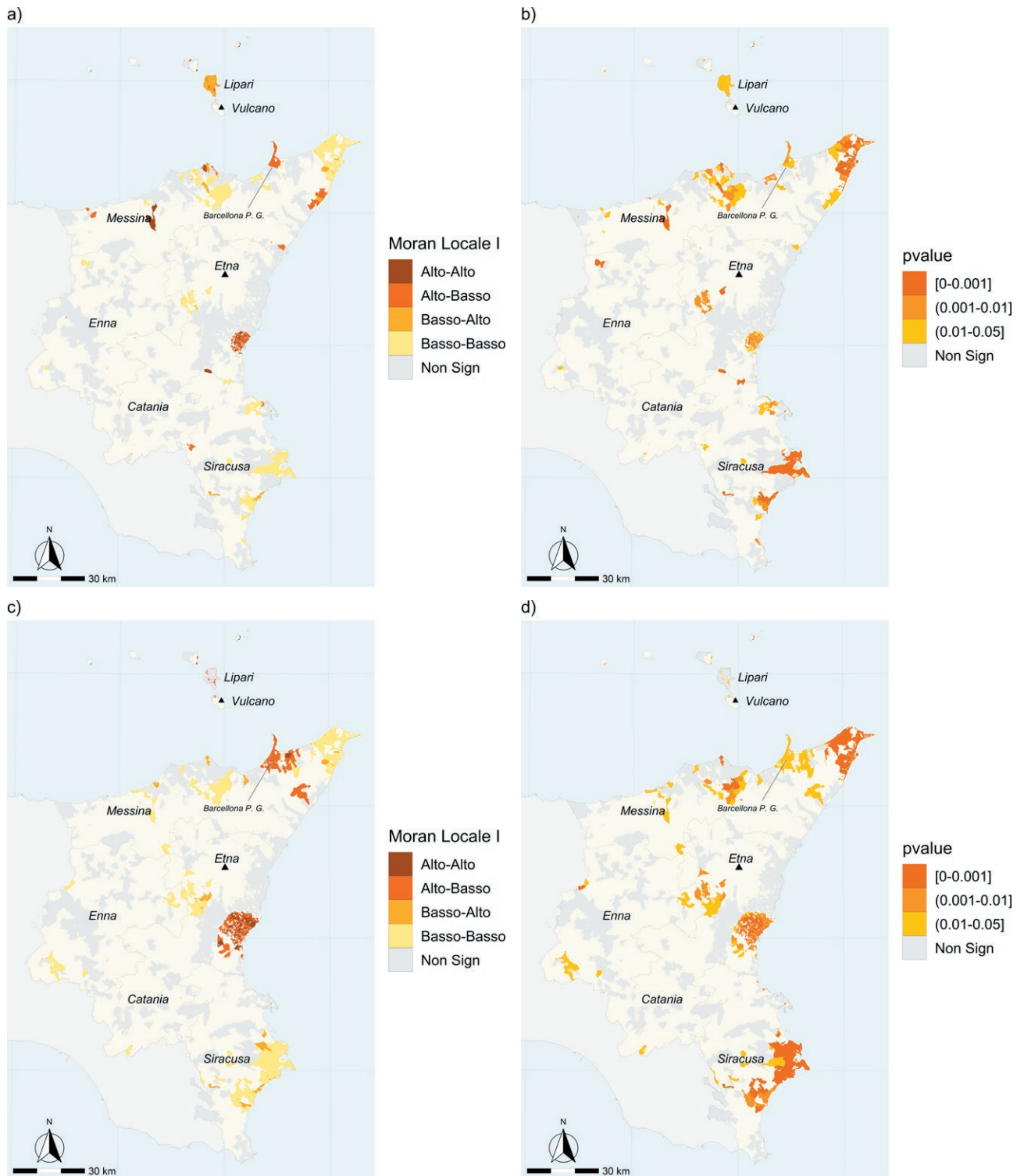


sezioni costituenti i gruppi di alto e basso rischio sono significative ad un livello pari al più ad  $\alpha = 0.05$ . Occorre notare, infine, che la gran parte delle sezioni considerate è risultata essere di rischio non significativo, come si evince dalle estese zone grigie presenti in tutte le mappe. In particolare, la presenza della zona di alto rischio di TT a sud-est dell'Etna potrebbe confermare l'ipotesi secondo cui venti persistenti in direzione SE spingerebbero le sostanze radioattive emesse dal vulcano proprio verso le zone che registrano un rischio elevato. Un suggerimento simile sembrerebbe porsi per le isole Eolie e le sezioni nei pressi di Barcellona Pozzo di Gotto.

