

Modelli di indagine per la valutazione di cluster spatio-temporali di tumori infantili

Investigative model for the evaluation of spatio-temporal clusters of childhood cancers

Paola Michelozzi,¹ Manuela De Sario,¹ Lucia Miligi,² Sara Piro,² Patrizia Schifano,¹ Luigi Bisanti³

¹ Dipartimento di epidemiologia del Servizio sanitario regionale, ASL Roma 1, Regione Lazio, Roma

² Istituto per lo studio e la prevenzione oncologica (ISPO), Firenze

³ Epidemiologo, Milano

Corrispondenza: Paola Michelozzi; p.michelozzi@deplazio.it

RIASSUNTO

Questo contributo descrive il quadro concettuale e le criticità delle indagini di *cluster* di tumori infantili e definisce un modello di indagine per le autorità sanitarie incaricate di valutare un sospetto *cluster*, tenendo conto delle linee guida disponibili e considerando gli sviluppi più recenti del Sistema informativo geografico e della metodologia statistica. Sono identificate tre fasi di indagine principali: la prima consiste nello studio preliminare sullo stato di salute della popolazione che vive nell'area in cui vengono definiti i casi e formulate ipotesi eziologiche; la seconda è la valutazione del *cluster* tramite metodi statistici che valutano l'eterogeneità spaziale e la raccolta di informazioni sui possibili fattori di rischio; la terza è lo studio epidemiologico analitico per testare le ipotesi eziologiche suggerite dall'indagine di *cluster*.

L'approccio di coorte residenziale è quello ad oggi più valido per valutare effetti a lungo termine, e permette di ricostruire l'intera storia residenziale dall'anagrafe della popolazione residente. La decisione del livello di approfondimento riguar-

do a un sospetto *cluster* deve tenere conto di considerazioni relative sia al livello di allarme presente nella popolazione sia alle limitate risorse disponibili. All'allarme per un sospetto *cluster* di casi di tumore si dovrebbe essere sempre capaci di dare una risposta, anche se ciò implica l'ammissione dei limiti delle proprie conoscenze e l'incertezza nell'interpretazione dei risultati disponibili.

Parole chiave: tumori infantili, *cluster*, studio analitico, fattori di rischio, incidenza di tumore

ABSTRACT

This paper describes the conceptual framework and the critical issues of investigations of clusters of childhood cancers and defines an investigative model for the health authorities responsible for assessing a suspected cluster, taking into account the guidelines available and considering the most recent advances of the Geographical Information System and of the specific statistical methodology. Three main investigation phases are identified: the first consists in the preliminary

study on the health of population living in the area where the cases are defined and aetiological hypotheses are formulated; the second is the cluster evaluation study using statistical methods assessing the spatial heterogeneity and collecting information about potential risk factors; the third is the analytical epidemiological study to test aetiological hypotheses suggested by the previous phases.

The residential cohort approach is the most valid to date to assess long-term effects, and allows to reconstruct the lifetime residential history from the population registry. The researchers' decision on how detailed about a suspected cluster the investigation has to be needs to take into account both the level of alarm in the population and the limited resources available. The concern about a suspected cluster of cancer cases should always be addressed, even if this implies to acknowledge limits of research and uncertainty in results interpretation.

Keywords: childhood cancers, cluster, analytical study, risk factors, cancer incidence

INTRODUZIONE

La valutazione di *cluster* spatio-temporali di malattia in una popolazione rappresenta una parte importante della ricerca epidemiologica. I primi studi sui *cluster* si riferiscono a esposizioni in ambito occupazionale; un drammatico caso storico è rappresentato dall'elevato numero di casi di osteosarcoma della mandibola in donne che lavoravano nelle industrie di orologi in New Jersey e Connecticut esposte a radiazioni ionizzanti.¹ Dagli anni Cinquanta sono stati studiati molti *cluster*, soprattutto di leucemie infantili in prossimità di stabilimenti industriali e centrali nucleari (vd. capitolo 8 «Revisione degli studi di cluster di leucemia infantile», pp. 38-41).²

I Centres for Disease Control and Prevention (CDC) di Atlanta sono stati tra i primi a sviluppare un approccio sistematico per condurre indagini su sospetti *cluster*,^{3,4} seguiti da altri Paesi.⁵⁻⁹

Gli studi di *cluster* rappresentano ancora oggi un tema molto controverso in epidemiologia e già nel 1990 Kenneth J. Rothman raccomandava cautela nell'investire risorse in questo tipo di indagini, in particolare quando non fosse presente un'ipotesi su un possibile meccanismo causale comune a tutti i casi del *cluster*.^{10,11} Poiché l'analisi di *cluster* rientra nella categoria degli studi di correlazione spaziale, non è adatta a testare ipotesi eziologiche; tuttavia, in rari casi ha portato al riconoscimento di un agente causale responsabile di tutti i casi osservati (vd. capitolo 8).

Gli studi che hanno permesso di identificare un agente causale comune riguardano in genere popolazioni di lavoratori (per esempio, casi di mesotelioma in lavoratori dell'amianto),¹² mentre la maggior parte delle indagini di *cluster* condotte su popolazioni residenti in una stessa area rimangono senza spiegazione. Una revi-

sione sistematica degli studi di *cluster* condotti negli Stati Uniti tra il 1990 e il 2011 ha mostrato che su 428 indagini che riguardavano 567 singole patologie tumorali o gruppi di patologie, solo in 72 era stato confermato il *cluster*, in tre casi veniva identificato il possibile fattore eziologico, ma in un solo caso veniva confermato.² D'altra parte la segnalazione di *cluster* di tumori è un evento molto frequente e fonte di grave allarme tra i residenti nel quartiere, nella scuola o nel luogo di lavoro in cui il *cluster* si verifica; dunque è un problema molto rilevante per la sanità pubblica.¹³ I CDC hanno osservato che «la percezione del rischio in una comunità può essere importante tanto quanto, o più importante di, un *cluster* reale» e che in nessun caso la denuncia di un possibile *cluster* di tumori dovrebbe essere trascurata dall'autorità sanitaria, anche quando la sua effettiva sussistenza è a priori poco probabile.^{3,4} Le segnalazioni di *cluster* hanno in genere molta visibilità sui media, poiché colpiscono l'opinione pubblica a causa della gravità della malattia, della popolazione colpita (spesso quella infantile, come nel caso della leucemia), dell'incertezza sulle cause, della difficoltà da parte degli esperti di condurre l'indagine e della conseguente mancanza di risposte tempestive e risolutive. Inoltre, per la seduzione esercitata sugli organi d'informazione da ogni evento nocivo e misterioso, molto spesso l'attenzione dei media non è causa di una riduzione dell'ansia, ma piuttosto segna il passaggio da un giusto allarme a un incontrollato allarmismo. È, quindi, indispensabile identificare il quadro concettuale e le criticità di questo tipo di indagine, in particolare approfondendo gli strumenti di cui dispone l'epidemiologia per indagini di *cluster* di tumori infantili, soprattutto in relazione ai fattori di rischio ambientali, e mettendone in luce i limiti.

QUADRO CONCETTUALE DEGLI STUDI SU CLUSTER DI TUMORI ED ESPOSIZIONI AMBIENTALI

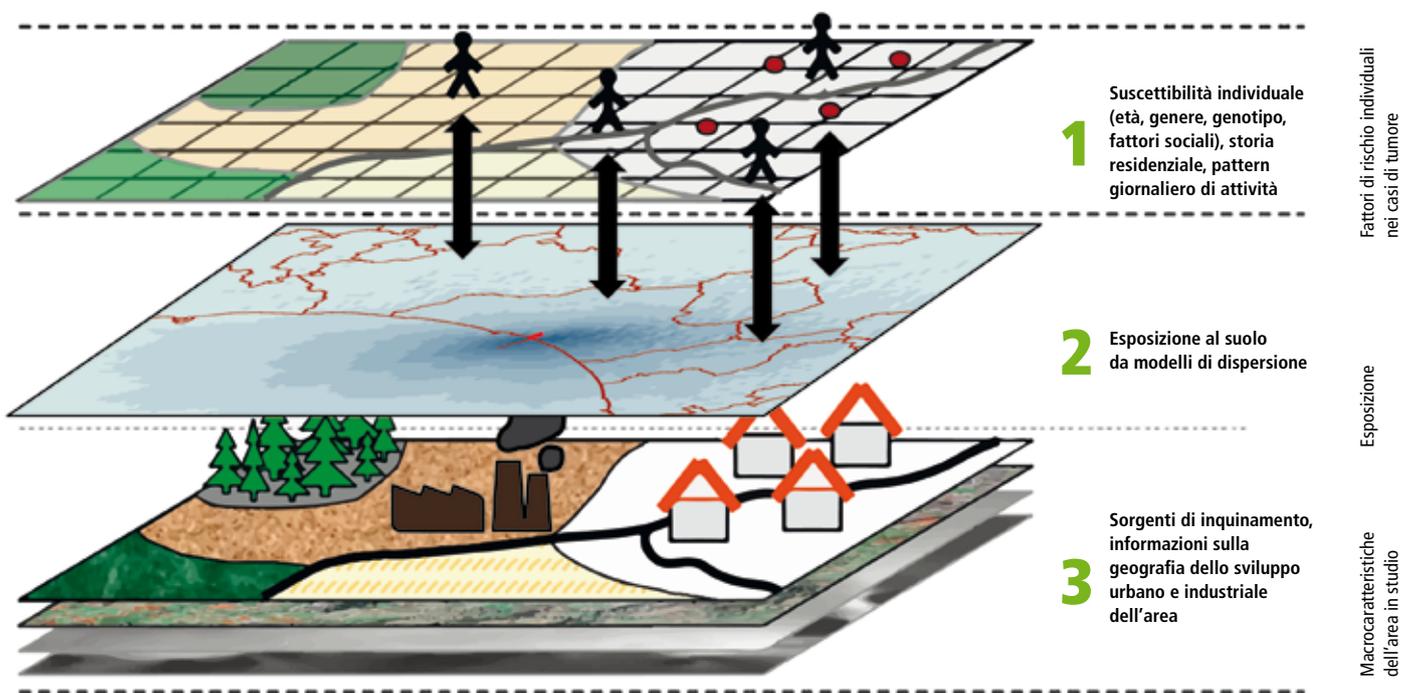
Per comprendere meglio le criticità legate all'indagine di *cluster*, è utile definire un quadro concettuale della relazione tra stato di salute di una popolazione, esposizione ambientale e area di residenza. Un *cluster* si verifica quando gli individui affetti da una certa malattia si distribuiscono in modo non casuale, ovvero formano un aggregato geografico e/o temporale. Nel caso di malattie tumorali, questo si può verificare se gli individui sono venuti in contatto con un agente cancerogeno che, per qualche motivo, è presente in concentrazione biologicamente rilevante nell'area di residenza, per esempio se vivono in prossimità di una sorgente puntiforme di un'esposizione altrimenti rara. A seconda dei meccanismi eziologici, varia il tempo intercorso tra esposizione e insorgenza della malattia (tempo di induzione), e anche il periodo di tempo tra insorgenza della malattia e manifestazione clinica della stessa (tempo di latenza). In questo contesto l'esposizione può essere rappresentata geograficamente, identificando con un punto nello spazio la fonte emissiva; la residenza degli individui può essere considerata un *proxy* della loro posizione nello spazio, di conseguenza l'intensità dell'esposizione è misurata in funzione della distanza dalla sorgente di inquinanti. Bisogna considerare che gli individui variano la loro posizione nello spazio per effetto della migrazione e per i loro specifici *pattern* giornalieri di attività. Il rischio è definito dalla sovrapposizione di tre superfici (figura 1): quella relativa all'esposizione (per esempio, stimata da modelli di dispersione degli inquinanti), quella relativa alla suscettibilità degli individui (legata a età, genere, genotipo, fattori so-

ciali) e quella dovuta a macrocaratteristiche dell'area in studio (come altre sorgenti di inquinamento, altri macrodeterminanti dello stato di salute della popolazione). Da questo quadro emerge che per studiare la correlazione spaziale servono dati validi di popolazione (in particolare rispetto alla residenza degli individui nel tempo), di malattia e delle esposizioni di interesse. Come negli altri studi epidemiologici, le stime ottenute possono essere affette da confondimento e altri *bias*.

STIMA DELL'ESPOSIZIONE SU BASE GEOGRAFICA

Molte esposizioni ambientali, come quelle derivanti dall'aria o dall'acqua potabile, per loro natura sono associabili a fattori geografici, quali il luogo di residenza, la sede di lavoro o di scuola degli individui.^{14,15} Per ogni individuo il livello di esposizione varia nel tempo e nello spazio ed è influenzato da fattori demografici e legati allo stile di vita (per esempio, età, genere, livello socioeconomico, tipo di occupazione). L'esposizione di un individuo a un fattore nocivo di origine ambientale può essere stimata dalla concentrazione di inquinante nell'ambiente (aria, acqua, suolo e alimentazione) e dalla concentrazione (*intake*) di inquinante nel corpo umano, ed è sottoposta ai meccanismi di assorbimento, distribuzione ed eliminazione. Per la stima dell'esposizione vanno considerati due aspetti: quello di induzione/latenza intercorsa tra l'esposizione e il danno considerato e quello di finestra di suscettibilità, ovvero il periodo nella vita di una persona in cui l'esposizione è associata al maggior rischio di sviluppare un danno.

Come in altri studi ecologici, negli studi di *cluster* sono considerate di interesse le esposizioni che caratterizzano l'ambiente di vita degli individui, se possibile rilevate a livello individua-



Adattata da: Leyk 2009¹⁶ (Creative Commons License)

Figura 1. Quadro concettuale per la valutazione di cluster spaziali di tumori.
Figure 1. Conceptual framework to evaluate spatial clusters of cancers.

le, quando non possibile a livello di piccola area (per esempio, sezione di censimento).^{14,15} Il Sistema informativo geografico (GIS) viene usato per stimare l'esposizione a livello di piccole aree o in specifici punti dello spazio utilizzando metodi di analisi spaziale descritti nel capitolo 10 («L'analisi statistica dei cluster in epidemiologia», pp. 55-64).

Per quanto riguarda la stima dell'esposizione possono essere utilizzati:

■ modelli di dispersione e trasporto: modelli complessi che integrano informazioni su emissioni, caratteristiche della sorgente delle emissioni, condizioni meteorologiche, topografia, chimica, insieme a misurazioni osservate, per stimare la distribuzione spaziale o temporale della concentrazione dell'esposizione emessa da una o più sorgenti locali;

■ modelli di regressione spaziali (come *land-use regression model*): sono costruiti usando dati sull'uso del suolo, sulla copertura vegetale, dati meteorologici, altitudine, densità di popolazione; a partire da dati di esposizione ambientale campionati nell'area in studio, il modello predice i valori per gli altri punti dell'area in studio.

In teoria, nello studio di un'associazione tra un'esposizione e una malattia con lungo periodo di induzione/latenza la misura ideale dell'esposizione è un marcatore biologico. Questo è possibile solo in pochi casi (per esempio, dal tasso di mutazioni somatiche nel DNA può essere ricostruita l'esposizione pregressa ad agenti genotossici come le radiazioni ionizzanti). Misure alternative dell'esposizione includono interviste attraverso questionario, la rilevazione dei dati ambientali disponibili, sia attuali sia storici, e la misura di marcatori biologici dell'esposizione attuale, assumendo che siano correlati all'esposizione del passato.