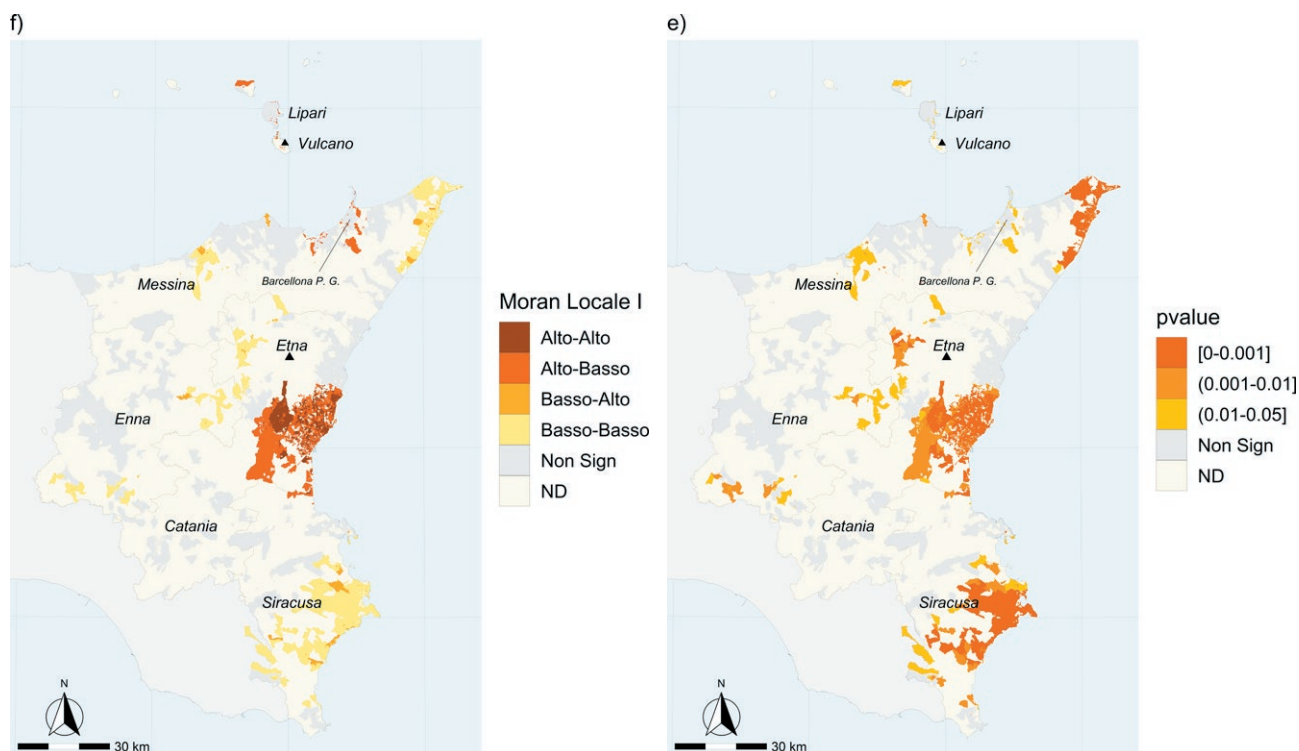
## 5. Conclusioni

Lo studio della diffusione geografica delle malattie infettive vanta una tradizione consolidata. La crescente incidenza di malattie cronico-degenerative (per lo più tumori e malattie cardiovascolari) ha portato all'applicazione, anche in questo ambito, delle metodologie tipiche per lo studio della diffusione delle infezioni. Quando tra le concause di simili patologie, come il TT, vengono annoverati dei fattori ambientali, l'analisi geografica costituisce un passo fondamentale per ottenere maggiore comprensione della distribuzione del rischio e dell'incidenza. Uno dei fattori ambientali che spesso viene citato come possibile causa dell'insorgenza del TT è la presenza di un vulcano attivo. In diverse regioni vulcaniche nel mondo molteplici studi effettuati sui dati dei registri tumori locali hanno riportato un aumento significativo dell'incidenza di TT (si veda sezione 1). La disponibilità dei GIS ha permesso di integrare dati ambientali e sanitari eterogenei e di diverso grado di aggregazione spaziale (puntuale per i casi di TT e areale per i dati riferiti all'intera popolazione). Inoltre, la compatibilità tra GIS e la piattaforma di calcolo statistico adoperata ha reso agevole l'implementazione di metodologie di analisi spaziale molto complesse sia da un punto di vista analitico che computazionale. Queste tecnologie facilitano la comprensione di problematiche di salute pubblica, come la distribuzione del TT, e favoriscono la collaborazione tra studiosi di ambiti differenti, quali medicina, epidemiologia e geografia, per mappare e prevedere il rischio di malattia. Nello specifico, questo lavoro ha portato alla mappatura dei casi di TT nella Sicilia orientale con il fine di visualizzare le zone di rischio e metterle in relazione con la vicinanza all'Etna. I dati sanitari analizzati sono stati rilasciati dal Registro Tumori della Sicilia Orientale. La geocodifica degli indirizzi dei casi di TT ha reso possibile lavorare a livello puntuale per evidenziare l'eventuale concentrazione residua per mez-