Le stime geografiche dell'esposizione necessitano di due livelli di validazione: rispetto a misure della concentrazione dell'esposizione nell'ambiente e rispetto a misure di concentrazione rilevate da sensori indossati dagli individui. In particolare, i *pattern* giornalieri di attività dei soggetti rappresentano una fonte di variabilità dell'esposizione individuale. Uno studio italiano che ha utilizzato il GPS per tracciare gli spostamenti di 103 bambini di aree urbane e rurali ai quali erano stati fatti indossare dispositivi per misurare l'esposizione a particolato atmosferico ha stimato che il contributo all'esposizione totale viene per il 15%-27% da attività svolte a scuola, per il 21%-27% da attività svolte in casa, per il 7%-16% da altre attività in ambienti confinati e solo per il 5%-13% da attività all'aperto.¹⁷

In mancanza di validazione dei dati di esposizione stimati dai modelli geografici, dovrebbero essere presentate stime quantitative dell'incertezza, come le varianze dei parametri tramite simulazioni Monte Carlo.¹⁸ Un approccio innovativo viene dall'integrazione dei dati sulla mobilità giornaliera degli individui, misurata tramite GPS e accelerometri in campioni della popolazione in studio, con analisi che utilizzano algoritmi automatici per identificare l'attività svolta nei diversi luoghi di vita (scuola, lavoro, percorsi da casa).¹⁹

Nuove prospettive nella stima dell'esposizione vengono dal recente sviluppo del concetto di esposoma, complementare a quello di genoma negli studi eziologici, che si riferisce alla totalità di esposizioni che ogni individuo subisce in tutta la sua

vita, integrando esposizioni dell'ambiente esterno e interno all'individuo, barriere sociali, culturali ed ecologiche.²⁰ Un'altra innovazione viene dai modelli di previsione *high-throughput* sviluppati dall'Agenzia per la protezione ambientale americana (EPA) per stimare i livelli di esposizione e il rischio associato all'esposizione a migliaia di sostanze chimiche contemporaneamente. Tali modelli permettono di stimare la concentrazione media di una sostanza chimica nell'aria, nell'acqua e nel suolo, e la distribuzione nella popolazione dell'esposizione dovuta a prodotti di uso comune nelle abitazioni. In Europa questo tipo di ricerche sono portate avanti dal consorzio *Health and Environment-wide Associations based on Large population Surveys* (HEALS).²¹

CRITICITÀ NEGLI STUDI DI CLUSTER DEI TUMORI INFANTILI RARITÀ DELLA MALATTIA

La maggiore difficoltà per lo studio dei *cluster* dei tumori infantili deriva dalla rarità di queste malattie; l'incidenza annuale attesa è, infatti, inferiore a 60 casi per milione di abitanti.²² In Italia, il numero di casi incidenti è stato di 164 nuovi casi all'anno per milione di bambini fino a 14 anni (periodo 2003-2008) (vd. capitolo 4 «Le basi dati per la stima dell'incidenza dei tumori infantili in Italia», p. 22). Un *cluster* di tumori infantili in una specifica area è di solito costituito da pochi casi, con conseguente riduzione della potenza statistica dei test di *cluster* e *clustering*.

STIMA DELL'ESPOSIZIONE NEGLI STUDI DI CLUSTER DI TUMORI INFANTILI

È stato stimato che nei bambini solo una porzione limitata di casi di tumore è attribuibile a fattori di rischio noti: una quota tra il 5% e il 10% a fattori genetici e familiari, una quota tra il 5% e il 15% a fattori ambientali, il restante numero di nuovi casi di tumore rimane non spiegato (vd. capitolo 2 «I fattori causali documentati o sospetti nell'eziologia dei tumori infantili», pp. 10-16). Nella ricerca dei fattori di rischio dei tumori infantili è importante considerare la maggiore suscettibilità dei bambini rispetto agli adulti a fattori nocivi di origine ambientale, la minore durata del periodo di induzione/latenza, la maggiore interazione tra fattori genetici e ambientali legati a diversi periodi della vita del bambino, la maggiore propensione dei sistemi cellulari in intensa replicazione a subire mutazioni.^{23,24} La più grande suscettibilità dei bambini rispetto agli adulti è dovuta principalmente a una quantità superiore di cibo, acqua e fluidi per unità di peso corporeo e a un più elevato volume di aria inspirata. Inoltre, durante l'infanzia si verificano una serie di cambiamenti fisiologici che influenzano l'assorbimento e il tipo di effetto del fattore nocivo di origine ambientale. Per esempio, i bambini hanno minori livelli epatici di enzimi detossificanti come il citocromo P450 e minore attività degli enzimi che riparano il DNA. I momenti dello sviluppo in cui avvengono i principali cambiamenti fisiologici sono poco noti. Anche se le evidenze sono poche, alcuni studi sperimentali confermano una maggiore suscettibilità alla formazione di tumori negli animali più giovani esposti a un agente cancerogeno, anche se la relazione tra età e suscettibilità ai tumori è complessa e dipendente dal tempo.

Un aspetto importante riguarda l'identificazione della finestra di rischio biologicamente rilevante. Le finestre di suscettibilità sono diverse in base al tipo di esposizione e di tumore. Le principali finestre di suscettibilità all'esposizione ambientale nociva da considerare sono:

- le esposizioni del padre o della madre prima del concepimento;
- le esposizioni *in utero*;
- le esposizioni durante i primi anni di vita;
- le esposizioni immediatamente precedenti all'esordio clinico della malattia.²⁵

Le prime fasi del processo evolutivo dei tumori della prima infanzia sono molto verosimilmente già avvenute *in utero*. Nel processo di cancerogenesi dovuta a fattori di rischio ambientali sembrano svolgere un ruolo importante meccanismi epigenetici prenatali. Per esempio, nelle forme di leucemia che si manifestano entro il primo anno di vita si sono riscontrati frequentemente riarrangiamenti del gene MLL1, codificante per un enzima chiave per l'attuazione di modifiche epigenetiche nelle prime fasi dello sviluppo fetale e tissutale e in grado di interferire con la differenziazione delle cellule staminali ematopoietiche pluripotenti.

Nel caso della leucemia linfoide acuta, si osserva un picco di incidenza nella prima infanzia (2-5 anni di età), con netta prevalenza di cloni mutati dei precursori dei linfociti B. Per questa malattia è stato proposto un modello di cancerogenesi a due stadi, con una prima mutazione *in utero* e una più tardiva, specialmente nei bambini che non hanno avuto un'esposizione precoce ai comuni agenti infettivi (*delayed infection* di Greaves).²⁶

Un modello alternativo spiega la clusterizzazione della leucemia infantile in aree a intenso ripopolamento, come l'effetto di nuovi virus in una specifica popolazione (*population mixing* di Kinlen).²⁶

L'esposizione può essere rilevata ricostruendo la storia residenziale con l'utilizzo di dati anagrafici e tramite intervista (che consente di valutare anche il domicilio effettivo del bambino). Un fattore da tenere in considerazione è, inoltre, l'eventuale associazione fra la mobilità residenziale e l'età, il livello socioeconomico, l'etnia o gli altri fattori individuali caratterizzanti i soggetti che formano il *cluster*. Informazioni accurate sulle esposizioni sono comunque difficilmente deducibili da dati autoriferiti, soprattutto in presenza di finestre strette di suscettibilità.

I biomarcatori in questo contesto non sono utili, perché indicatori di esposizioni avvenute nel breve-medio termine, a meno che non siano stati rilevati nell'ambito di una coorte longitudinale di nati. Possono essere, invece, validi per identificare polimorfismi genetici che conferiscono suscettibilità alle esposizioni ambientali; consentono, inoltre, di analizzare l'effetto dell'esposizione in sottogruppi con diversa suscettibilità genetica.

Grazie alle nuove tecnologie di *microarray* del DNA può essere effettuato lo screening di migliaia di polimorfismi di un singolo nucleotide e di geni marcatori di malattia, per esempio nel caso dei geni di fusione ALL1/MLL/HRX che sono altamente predittivi della leucemia acuta nei bambini.²⁷ In questo ambito ci sono criticità etiche e pratiche legate al fatto che le ricerche coinvolgono i bambini.

FASI DELL'INDAGINE DI CLUSTER

Vengono identificate le seguenti fasi dell'indagine di *cluster*:^{28,29}

1. studio preliminare sullo stato di salute della popolazione / identificazione dei casi / popolazione di riferimento;
2. valutazione della presenza di un *cluster* / identificazione dei fattori di rischio ambientali / sorveglianza sanitaria;
3. studio eziologico.