zo della funzione K di Ripley non omogenea. Secondo la presente analisi empirica il rischio di TT non si distribuisce casualmente nell'area di interesse, ma presenta delle discontinuità che fanno emergere zone localizzate di alto rischio in un contesto che appare in generale di basso rischio o rischio non significativo. In particolare, le aree di alto rischio significativo sono emerse dopo aver controllato i fattori demografici che concorrono ad aumentare l'incidenza del TT, e cioè la popolazione residente in una specifica zona, la proporzione di popolazione femminile e l'età. Per quantificare il rischio è stato adottato l'RSI pesato per diversi strati della popolazione e calcolato mediante standardizzazione indiretta. Dalle prime mappe ottenute è stata rilevata una possibile zona di rischio significativo ai piedi dell'Etna. Tuttavia, i limiti intrinseci dell'RSI e dei suoi intervalli di confidenza potrebbero aver inciso sulla rilevazione delle aree di rischio. Inoltre, l'impossibilità di valutare i valori dell'RSI in relazione alla disposizione geografica ha reso necessario corredare la fase preliminare di mappatura dell'indice RSI con un'analisi della distribuzione spaziale del rischio. In primo luogo, è stata applicata la funzione K di Ripley generalizzata con bande di confidenza basate su 999 simulazioni Monte Carlo riferite all'ipotesi nulla di *randomness* ( $\alpha = 0.05$ ). Abbiamo riscontrato uno scostamento statisticamente significativo rispetto alla disposizione dei casi di TT che si avrebbe sotto l'ipotesi nulla di casualità. Dopodiché è stato calcolato l'indice locale Moran I sugli RSI ottenuti in precedenza al fine di creare delle mappe di *cluster* di alto e basso rischio, e di cambiamento del rischio. Tali mappe hanno evidenziato la presenza di gruppi di sezioni di alto rischio a SSE dell'Etna, nelle isole Eolie e nei pressi di Barcellona Pozzo di Gotto. Nel resto dell'area non sono emersi altri importanti *cluster* di zone ad alto rischio. La rilevazione di aree di maggior rischio localizzate nei pressi dell'Etna sembra avvalorare l'ipotesi che la presenza di un vulcano possa influenzare l'incidenza del TT. In aggiunta a ciò, anche le aree di rischio emerse nelle Eolie e lungo la costa settentrionale della Sicilia sembrano indicare una possibile influenza del vicino Vulcano. I risultati ottenuti mediante mappatura dell'RSI e dell'indice locale Moran I sembrano registrare le medesime aree di alto e basso rischio con una variabilità maggiore nell'RSI probabilmente dovuta alla sua instabilità intrinseca. La funzione K di Ripley non omogenea invece riporta un andamento dei casi di TT a gruppi che trova conferma nella mappa del Moran locale. L'analisi della distribuzione spaziale del rischio per mezzo dell'indice locale Moran I (Fig. 6) andrebbe interpretata nell'ottica della prima legge della geografia formulata da Tobler, secondo la quale le cose vicine tra loro sono maggiormente collegate tra loro



**Figura 6a-d.** Mappa dei cluster di rischio basati sull'indice Moran Locale I (a-b-c) e relativi pvalue (d-e-f) calcolati tramite permutazione condizionata per un bandwidth di 5 km (Figg. 6a-d), 10 km (Figg.6b-e) e 20 km (Figg.6c-f). Fonte: elaborazione originale degli autori.



**Figura 6e-f.** Mappa dei cluster di rischio basati sull'indice Moran Locale I (a-b-c) e relativi pvalue (d-e-f) calcolati tramite permutazione condizionata per un bandwidth di 5 km (Figg. 6a-d), 10 km (Figg.6b-e) e 20 km (Figg.6c-f). Fonte: elaborazione originale degli autori.

rispetto a quelle lontane (Sui 2004). La legge di Tobler è stata utilizzata come base per lo studio delle malattie, poiché le persone che vivono l'una accanto all'altra tendono ad essere esposte agli stessi fattori ambientali che favoriscono l'insorgenza di determinate patologie (Tlacuilo-Parra et al. 2017). Poiché, come già segnalato in letteratura, la prossimità a un vulcano attivo come l'Etna potrebbe concorrere all'eziologia del TT, occorrerebbe pianificare misure sanitarie di sicurezza. Questo aspetto dovrebbe essere di cruciale interesse per la sanità pubblica che potrebbe ottimizzare la distribuzione di servizi sanitari locali e mettere in atto mirate campagne di screening, monitoraggio e prevenzione sfruttando in maniera efficiente le risorse a disposizione. L'importanza di politiche localizzate sul territorio è stata ormai riconosciuta in numerosi ambiti (Fuchs 1984; Carter et al. 2000; Bradshaw, Muller 2004; Bell 2015). L'ausilio di nuovi software e routine in grado di collegare gli eventi di tumore ad un preciso punto nello spazio e le odierne tecniche di gestione di grandi moli di dati non possono dunque essere ignorate per favorire l'efficacia delle iniziative sanitarie per la lotta contro i tumori.

L'intento del presente lavoro è stato di tipo esplorativo, pertanto ulteriori analisi andrebbero condotte per

ottenere una comprensione più approfondita per distinguere con un dettaglio geografico e una precisione statistica maggiore i fattori ambientali che contribuiscono all'insorgenza del TT.

### Riferimenti bibliografici

Accordo di Partenariato 2014-2020. (2013). *Strategia nazionale per le Aree interne: definizione, obiettivi, strumenti e governance*. [https://www.miur.gov.it/documents/20182/890263/strategia\\_nazionale\\_aree\\_interne.pdf/d10fc111-65c0-4acd-b253-63efae626b19](https://www.miur.gov.it/documents/20182/890263/strategia_nazionale_aree_interne.pdf/d10fc111-65c0-4acd-b253-63efae626b19)

Alanis, P. K. B., Yamaya, Y., Takeuchi, A., Sasai, Y., Okada, Y., Nagao, T. (2013). A large hydrothermal reservoir beneath taal volcano (Philippines) revealed by magnetotelluric observations and its implications to the volcanic activity. *Proceedings of the Japan Academy Series B: Physical and Biological Sciences*, 89(8), 383–389. <https://doi.org/10.2183/pjab.89.383>