In questo paragrafo l'approccio in fasi è accompagnato da esempi (vd. box 1 e 2) che provengono da studi di *cluster* condotti in Italia che spesso non hanno prodotto un articolo scientifico e i cui risultati sono disponibili solo in forma di rapporto. Gli esempi riportati possono essere utili per illustrare come i risultati ottenuti e la loro interpretazione consentano di procedere attraverso le fasi successive dello studio di *cluster*.

FASE 1: STUDIO PRELIMINARE SULLO STATO DI SALUTE DELLA POPOLAZIONE / IDENTIFICAZIONE DEI CASI / POPOLAZIONE DI RIFERIMENTO

Questa fase affronta diversi aspetti:

■ **valutazione preliminare del sospetto *cluster***: spesso viene riferito un generico eccesso di tumori o un eccesso di casi di patologie diverse poco verosimilmente riconducibili a un'unica eziologia. Devono essere raccolti i dati rilevanti disponibili per ricavarne indicatori (tassi di mortalità, di ospedalizzazione, incidenza di tumori) capaci di descrivere lo stato di salute della popolazione in cui si è verificato il presunto *cluster*. Questa valutazione, che non riguarda solo la patologia in questione, può essere condotta sull'intera area amministrativa in cui si è verificato il presunto eccesso (ASL, comune) e consente di formare un quadro dello stato di salute della popolazione coinvolta rispetto ad aree di riferimento (comuni limitrofi o l'intero territorio comunale o regionale);

■ **revisione sistematica della letteratura** sulle esposizioni associate ai tumori inclusi nel *cluster* segnalato e sui tumori associati alle esposizioni ambientali rilevate nell'area;

■ **definizione di caso**: deve avere un'elevata specificità, mentre spesso i *cluster* segnalati includono diversi tipi di tumori;

■ **conferma della diagnosi dei casi**: i casi di tumore nella popolazione in studio dovrebbero essere identificati attraverso i registri tumori, dove disponibili (vd. capitolo 6 «Il contributo dell'Associazione italiana dei registri tumori (AIRTUM)», pp. 28-30). Il problema principale per i registri tumori è rappresentato da ritardi della registrazione dei casi, anche di diversi anni. Nel caso dei tumori infantili un'altra fonte è rappresentata dall'archivio AIEOP (vd. capitolo 5 «Il contributo dell'Associazione italiana di ematologia e oncologia pediatrica (AIEOP)», pp. 23-27). Il ricorso ad altre fonti di dati per l'accertamento dei casi, in particolare alle schede di dimissione ospedaliera e alle cartelle cliniche, è indispensabile in assenza di un registro tumori (vd. capitolo 7 «Un esempio di uso integrato di banche dati sanitarie correnti per lo studio dei cluster di tumori infantili», pp. 31-35). In quest'ultimo caso, la criticità principale riguarda la qualità dei dati, che può essere diminuita per errori di diagnosi o di codifica, oppure per la presenza di *record* duplicati che portano a una sovrastima del rischio di malattia. Se le inaccurately nei dati sono distribuite in modo eterogeneo nel-

lo spazio, questo può portare anche a un rischio di falsi positivi nella rilevazione del *cluster*;

■ **definizione dell'area geografica e del periodo temporale:** la decisione sui confini dell'area in studio è vincolata dalla disponibilità di dati a livello di aree definite per scopi amministrativi. In Italia, l'unità più piccola a livello amministrativo presente nei dati di popolazione è la sezione di censimento. L'errata definizione dell'area in studio può limitare la capacità di identificare il *cluster* e di comprenderne l'eziologia, portando, per esempio, al mancato riconoscimento di un vero *cluster* (a causa dell'esclusione di casi potenzialmente rilevanti) o alla rilevazione di uno falso (a causa dell'esclusione di persone esposte non malate).³⁰ Le decisioni sui limiti temporali e geografici dell'indagine dovrebbero tener conto del fatto che l'analisi statistica dei dati a livello di quartiere o di aree geografiche poco abitate potrebbe non essere praticabile a causa della bassa frequenza degli eventi considerati e, quindi, dell'instabilità dei tassi stimati e dell'insufficiente potenza statistica.³¹ Si deve tenere presente il rischio di *boundary shrinkage* che si verifica quando l'area in studio è talmente piccola da portare a una sovrastima del rischio di malattia, poiché i casi attesi sono pochi e la significatività statistica può essere amplificata in modo artefatto.²⁸

■ **definizione della popolazione di riferimento:** i denominatori, se possibile, dovrebbero derivare dai registri di popolazione; la criticità principale risiede nella qualità dei dati che può limitare la capacità di identificare il *cluster*.

In tabella 1 sono riassunti i limiti e i punti di forza di questa prima fase.

LIMITI	PUNTI DI FORZA
Ritardi nella registrazione dei casi nel registro tumori. Errori di diagnosi e di codifica e record duplicati usando altre fonti di dati di incidenza di tumori.	Consistenza con altri studi (revisione della letteratura)
Rischio di <i>boundary shrinkage</i> in piccole aree	Formulazione di ipotesi eziologiche (revisione della letteratura)
Errori nella stima dei denominatori	Plausibilità biologica (revisione della letteratura)
Instabilità dei tassi in caso di bassa frequenza della malattia in piccole aree	

Tabella 1. Limiti e punti di forza della fase 1.
Table 1. Limits and strengths of phase 1.

Box 1. Fase 1: dalla denuncia del sospetto cluster alla valutazione dell'eccesso di rischio nella popolazione in studio.
Box 1. Phase 1: from the report of suspected cluster to the evaluation of the excess risk in the study population.

A. Denuncia di un eccesso di mortalità per tumori in prossimità di una stazione radio a Roma

A Cesano, centro abitato alle porte di Roma, dal 1957 è attiva la stazione di Radio Vaticana, proprietà extraterritoriale della Santa Sede. A metà degli anni Novanta pochi studi epidemiologici, soprattutto di tipo occupazionale, avevano indagato l'associazione tra esposizione a radiofrequenze e tumori. Un medico di famiglia di Cesano annota un numero molto elevato di morti per cancro e comunica alla ASL di competenza che la mortalità per tumori a Cesano, secondo i suoi calcoli, è tre volte superiore a quella nazionale. La denuncia del medico determina una richiesta di indagine epidemiologica all'Osservatorio epidemiologico della Regione Lazio: è il 1996. L'eccesso rilevato dal medico di Cesano non trova conferma nell'analisi della mortalità per causa che non mette in luce eccessi per patologie tumorali nella popolazione residente. L'allarme, tuttavia, rimane alto nella popolazione. Negli stessi anni vengono pubblicati due studi inglesi che suggeriscono un'associazione tra esposizione a radiofrequenze e incidenza e mortalità per leucemie.^{32,33} Si decide, quindi, di replicare lo studio inglese a Cesano utilizzando la stessa metodologia e i risultati riportano un eccesso di mortalità e di incidenza di leucemia infantile.³⁴

B. Identificazione di un eccesso di casi di leucemia linfoide acuta (LLA) in una scuola di Milano

Nel gennaio 2010 un dirigente scolastico segnala alla ASL di Milano 3 casi di LLA tra gli alunni della scuola e un altro caso in un bambino in età prescolare residente nello stesso quartiere. Per tutti i 4 casi l'esordio di malattia è avvenuto nelle quattro settimane comprese tra il 15.12.2009 e il 14.01.2010. A seguito della segnalazione, la struttura ospedaliera dove vengono ricoverati i casi di leucemia infantile segnala altri 3 casi di LLA diagnosticati nelle stesse settimane in bambini residenti a Milano ma in quartieri diversi, portando le dimensioni complessive del *cluster* temporale/spaziale a 7 casi: tutti residenti a Milano, 4 in uno stesso quartiere e, di questi, 3 alunni di una stessa scuola elementare. Una delle prime fasi dell'indagine avviata dalla ASL ha riguardato l'analisi dell'andamento della stessa malattia negli anni precedenti a Milano e in aree di riferimento. In tabella 2 sono riportati i tassi standardizzati di tutte le leucemie infantili (0-14 anni) osservati a Milano, nel Nord Italia e nell'intero Paese.

	MILANO			NORD ITALIA			ITALIA		
	Dati ASL			Dati AIRTUM			Dati AIRTUM		
	Maschi	Femmine	Totale	Maschi	Femmine	Totale	Maschi	Femmine	Totale
Leucemie	72,3	44,3	58,7	n.d.	n.d.	56,1	59,2	49,1	54,3
LLA	56,1	38,9	47,8	n.d.	n.d.	44,9	47,0	38,9	43,0

n.d.: non disponibile / not available

Tabella 2. Tassi standardizzati di incidenza (x1.000.000/anno) per tutte le leucemie infantili (casi 0-14 anni) a Milano, nel Nord Italia e nell'intero Paese. Dati del Registro tumori di Milano 1999-2006; dati AIRTUM 1998-2002. / **Table 2.** Standardized incidence rates (x1,000,000/year) for all childhood (0-14 years) leukaemias in Milan, in Northern Italy, and in the whole Country. Data from Cancer registry of Milan 1999-2006; data from AIRTUM database 1998-2002.

I tassi di incidenza di Milano risultavano più alti sia di quelli nazionali sia di quelli riferiti dagli altri registri del Nord, ma non in modo statisticamente significativo. L'eccesso di incidenza osservato riguardava solo i maschi. La ASL decide di proseguire l'indagine di *cluster*.

(Fonte: Bisanti 2010)³⁵

C. Aumento dell'incidenza di leucemia infantile in un municipio di Roma

Nel periodo 2000-2011 viene segnalato al Dipartimento di epidemiologia del Servizio sanitario regionale del Lazio (DEP Lazio), ASL Roma 1 (ex ASL Roma E) un presunto *cluster* di casi pediatrici (0-14 anni) di leucemia linfoide acuta (LLA) tra i residenti in un municipio del Comune di Roma (Municipio XIII, ora Municipio X). Il DEP Lazio decide di valutare l'incidenza di malattia nel comune di Roma e nei diversi municipi: la tabella riporta i tassi di incidenza standardizzati con metodo diretto sulla popolazione italiana al Censimento 2001 per il comune di Roma e i rischi relativi (RR) di ciascun municipio rispetto al comune di Roma calcolati come rapporto dei rispettivi tassi standardizzati.

I 27 casi pediatrici di LLA nel Municipio XIII osservati nel 2000-2011 mostrano un eccesso rispetto all'atteso (riferimento: comune di Roma). Si decide, dunque, di proseguire nell'analisi eseguendo test di *cluster* per identificare se all'interno del municipio siano presenti o meno aggregati di casi.

(Fonte: Dipartimento di epidemiologia del Servizio sanitario regionale, ASL Roma 1 2013)³⁶

	CASI (n.)	TASSO STANDARDIZZATO (IC95%)	RISCHIO RELATIVO (IC95%)
Roma	194	44,1 (38,3-50,8)	1,0 –
Municipio I	5	35,0 (14,5-84,4)	0,8 (0,3-1,9)
Municipio II	5	28,2 (11,7-67,8)	0,6 (0,3-1,6)
Municipio III	2	27,9 (7,0-111,5)	0,6 (0,2-2,6)
Municipio IV	14	47,6 (28,2-80,4)	1,1 (0,6-1,9)
Municipio V	10	34,8 (18,7-64,7)	0,8 (0,4-1,5)
Municipio VI	4	22,3 (8,4-59,4)	0,5 (0,2-1,4)
Municipio VII	9	44,4 (23,1-85,3)	1,0 (0,5-2,0)
Municipio VIII	17	42,3 (26,2-68,1)	1,0 (0,6-1,6)
Municipio IX	12	67,7 (38,4-119,3)	1,5 (0,9-2,8)
Municipio X	8	27,3 (13,6-54,6)	0,6 (0,3-1,3)
Municipio XI	8	40,7 (20,3-81,4)	0,9 (0,5-1,9)
Municipio XII	16	54,4 (33,3-88,8)	1,2 (0,7-2,1)
Municipio XIII	27	72,3 (49,5-105,5)	1,6 (1,1-2,5)
Municipio XV	7	29,2 (13,9-61,3)	0,7 (0,3-1,4)
Municipio XVI	9	41,4 (21,6-79,7)	0,9 (0,5-1,8)
Municipio XVII	3	33,5 (10,8-104,0)	0,8 (0,2-2,4)
Municipio XVIII	11	52,3 (28,9-94,5)	1,2 (0,7-2,2)
Municipio XIX	18	60,9 (38,4-96,7)	1,4 (0,9-2,2)
Municipio XX	9	35,3 (18,4-67,9)	0,8 (0,4-1,6)

Tabella 3. Tassi standardizzati di incidenza (x1.000.000/anno) (IC95%) di LLA e rischi relativi (IC95%) nel comune di Roma e nei singoli municipi. Bambini 0-14 anni.

Table 3. Standardized incidence rates (x1,000,000/year) (95%CI) of ALL and relative risks (95%CI) in the municipality of Rome and in every city district. Children 0-14 years.

FASE 2. VALUTAZIONE DELLA PRESENZA DI UN CLUSTER / IDENTIFICAZIONE DEI FATTORI DI RISCHIO AMBIENTALI / VALUTAZIONE DELLA CORRELAZIONE TRA ESPOSIZIONI IDENTIFICATE E CASI DI MALATTIA

Questa fase affronta diversi aspetti:

■ **valutazione della presenza di un cluster:** una serie di test di analisi spaziale è utile per verificare se i casi di tumore in studio si distribuiscono in agglomerati spaziali, ossia per testare l'eterogeneità spaziale, ed è descritta nel capitolo 10. Questi test si basano sull'ipotesi nulla di una distribuzione casuale dei casi nel-

lo spazio, ma la *performance* dei test dipende dalla validità delle assunzioni sottostanti.³⁷ La difficoltà di definire valori attesi a priori può essere risolta ricorrendo all'analisi bayesiana che, per prima cosa, incorpora la distribuzione a priori dell'area in studio e del momento temporale di interesse, poi applica uno dei test statistici.^{38,39} Sono disponibili anche metodi per valutare il *clustering* spazio-temporale tra sottogruppi diagnostici;⁴⁰

■ **il problema dei confronti multipli:** nell'analisi di *cluster*, l'alto numero di test statistici (per esempio quando si analizzano più cause di malattia in una stessa area) può portare a un ec-

cesso di falsi positivi utilizzando il *p-value*. È stato recentemente suggerito l'utilizzo del *false discovery rate* bayesiano, o *q-value*, che rappresenta un approccio più adeguato rispetto al metodo classico di Bonferroni;⁴¹

■ **il problema dell'inferenza post-hoc:** Coory et al. sostengono che negli studi di *cluster* l'uso del *p-value* e degli intervalli di confidenza (IC) non dovrebbe proprio essere preso in considerazione, trattandosi di un disegno di studio post-hoc, per il quale questi strumenti sono impropri;⁴²

■ **raccolta dei dati sui potenziali fattori di rischio ambientali** identificati e sui fattori di rischio individuali (come dieta, infezioni e storia familiare): nella ricostruzione dell'esposizione va tenuto conto del *recall bias*, che interessa in particolar modo le popolazioni che vivono in aree con criticità ambientali e che, quindi, sono più attente alla segnalazione sia di potenziali *cluster* sia di contaminanti nell'ambiente;⁴³

■ **plausibilità biologica dell'associazione tra esposizione e malattia:** i casi che costituiscono il sospetto *cluster* devono condividere un'eziologia comune compatibile con i fattori di rischio rilevati, con il tipo di tumore, con il tempo di induzione/latenza e con i *pattern* migratori entro e fuori dalla comunità;

■ **valutazione della possibile correlazione tra fattori di rischio identificati e cluster di malattia:** gli studi di *cluster*, come già ricordato, non sono adatti a testare ipotesi eziologiche e solo in rari casi hanno portato al riconoscimento dell'agente causale; possono essere utili, però, a generare ipotesi da testare con studi ad hoc (si veda fase 3). Tuttavia, nella pratica comune sono effettuate misure di esposizione dei casi e della popolazione di riferimento ai fattori di rischio postulati. Questa pratica può rivelarsi utile sia per escludere l'azione sui casi di un fattore di rischio debole, sia, al contrario, per suggerire studi più approfonditi qualora l'esposizione tra i casi risulti elevata. Va comunque tenuto conto dell'effetto rassicurante prodotto sulla popolazione interessata che la conduzione di misurazioni dell'esposizione ai fattori di rischio postulati produce;