Anand, S., Staniec, A., Montez-Rath, M., Vlahos, P. (2020). Using GIS Mapping to Track Hot Spots of Kidney Disease in California. *New England Journal of*

- Medicine*, 382(23), 2265–2267. <https://doi.org/10.1056/nejmc2001023>
- Andronico, D., Spinetti, C., Cristaldi, A., Buongiorno, M. F. (2009). Observations of Mt. Etna volcanic ash plumes in 2006: An integrated approach from ground-based and polar satellite NOAA-AVHRR monitoring system. *Journal of Volcanology and Geothermal Research*, 180, 135–147. <https://doi.org/10.1016/j.jvolgeores.2008.11.013>
- Anselin, L. (1995). Local Indicators of Spatial Association-LISA. *Geographical Analysis*, 27, 93–115.
- Anselin, L. (2005). *Exploring Spatial Data with GeoDa: A Workbook*. <https://www.geos.ed.ac.uk/~gisteac/fspat/geodaworkbook.pdf>
- Anselin, L., Bera, A. K. (1998). Spatial dependence in linear regression models with an introduction to spatial econometrics. In Ullah A., Giles D. (Eds.), *Handbook of Applied Economic Statistics* (pp. 237–289). New York, Marcel Dekker.
- Anselin, L., Rey, S. J. (2010). *Perspectives on spatial data analysis. Perspectives on Spatial Data Analysis*. Berlin, Heidelberg, Springer.
- Aragon, T. J. (2020). epitools: Epidemiology Tools. R Package Version 0.5-10.1. <https://CRAN.R-Project.Org/Package=epitools>.
- Arnbjörnsson, E., Arnbjörnsson, A., Ólafsson, A. (1986). Thyroid cancer incidence in relation to volcanic activity. *Archives of Environmental Health*, 41(1), 36–40.
- Assessorato Regionale alla Salute. (2016). *Atlante Sanitario della Sicilia*. [https://www.epicentro.iss.it/tumori/pdf/ASS%20SALUTE\\_Atlante%20Tumori%202016.pdf](https://www.epicentro.iss.it/tumori/pdf/ASS%20SALUTE_Atlante%20Tumori%202016.pdf)
- Baddeley, A. J., Møller, J., Waagepetersen, R. (2000). Non- and semi-parametric estimation of interaction in inhomogeneous point patterns. *Statistica Neerlandica*, 54(3), 329–350. <https://doi.org/10.1111/1467-9574.00144>
- Baddeley, A., Turner, R. (2000). Practical maximum pseudolikelihood for spatial point patterns. *Australian and New Zealand Journal of Statistics*, 42(3), 283–322. <https://doi.org/10.1017/s000186780004698x>
- Banerjee, S., Carlin, B. P., Gelfand, A. E. (2004). *Hierarchical Modeling and Analysis for Spatial Data*. Boca Raton/ London, Chapman & Hall/CRC.
- Bell, M. (2015). Demography, time and space. *Journal of Population Research*, 32, 173–186. <https://doi.org/10.1007/s12546-015-9148-6>
- Berke, E. M. (2010). Geographic information systems (GIS): Recognizing the importance of place in primary care research and practice. *Journal of the American Board of Family Medicine*, 23(1), 9–12. <https://doi.org/10.3122/jabfm.2010.01.090119>
- Bertazzon, S. (2009). Malattie cardiovascolari e fattori socioeconomici di rischio: un'analisi spaziale empirica su Calgary (Canada). *Rivista Geografica Italiana*, 116(3), 379–400.
- Bivand, R. S., Pebesma, E., Gómez-Rubio, V. (2008). *Applied Spatial Data Analysis with R*. New York, Springer.
- Boffetta, P., Memeo, L., Giuffrida, D., Ferrante, M., & Sciacca, S. (2020). Exposure to emissions from Mount Etna (Sicily, Italy) and incidence of thyroid cancer: a geographic analysis. *Scientific Reports*, 10(21298).
- Borruso, G., Murgante, B. (2012). Analisi dei fenomeni immigratori e tecniche di autocorrelazione spaziale. Primi risultati e riflessioni. *Geotema*, 43-44-45, 182-189.
- Bradshaw, T. K., Muller, B. (2004). Shaping policy decisions with spatial analysis. In Janelle G. M. F, D. G. (Eds.), *Spatially Integrated Social Science: examples in best practice*. New York, Oxford University Press.
- Bray, F., Colombet, M., Mery, L., Piñeros, M., Znaor, A., Zanetti, R., & Ferlay, J. (Eds.). (2017). *Cancer incidence in five continents Vol. XI*. Lyon: International Agency for Research on Cancer.
- Bristow, R. E., Chang, J., Ziogas, A., Gillen, D. L., Bai, L., & Vieira, V. M. (2015). Spatial analysis of advanced-stage ovarian cancer mortality in California. *American Journal of Obstetrics and Gynecology*, 213(1), 43.e1-43.e8. <https://doi.org/10.1016/j.ajog.2015.01.045>
- Buat-Ménard, P., Arnold, M. (1978). The heavy metal chemistry of atmospheric particulate matter emitted by Mount Etna Volcano. *Geophysical Research Letters*, 5(4), 245–248. <https://doi.org/10.1029/GL005i004p00245>
- Caguioa, P. B., Bebero, K. G. M., Bendebel, M. T. B., Saldana, J. S. (2019). Incidence of thyroid carcinoma in the Philippines: A retrospective study from a tertiary university hospital. *Annals of Oncology*, 30. <https://doi.org/10.1093/annonc/mdz428.024>
- Caltabiano, T., Burton, M., Giammanco, S., Allard, P., Bruno, N., Murè, F., Romano, R. (2004). Volcanic gas emissions from the summit craters and flanks of mt. Etna, 1987-2000. *Geophysical Monograph Series*, 143, 111–128. <https://doi.org/10.1029/143GM08>
- Carter, R., Mendis, K. N., Roberts, D. (2000). Spatial targeting of interventions against malaria. *Bulletin of the World Health Organization*, 78(12), 1401–1411.