■ **confondimento attribuibile al livello socioeconomico e ad altre caratteristiche della popolazione:** nella valutazione dei fattori di rischio è importante tener conto dei potenziali confondenti, in particolare del livello socioeconomico della popolazione in studio e di suoi sottogruppi. I sistemi informativi utilizzati per la delimitazione della popolazione di interesse e per la stima dell'incidenza di malattia non hanno, di regola, informazioni su importanti confondenti come il livello socioeconomico. Tuttavia, spesso si rivela utile il ricorso a *proxy* del livello socioeconomico (per esempio, il reddito medio della popolazione residente nelle aree statistiche in cui è suddiviso il territorio dei maggiori comuni italiani);

■ **sorveglianza epidemiologica:** se l'eccesso di casi segnalato ha una bassa significatività statistica e l'esposizione ha una bassa plausibilità biologica, è ragionevole avviare una sorveglianza epidemiologica che avrà per obiettivo la verifica che l'eccesso si mantenga nel tempo e che la distribuzione spaziale dei nuovi casi sia coerente con quella del supposto *cluster*. La sorveglianza epidemiologica deve in ogni caso essere parte

integrante del piano di indagini avviate a seguito della segnalazione di un possibile *cluster*. Il riscontro, negli anni successivi alla segnalazione, dell'assenza o della sensibile riduzione del tipo di tumore infantile in studio sarà un elemento a sostegno della veridicità del *cluster* stesso e della sua dipendenza causale da un fattore ambientale a comparsa improvvisa e a durata limitata nel tempo (come un'epidemia batterica o virale). Secondo il *two hits model* di Mel Greaves,⁴⁴ infatti, il fattore promotore (secondo impulso) agisce sui soggetti suscettibili portatori di specifiche mutazioni geniche (primo impulso) attivando cloni cellulari pre-leucemici e sincronizzando l'esordio clinico della malattia (comparsa del *cluster*) in soggetti che, altrimenti, avrebbero manifestato i sintomi in diversi momenti successivi o non si sarebbero ammalati affatto. La sincronizzazione dell'esordio ha un effetto *harvesting* dei suscettibili, rivelato dal deficit di casi nell'anno (o negli anni) successivo. Per la sorveglianza epidemiologica può essere sufficiente la consultazione di fonti già disponibili (per esempio, i registri tumori) oppure può essere necessario allestire sistemi di rilevazione ad hoc. La sorveglianza permetterà di valutare gli scostamenti rispetto all'atteso della distribuzione degli eventi in studio nel tempo e nello spazio. Da un punto di vista statistico ciò significa calcolare due parametri: un tasso di riferimento basato sulle serie storiche e una soglia di allarme (per esempio, una deviazione standard del *moving average range*). Il metodo deve essere sensibile nel monitorare cambiamenti nel tasso (basso rischio di falsi negativi) e specifico, ovvero ridurre il rischio di falso allarme (falsi positivi). La sorveglianza attiva permanente, o di lungo periodo, permette di identificare i possibili *cluster* via via che si presentano; la sorveglianza limitata ai 2-3 anni successivi al *cluster* fornirà le informazioni utili all'analisi eziologica dell'evento. Tuttavia, vanno accuratamente considerati sia i costi di gestione della sorveglianza e della conseguente comunicazione, sia i rischi di generare falsi allarmi o di produrre risultati distorti dagli effetti dell'induzione/latenza, della mobilità, dei confronti multipli e dell'instabilità dei tassi, sia gli aspetti etici collegati.

In tabella 4 sono riassunti i limiti e i punti di forza di questa seconda fase.

LIMITI	PUNTI DI FORZA
Problema dei confronti multipli	Approccio bayesiano utilizza anche le informazioni fornite dagli allarmisti
Maggior rischio di <i>recall bias</i> in popolazioni che vivono in aree con criticità ambientali	Approccio basato sul GIS tramite georeferenziazione
Problema dell'inferenza <i>post hoc</i>	Attivazione di una sorveglianza sanitaria nell'area
Scarse informazioni su potenziali confondenti da sistemi informativi, in particolare sul livello socioeconomico	

Tabella 4. Limiti e punti di forza della fase 2.

Table 4. Limits and strengths of phase 2.

Box 2. Fase 2: dalla valutazione della presenza del cluster alla stima dei possibili fattori di rischio.
Box 2. Phase 2: from the evaluation of the existence of a cluster to the estimation of potential risk factors.

A. Rilievi ambientali nello studio del cluster di LLA in una scuola di Milano

In riferimento all'occorrenza del cluster verificatosi tra il 2009 e il 2010 nel comune di Milano, nell'indagine sono stati presi in considerazione tutti i fattori di rischio (certi o sospetti) per le leucemie infantili segnalati in letteratura, tra cui i campi magnetici a bassa frequenza (ELF). Sono state effettuate misurazioni ambientali finalizzate a quantificare in modo diretto l'esposizione a ELF nelle abitazioni (al momento della diagnosi) dei 7 casi e nei plessi scolastici; sono poi state raccolte informazioni sull'esposizione pregressa (per esempio, a linee elettriche in prossimità delle abitazioni) mediante un questionario somministrato ai genitori dei bambini.

Le misurazioni effettuate sia nelle scuole sia nelle abitazioni sono risultate ampiamente al di sotto di 0,3-0,4 μT (valore di esposizione per almeno quattro ore al giorno per alcuni anni associato a un eccesso di rischio di leucemia in due metanalisi);^{45,46} valori di campo magnetico superiori sono stati misurati solo in due abitazioni in particolari condizioni di carico. Per quanto riguarda le esposizioni pregresse rilevate dal questionario, non si è rilevata la presenza di linee elettriche in prossimità delle abitazioni.

(Adattato da: Bisanti 2010)³⁵

B. Incidenza di leucemia ed emissione di campi elettromagnetici in un municipio di Roma

Nello studio dell'eccesso di leucemie infantili identificato in un municipio di Roma³⁶ è stata condotta un'analisi per valutare il rischio in prossimità di fonti di emissione di campi magnetici ad alta e bassa frequenza presenti nell'area. Le fonti di esposizione identificate sono state otto: una sorgente di campi ELF (denominata ELF) e sette fonti di radiofrequenze (denominate RF da 1 a 7). Tutte le fonti di emissione sono state georeferenziate (figura 2) e sono state definite aree concentriche di 0-0,2 km, 0,2-0,5 km e due aree cumulate rispettivamente di 0-0,5 km e 0-1 km da ogni fonte di emissione. Per valutare l'esposizione dei casi è stato utilizzato e georeferenziato l'indirizzo al momento della diagnosi.