- Cerioli, A., Riani, M. (1999). The Ordering of Spatial Data and the Detection of Multiple Outliers. *Journal of Computational and Graphical Statistics*, 8(2), 239–258. <https://doi.org/10.1080/10618600.1999.10474812>
- Cimino, G., Ziino, M. (1983). Heavy metal pollution. Part VII. Emissions from Mount Etna volcano. *Geophysical Research Letters*, 10(1), 31–34.
- Cooley, D. (2020). googleway: Accesses Google Maps APIs to Retrieve Data and Plot Maps. R package version 2.7.3. Retrieved from <https://cran.r-project.org/package=googleway>
- Craglia, M., & Maheswaran, R. (2004). *GIS in public health practice: opportunities and pitfalls*. Boca Raton, FL, CRC Press.
- Cressie, N. (1991). *Statistics for Spatial Data*. New York, Wiley.
- Croner, C. M., Sperling, J., Broome, F. R. (1996). Geographic information systems (GIS): New perspectives in understanding human health and environmental relationships. *Statistics in Medicine*, 15(18), 1961–1977.
- Curado, M.P., Edwards, B., Shin, H. R., Storm, H., Ferlay, J., Heanue, M., Boyle, P. (2007). *Cancer incidence in five continents, Volume IX*.
- D'Aleo, R., Bitetto, M., Delle Donne, D., Tamburello, G., Battaglia, A., Coltelli, M., ... Aiuppa, A. (2016). Spatially resolved SO<sub>2</sub> flux emissions from Mt Etna. *Geophys. Res. Lett.*, 43(14), 7511–7519. <https://doi.org/10.1002/2016GL069938>. Received
- Dangermond, J., De Vito, C., Pesaresi, C. (2020). Using GIS in the Time of the COVID-19 Crisis, casting a glance at the future. A joint discussion. *J-Reading Journal of Research and Didactics in Geography*, 1(December 2019), 195–205. <https://doi.org/10.4458/3099-16>
- Delaunay, M., van der Westhuizen, H., Godard, V., Agius, R., Le Barbier, M., Godderis, L., Bonnetterre, V. (2015). Use of GIS in visualization of work-related health problems. *Occupational Medicine*, 65(8), 682–692. <https://doi.org/10.1093/occmed/kqv152>
- Dixon, P. M. (2002). Ripley's K function. In El-Shaarawi A. H. , Piegorsch W. W. (Eds.), *Encyclopedia of Environmetrics* (pp. 1796–1803). Chichester, John Wiley & Sons, Ltd.
- Duntas, L. H., Doumas, C. (2009). The “rings of fire” and thyroid cancer. *Hormones*, 8(4), 249–253.
- Eurostat. (2013). *Revision of the European Standard Population Report of Eurostat's task force*. <https://ec.europa.eu/eurostat/web/products-manuals-and-guidelines/-/ks-ra-13-028>
- Fagin, J. A., Wells, S. A. (2016). Biologic and Clinical Perspectives on Thyroid Cancer. *New England Journal of Medicine*, 375(11), 1054–1067. <https://doi.org/10.1056/nejmra1501993>
- Ferrante, M., Fiore, M., Ledda, C., Cicciù, F., Alonzo, E., Fallico, R., ... Sciacca, S. (2013). Monitoraggio di metalli pesanti ed oligoelementi in aria, ortofrutta e terreno nella provincia di Catania. *Igiene e Sanità Pubblica*, 69(1), 47–54.
- Fiore, M., Conti, G. O., Caltabiano, R., Buffone, A., Zucarello, P., Cormaci, L., ... Ferrante, M. (2019). Role of Emerging Environmental Risk Factors in Thyroid Cancer: A Brief Review. *International Journal of Environmental Research and Public Health*, 16(1185).
- Fitzmaurice, C., Dicker, D., Pain, A., Hamavid, H., Moradi-Lakeh, M., MacIntyre, M. F., ... Naghavi, M. (2015). The Global Burden of Cancer 2013. *JAMA Oncology*, 1(4), 505–527. <https://doi.org/10.1001/jamaoncol.2015.0735>
- Fondazione AIOM. (2020). *I numeri del cancro in Italia 2020*. [https://www.aiom.it/wp-content/uploads/2020/10/2020\\_Numeri\\_Cancro-operatori\\_web.pdf](https://www.aiom.it/wp-content/uploads/2020/10/2020_Numeri_Cancro-operatori_web.pdf)
- Franzini, M., Annovazzi-Lodi, L., Casella, V. (2020). Assessment of the Completeness of OpenStreetMap and Google Maps for the Province of Pavia (Italy). *GISTAM 2020 - Proceedings of the 6th International Conference on Geographical Information Systems Theory, Applications and Management*, 270–277. <https://doi.org/10.5220/0009564302700277>
- Fuchs, R. J. (1984). Government policy and population distribution. In Clarke J. I. (Ed.), *Geography and population: approaches and applications* (pp. 127–137). New York, Pergamon Press.
- Geraghty, E. M., Balsbaugh, T., Nuovo, J., & Tandon, S. (2010). Using geographic information systems (GIS) to assess outcome disparities in patients with type 2 diabetes and hyperlipidemia. *Journal of the American Board of Family Medicine*, 23(1), 88–96. <https://doi.org/10.3122/jabfm.2010.01.090149>
- Ghosh, M., Natarajan, K., Waller, L. A., & Kim, D. (1999). Hierarchical Bayes GLMs for the analysis of spatial data: An application to disease mapping. *Journal of Statistical Planning and Inference* (Vol. 75), 305–318
- Goodman, M. T., Yoshizawa, C. N., Kolonel, L. N. (1988). Descriptive epidemiology of thyroid cancer in Hawaii. *Cancer*, 61, 1272–1281.

- Hrafnelsson, J., Tulinius, H., Jonasson, J. G., Ólafsdóttir, G., & Sigvaldason, H. (1989). Papillary Thyroid Carcinoma in Iceland: A study of the Occurrence in families and the coexistence of other primary tumours. *Acta Oncologica*, 28(6), 785–788.
- OpenStreetMap, <https://www.openstreetmap.org>, ultimo accesso 06/11/2021
- Istat. (2013). *La Sicilia, un territorio che cambia*. <https://www.istat.it/it/archivio/183255>
- Kahle, D., Wickham, H. (2013). ggmap: Spatial Visualization with ggplot2. *The R Journal*, 5(1), 144–161. Retrieved from <http://journal.r-project.org/archive/2013-1/kahle-wickham.pdf>
- Karmacharya, A. (2018). Positional Accuracy of Online Geocoding Services: Case Study of Bhaktapur District. *Nepalese Journal of Geoinformatics*, 17(1), 16–21.
- Keddem, S., Barg, F. K., Glanz, K., Jackson, T., Green, S., George, M. (2015). Mapping the urban asthma experience: Using qualitative GIS to understand contextual factors affecting asthma control. *Social Science and Medicine*, 140, 9–17. <https://doi.org/10.1016/j.socscimed.2015.06.039>
- Keiding, J. K., Sigmarsson, O. (2012). Geothermobarometry of the 2010 Eyjafjallajökull eruption: New constraints on Icelandic magma plumbing systems. *Journal of Geophysical Research: Solid Earth*, 117(3), 1–15. <https://doi.org/10.1029/2011JB008829>
- Kilfoy, B. A., Zheng, T., Holford, T. R., Han, X., Ward, M. H., Sjodin, A., ... Zhang, Y. (2009). International patterns and trends in thyroid cancer incidence, 1973-2002. *Cancer Causes and Control*, 20(5), 525–531.
- Kitahara, C. M., Sosa, J. A. (2016). The changing incidence of thyroid cancer. *Nature Reviews Endocrinology*, 12(11), 646–653. <https://doi.org/10.1038/nrendo.2016.110>
- Kolonel, L. N., Hankin, J. H., Wilkens, L. R., Fukunaga, F. H., Hinds, M. W. (1990). An epidemiologic study of thyroid cancer in Hawaii. *Cancer Causes and Control*, 1, 223–234.
- Kung, T. M., Ng, W. L., Gibson, J. B. (1981). Volcanoes and carcinoma of the thyroid: A possible association. *Archives of Environmental Health*, 36(5), 265–267.
- Li, X., Anselin, L. (2021). rgeoda: R Library for Spatial Data Analysis. R package version 0.0.8-1. <https://CRAN.R-Project.Org/Package=rgeoda>.
- Liu, Y., Su, L., Xiao, H. (2017). Review of Factors Related to the Thyroid Cancer Epidemic. *International Journal of Endocrinology*. <https://doi.org/10.1155/2017/5308635>
- Macdonald, G. A. (1968). Composition and origin of Hawaiian lavas. *Memoir of the Geological Society of America*, 116, 477–522. <https://doi.org/10.1130/MEM116-p477>
- MahdaviFar, N., Pakzad, R., Ghoncheh, M., Pakzad, I., Moudi, A., Salehiniya, H. (2016). Spatial analysis of breast cancer incidence in Iran. *Asian Pacific Journal of Cancer Prevention*, 17, 59–64. <https://doi.org/10.7314/APJCP.2016.17.S3.59>
- Malandrino, P., Russo, M., Ronchi, A., Minoia, C., Cataldo, D., Regalbutto, C., ... Vigneri, R. (2016). Increased thyroid cancer incidence in a basaltic volcanic area is associated with non-anthropogenic pollution and biocontamination. *Endocrine*, 53, 471–479.
- Malandrino, P., Scollo, C., Marturano, I., Russo, M., Tavarelli, M., Attard, M., ... Mian, C. (2013). Descriptive epidemiology of human thyroid cancer: experience from a regional registry and the “volcanic factor.” *Frontiers in Endocrinology*, 4(65), 1–7.
- May, J. M. (1950). Medical Geography: Its Methods and Objectives. *Geographical Review*, 40(1), 9–41. <https://doi.org/10.2307/210990>
- Menck, H. R., Henderson, B. E. (1985). Cancer incidence in the Pacific Basin. *National Cancer Institute Monograph*, 69, 105–111.
- Moore, R. B., Trusdell, F. A. (1993). Geology of Kilauea Volcano. *Geothermics*, 22(4), 243–254.
- Musa, G. J., Chiang, P. H., Sylk, T., Bavley, R., Keating, W., Lakew, B., ... Hoven, C. W. (2013). Use of GIS Mapping as a Public Health Tool--From Cholera to Cancer. *Health Services Insights*, 6. <https://doi.org/10.4137/HSI.S10471MANCA NUMERO DI PAGINE>
- Najafabadi, A. T., Pourhassan, M. (2011). Integrating the geographic information system into cancer research. *Indian Journal of Cancer*, 48(1), 105–109. <https://doi.org/10.4103/0019-509X.75834>
- Newhall, C. G. (1979). Temporal variation in the lavas of Mayon volcano, Philippines. *Journal of Volcanology and Geothermal Research*, 6(1–2), 61–83. [https://doi.org/10.1016/0377-0273\(79\)90047-7](https://doi.org/10.1016/0377-0273(79)90047-7)
- Novembre, C. (2015). Le aree interne della Sicilia tra problemi di sviluppo e ricomposizione territoriale. *Rivista Geografica Italiana*, 122(2), 235–253.
- Nykiforuk, C. I. J., Flaman, L. M. (2011). Geographic Information Systems (GIS) for Health Promotion and Public Health: A Review. *Health Promotion Practice*, 12(1), 63–73. <https://doi.org/10.1177/1524839909334624>