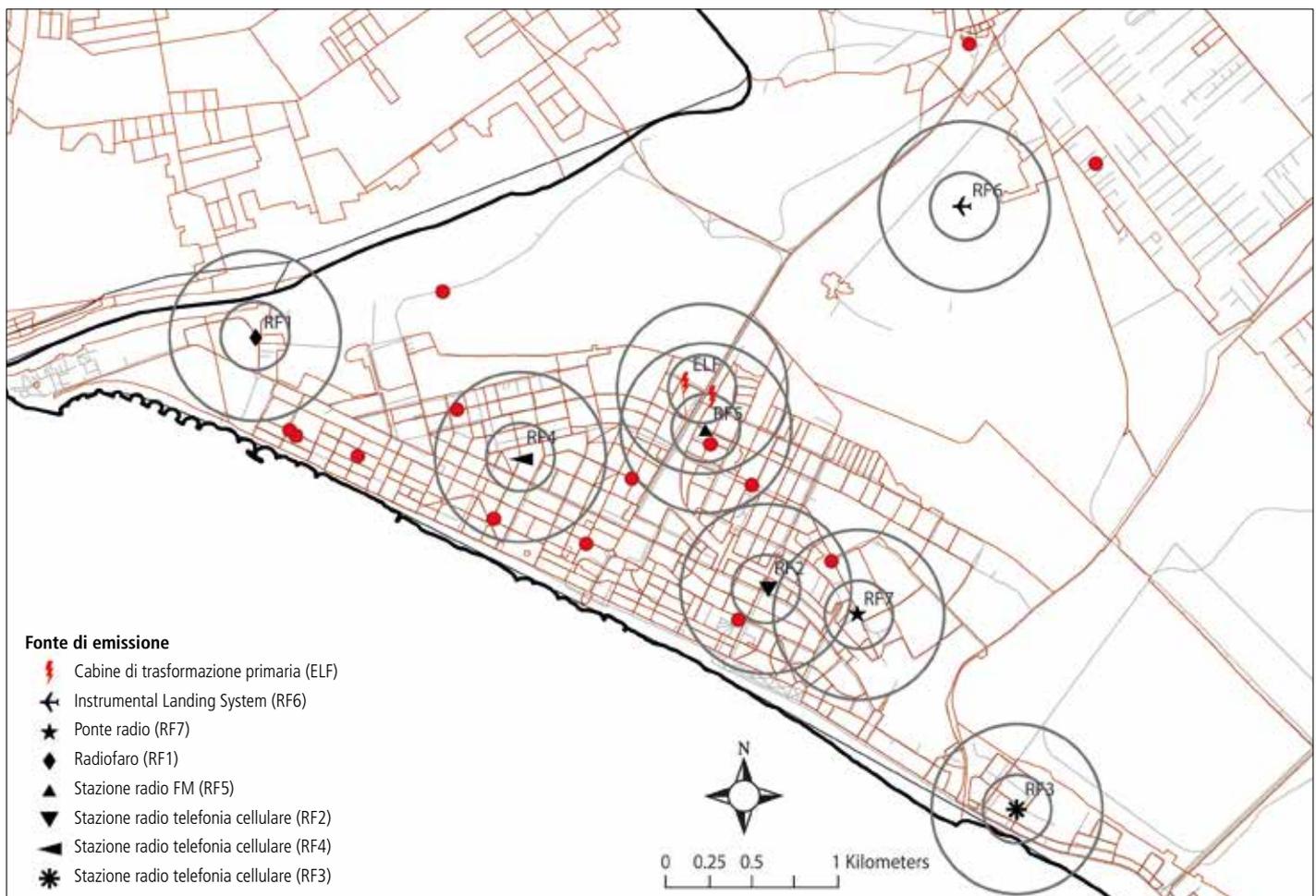


Figura 2. Casi di LLA infantile (codice ICD-9: 204), fonti di emissione e aree di 0,2 e 0,5 km da ciascuna fonte.
Figure 2. Cases of childhood ALL (ICD-9 code: 204), emission sources, and areas of 0.2 and 0.5 km from each source.

Dall'analisi dei rapporti standardizzati di incidenza (SIR) si nota che nell'area entro 0,2 km dalle fonti di emissione è presente solo la residenza di 1 caso (fonte RF5). Nell'area entro 500 metri si è osservato 1 caso per le fonti ELF e per l'emittente RF7 e 2 casi per le emittenti RF2, RF4 e RF5. In nessun caso le stime raggiungono la significatività statistica. Considerando insieme tutte le aree tra 0,2 e 0,5 km dalle fonti, i tassi di incidenza sono superiori all'atteso, ma tale eccesso non risulta statisticamente significativo (0-0,2 km: SIR: 1,65; IC95% 0,02-9,19 – 0-0,5 km: SIR: 1,73; IC95% 0,63-3,78). L'aggiustamento per livello socioeconomico non modificava le stime. I SIR più elevati si osservano a maggiore distanza dalle fonti: tra 500 metri e 1 km, 3 casi sono localizzati da RF1 (SIR: 4,17; IC95% 0,84-12,19) e 2 casi da RF6 (SIR: 25,19; IC95% 2,83-90,95). Tali eccessi rilevati nell'area a oltre 500 metri dai due impianti sembrano non riferibili alle fonti di emissione e identificano probabilmente due diversi *cluster* spaziali.

C. Rilevazione di un *cluster* spazio-temporale e ipotesi infettiva nel *cluster* di LLA in una scuola di Milano

In riferimento all'occorrenza del *cluster* verificatosi tra il 2009 e il 2010 nel comune di Milano, per rilevare la presenza di eventuali *cluster* spazio-temporali è stata utilizzata la statistica Scan proposta da Kulldorff.⁴⁷

Nell'arco del periodo 1999-2010, lo Scan test ha individuato un unico *cluster* spazio-temporale statisticamente significativo (*p-value*: 0,017) di LLA, delimitato dal periodo 01.12.2009-31.01.2010 e che interessa tre aree amministrative adiacenti al plesso scolastico.

Per tutti i rischi ambientali considerati non è stata messa in luce alcuna associazione con l'eccesso di LLA osservato. Il *cluster* è stato spiegato ipotizzando l'azione di una causa infettiva epidemica sulla base della considerazione, avvalorata dalla letteratura,²⁶ che un clone cellulare leucemogeno può essere inibito nella sua capacità patogena dal sistema immunitario dell'ospite fino a quando questo non venga impegnato dall'attacco di un agente infettivo, come è dimostrato dall'aumento dei nuovi casi di leucemia in corrispondenza delle epidemie influenzali.

Il virus influenzale A(H1N1), che fino al 2009 era del tutto sconosciuto biologicamente, ha avuto una larga diffusione epidemica da giugno a dicembre 2009, esattamente nel semestre che ha preceduto il *cluster* di LLA.

La concentrazione di 4 dei 7 casi del *cluster* in uno stesso quartiere, di cui 3 in una stessa scuola elementare, può essere spiegata con la teoria della *delayed infection*²⁶ secondo la quale (in analogia con l'ipotesi igienica) i bambini geneticamente predisposti alla leucemia che nel primo anno di vita sono esposti a un basso numero di episodi infettivi (condizione tipica dei nuovi nati in famiglie con condizione socioeconomica elevata) sono a maggior rischio di LLA per l'incapacità del loro sistema immunitario, poco esercitato, di fronteggiare un evento infettivo e, allo stesso tempo, tenere a bada il clone cellulare pre-leucemico già presente nel tessuto emopoietico del bambino. Il quartiere e la scuola oggetto del *cluster* sono in un'area molto centrale ed esclusiva di Milano, abitata da persone di condizione socioeconomica molto elevata.

(Adattato da: Bisanti 2010)³⁵

D. Attivazione di una sorveglianza epidemiologica dei casi di leucemia in alcuni comuni della Toscana

I medici di medicina generale hanno segnalato alla ASL 11 della Regione Toscana 8 casi di tumori emolinfopoiетici in ragazzi e adulti. Viene istituito un gruppo di lavoro con rappresentanti della ASL, dell'Istituto per lo studio e la prevenzione oncologica (ISPO) di Firenze, dell'Università degli Studi di Pisa, del Consiglio nazionale delle ricerche (CNR), dei Sindaci dei comuni interessati e dell'Agenzia ambientale regionale (ARPAT).

L'area in esame ha una forte vocazione industriale (comprendente il cuoio), gli insediamenti produttivi e le attività di servizi accessori, come i consorzi di depurazione delle acque e di trattamento e recupero dei rifiuti, sono ancora in molte aree fortemente connessi al tessuto residenziale; di conseguenza, la preoccupazione per possibili ripercussioni sulla salute sono molto sentiti dalla popolazione e dalle istituzioni locali.

Il gruppo di lavoro ha effettuato: • l'aggiornamento dei dati di mortalità e di incidenza delle neoplasie del sistema emolinfopoiетico utilizzando le fonti informative disponibili; • l'analisi di mortalità per tutti i tumori e per il totale delle neoplasie del sistema emolinfopoiетico; • la ricostruzione della storia residenziale dei casi; • la definizione dei criteri per l'attribuzione univoca della residenza, con relativa georeferenziazione; • l'analisi microgeografica per valutare l'incidenza a livello di microaree sub-comunali; • la caratterizzazione ambientale territoriale per valutare la presenza di fattori di pressione per i quali esista una relazione con le patologie oggetto dell'indagine.