- Omran, A. R. (1971). The epidemiologic transition. A theory of the epidemiology of population change. *Milbank Mem Fund Q.*, 49(4), 509–538.
- Openshaw, S. (1981). The Modifiable Areal Unit Problem. *Quantitative Geography: A British View*, 60–69.
- Palagiano, C. (1990). Il sistema salute. In Palagiano, C. (a cura di). *Atti del convegno sul tema: "La geografia medica oggi". Problemi tecnico-metodologici e applicazioni* (pp. 7–14). Roma: SGI.
- Palagiano, C. (1991). Applications and Applicability of the Geographical Research on Mortality: the example of Italy. *Espace, Populations, Sociétés*, 9(1), 237–243. <https://doi.org/10.3406/espos.1991.1464>
- Palagiano, C., De Santis, G., Scifoni, M. G. (1992). La geografia medica e gli ecosistemi. In *Atti del Quarto Seminario internazionale di Geografia medica (Roma, 1991)*. Perugia, Rux.
- Palagiano, C., Pesaresi, C. (2011). *La salute nel mondo. Geografia medica e qualità della vita*. Roma: Carocci.
- Pallister, J. S., Hoblitt, R. P., Meeker, G. P., Knight, R. J., Siems, D. F. (1996). Magma Mixing at Mount Pinatubo: Petrographic and Chemical Evidence from the 1991 Deposits. *Fire and Mud: Eruptions and Lahars of Mount Pinatubo*. U.S. Geological Survey.
- Parkin, D. M., Bray, F., Ferlay, J., & Pisani, P. (2005). Global Cancer Statistics, 2002. *CA: A Cancer Journal for Clinicians*, 55(2), 74–108. <https://doi.org/10.3322/canjclin.55.2.74>
- Pellegriti, G., De Vathaire, F., Scollo, C., Attard, M., Giordano, C., Arena, S., ... Vigneri, R. (2009). Papillary thyroid cancer incidence in the volcanic area of Sicily. *Journal of the National Cancer Institute*, 101, 1575–1583.
- Plantin, J. C. (2018). Google maps as cartographic infrastructure: From participatory mapmaking to database maintenance. *International Journal of Communication*, 12, 489–506.
- Polacci, M., Andronico, D., de' Michieli Vitturi, M., Taddeucci, J., & Cristaldi, A. (2019). Mechanisms of Ash Generation at Basaltic Volcanoes: The Case of Mount Etna, Italy. *Frontiers in Earth Science*, 7(August). <https://doi.org/10.3389/feart.2019.00193>
- R Core Team. (2020). *R: A language and environment for statistical computing*. Vienna. Retrieved from <http://www.r-project.org/>
- Ripley, B. D. (1977). Modelling Spatial Patterns. *Journal of the Royal Statistical Society*, B(39), 172–212.
- Ripley, B. D. (1991). *Statistical inference for spatial processes*. Cambridge, Cambridge University Press.
- Sanders, C. L., McKay, K. H. (2013). the Search for "Strong Medicine": Pathways To Healthcare Development in Remote Nepal Using Gis. *Technology & Innovation*, 15(2), 109–124. <https://doi.org/10.3727/194982413x13650843068799>
- Scrofani, L., Di Blasi, E., Di Bella, A. S., Ruggiero, L., Arangio, A., Petino, G. (2020). Territorial Cohesion in the Sicily's Planning 2014-2020. What are Policy Choices for not Being Isolated/Islanders? In Prezioso M. (Ed.), *Territorial Impact Assessment of National and Regional Territorial Cohesion in Italy* (pp. 547–559). Bologna, Pàtron Editore.
- Seib, C. D., Sosa, J. A. (2019). Evolving Understanding of the Epidemiology of Thyroid Cancer. *Endocrinology and Metabolism Clinics of North America*, 48(1), 23–35. <https://doi.org/10.1016/j.ecl.2018.10.002>
- Skinner R. (2010). *GIS in hospital and healthcare emergency management*. Boca Raton, FL, CRC Press.
- Small, C. (2004). Global Population Distribution and Urban Land Use in Geophysical Parameter Space. *Earth Interactions*, 8(8), 1–18. [https://doi.org/10.1175/1087-3562\(2004\)008<0001:gpdaul>2.0.co;2](https://doi.org/10.1175/1087-3562(2004)008<0001:gpdaul>2.0.co;2)
- Snow, J. (1855). *On the mode of communication of cholera*. London, John Churchill.
- Stevenson, L. G. (1965). Putting Disease on the Map: The Early Use of Spot Maps in the Study of Yellow Fever. *Journal of the History of Medicine and Allied Sciences*, 20(3), 226–261.
- Sui, D. Z. (2004). Tobler's first law of geography: a big idea for a small world? *Annals of the American Association of Geographers*, 94, 269–277.
- Tlacuilo-Parra, A., Garibaldi-Covarrubias, R., Romo-Rubio, H., Soto-Sumuano, L., Ruiz-Chávez, C. F., Suárez-Arredondo, M., ... Gallegos-Castorena, S. (2017). Geographical distribution and cluster detection of childhood leukemia in the metropolitan area of Guadalajara, Mexico. *Revista de Investigacion Clinica*, 69(3), 159–165. <https://doi.org/10.24875/RIC.17002131>
- Tobler, W. R. (1979). Smooth pycnophylactic interpolation for geographical regions. *Journal of the American Statistical Association*, 74(367), 519–530. <https://doi.org/10.1080/01621459.1979.10481647>
- Truong, T., Rougier, Y., Dubourdieu, D., Guihenneuc-Jouyaux, C., Orsi, L., Hémon, D., Guénel, P. (1985). Time trends and geographic variations for thyroid cancer in New Caledonia, a very high incidence area (1985-1999). *European Journal of Cancer Prevention*, 16(1), 62–70.

- Varrica, D., Tamburo, E., Dongarrà, G., Sposito, F. (2014). Trace elements in scalp hair of children chronically exposed to volcanic activity (Mt. Etna, Italy). *Science of the Total Environment*, 470–471, 117–126. <https://doi.org/10.1016/j.scitotenv.2013.09.058>
- Vigneri, R., Malandrino, P., Giani, F., Russo, M., Vigneri, P. (2017). Heavy metals in the volcanic environment and thyroid cancer. *Molecular and Cellular Endocrinology*, 457, 73–80.
- Vigneri, R., Malandrino, P., Vigneri, P. (2015). The changing epidemiology of thyroid cancer: Why is incidence increasing? *Current Opinion in Oncology*, 27, 1–7.
- Vizzini, S., Andolina, C., Caruso, C., Corbo, A. (2020). Isole Eolie: i campi di emissioni vulcaniche sottomarine di CO<sub>2</sub> a Vulcano e Panarea. *Mem. Descr. Carta Geol. d'It.*, 105, 91–96.
- Vizzini, S., Leonardo, R. Di, Costa, V., Tramati, C. D., Luzzu, F., Mazzola, A. (2013). Trace element bias in the use of CO<sub>2</sub> vents as analogues for low pH environments: Implications for contamination levels in acidified oceans. *Estuarine, Coastal and Shelf Science Journal*, 134, 19–30. <https://doi.org/10.1016/j.ecss.2013.09.015>
- Wakefield, J. (2007). Disease mapping and spatial regression with count data. *Biostatistics*, 8(2), 158–183. <https://doi.org/10.1093/biostatistics/kxl008>
- Waller, L. A., Gotway, C. A. (2004). *Applied Spatial Statistics for Public Health Data*. Hoboken, John Wiley & Sons.
- Walter, S. D. (2000). Disease mapping: a historical perspective. In Elliott P., Wakefield J., Best N., Briggs D. (Eds.), *Spatial Epidemiology: Methods and Applications* (pp. 223–239). Oxford University Press.
- Waterhouse, J., Muir, C., Shanugaratnam, K., Powell, J. E. (1982). *Cancer Incidence in Five Continents*. Lyon, France.
- World Population Ageing*. (2017). Retrieved from [https://www.un.org/en/development/desa/population/publications/pdf/ageing/WPA2017\\_Highlights.pdf](https://www.un.org/en/development/desa/population/publications/pdf/ageing/WPA2017_Highlights.pdf)
- Ye, X., Torabi, M., Lix, L. M., Mahmud, S. M. (2017). Time and spatial trends in lymphoid leukemia and lymphoma incidence and survival among children and adolescents in Manitoba, Canada: 1984–2013. *PLoS ONE*, 12(4). <https://doi.org/10.1371/journal.pone.0175701>
- Zambrano, L. I., Vasquez-Bonilla, W. O., Fuentes-Barahona, I. C., Cláudio da Silva, J., Valle-Reconco, J. A., Medina, M. T., ... Rodríguez-Morales, A. J. (2019). Spatial distribution of Zika in Honduras during 2016–2017 using geographic information systems (GIS) – Implications for public health and travel medicine. *Travel Medicine and Infectious Disease*, 31(November 2018), 101382. <https://doi.org/10.1016/j.tmaid.2019.01.017>