È emersa una situazione complessiva in linea con i riferimenti esterni all'area di interesse e alcuni segnali di attenzione connessi a eccessi di mortalità e di incidenza in alcuni comuni della ASL. Tra le caratteristiche ambientali indagate (viabilità-traffico, uso di pesticidi, emissioni di solventi da imprese del territorio e presenza di elettrodotti e impianti di telecomunicazione) non è stata identificata una potenziale causa comune in grado di ipotizzare un legame con gli eccessi descritti. In conclusione, l'ASL 11 decide di proseguire la sorveglianza dello stato di salute della popolazione aggiornando periodicamente il nuovo archivio ASL dei casi incidenti delle neoplasie del sistema emolinfopoiетico e di rafforzare la raccolta di informazioni sulle pressioni ambientali esistenti.

(Fonte: ASL 11 Empoli 2012)⁴⁸

FASE 3: STUDIO EZIOLOGICO

Di seguito, gli aspetti affrontati da questa fase:

■ **estendere lo studio di *cluster* utilizzando una casistica più ampia:** l'obiettivo è di avere una maggiore potenza per valutare la possibile associazione tra esposizione a uno specifico contaminante ambientale sospetto di aver avuto un ruolo

nell'occorrenza del *cluster*. Un'associazione statisticamente significativa non rappresenta, però, una prova di causalità di per sé; può essere, quindi, necessario utilizzare un disegno di studio eziologico (coorte o caso-controllo) (tabella 5) insieme a studi clinici e di laboratorio.¹⁰ Un approccio tradizionale di analisi epidemiologica del *cluster* utilizzando questi disegni di studio

ASPETTO METODOLOGICO	RACCOMANDAZIONE
Nesso di causalità	Le ipotesi devono essere verificabili e deve esistere una plausibilità biologica.
Unità di analisi	Di solito sono definite in base alla più piccola unità amministrativa o statistica disponibile. Va verificata la corrispondenza con la distribuzione dell'esposizione e della malattia e con i <i>pattern</i> di attività degli individui.
Esposizione a livello ecologico	Va definita in modo concettuale e operativo.
Altri fattori a livello individuale	Vanno definiti a livello concettuale se confondenti, moderatori o mediatori e vanno usati metodi adeguati per l'aggiustamento.
Potenza, dimensione campionaria e rappresentatività	Vanno stimate nel protocollo di studio.
Modelli multilivello	Vanno applicati in modo appropriato, includendo gli effetti casuali a livello sia concettuale sia operativo.
Disegno dello studio	Il migliore approccio è quello longitudinale, meno affetto da <i>bias</i> di selezione e misclassificazione, in grado di valutare il nesso di causalità e monitorare le variazioni temporali delle relazioni tra individui e area di residenza.

(Adattata da: Riva 2007)⁴⁹

Tabella 5. Principali aspetti metodologici da considerare nella conduzione dello studio eziologico.
Table 5. Main methodological aspects to be considered in the conduction of an aetiological study.

(per esempio, analisi caso-controllo) è problematico per la piccola dimensione campionaria del *cluster* e la bassa informatività scientifica.¹⁰ Anche in questa fase l'analisi spaziale è un'alternativa fattibile e valida. La fase di studio eziologico richiede una valutazione preliminare della possibilità di utilizzare i dati spaziali a livello di piccola area come stima dell'esposizione a livello individuale.³⁰ Questo può avvenire tramite uno studio pilota di piccole dimensioni per investigare le fonti, l'entità delle esposizioni individuali, includendo se possibile dati ambientali, misurazioni dell'esposizione individuale tramite sensori individuali e biomarcatori per stimare la concentrazione della sostanza all'interno dell'organismo, tenendo anche conto dell'accumulo di mesi o anni (dose interna). Sulla base dello studio di fattibilità, i ricercatori devono definire un protocollo e implementare lo studio;

■ **la stima dell'esposizione su base geografica:** l'utilizzo di modelli di dispersione intorno a una sorgente di inquinamento puntiforme può essere più o meno adeguato, a seconda delle situazioni, poiché può determinare una misclassificazione dell'esposizione. Una misclassificazione non differenziale, che comporta una distorsione del rischio relativo verso l'ipotesi nulla, può aumentare il rischio di risultati falsamente negativi. Anche la misclassificazione differenziale può provocare una distorsione nella stima di effetto, di solito in direzione opposta rispetto all'ipotesi nulla. Questo problema si può verificare intorno a una sorgente di inquinanti, portando a rilevare un'associazione che in realtà non esiste. Persino un piccolo eccesso di rischio, se si verifica in una popolazione di grandi dimensioni, in termini assoluti si traduce in un impatto significativo in termini di casi attribuibili all'esposizione. Per malattie rare come i tumori infantili ciò ha conseguenze devastanti dal punto di vista della comunità. Un aspetto importante che deve essere soddisfatto affinché l'approccio di studio di piccole aree a fini eziologici sia ritenuto valido riguarda l'eterogeneità di esposizione tra le diverse aree: la condizione ideale è quella in cui le aree hanno esposizione nettamente diversa (per esempio, il confronto esposti/non esposti è migliore del contrasto tra basso e alto livello di esposizione all'inquinante);

■ **l'approccio semi-individuale** utilizza dati di esposizione a livello di area e confondenti a livello individuale, permettendo

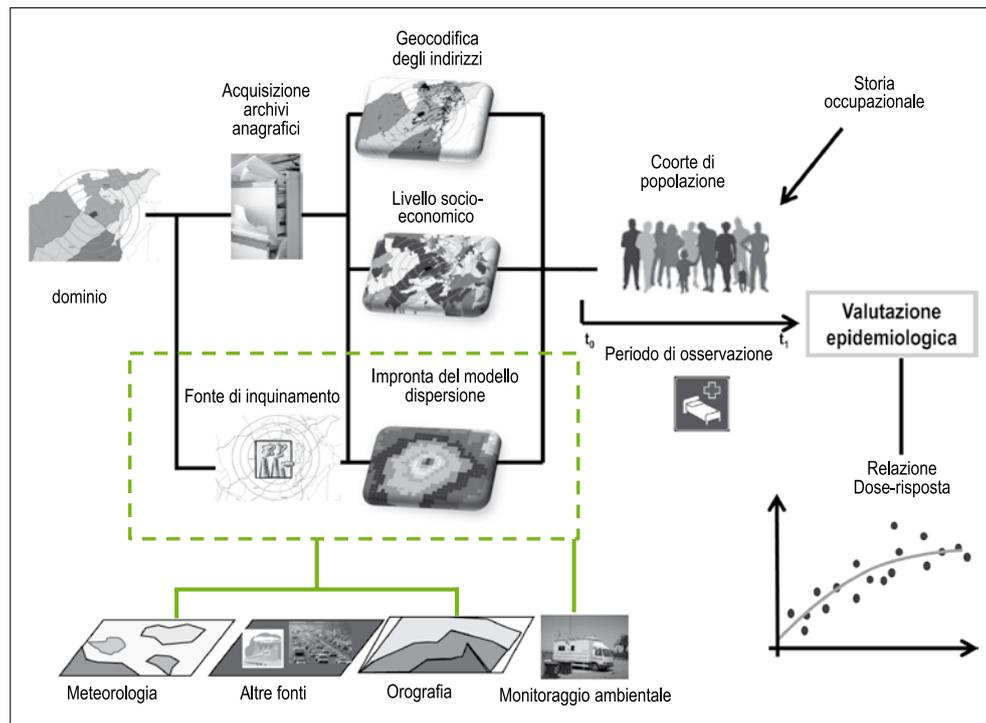
un controllo efficiente di questi ultimi, ma è suscettibile a confondimento o altro *bias* tra le diverse aree geografiche. Questi studi semi-individuali sono più efficienti se prevedono la rilevazione di dati di esposizione e di confondimento in una *survey* su un campione di popolazione. Un approccio promettente nelle analisi di questo tipo è costituito dai **modelli multilivello**, utilizzati per analizzare dati con una fonte di variabilità tra le aree (come le sezioni di censimento), tenendo conto della maggiore correlazione tra i soggetti che risiedono all'interno di una stessa area. Permettono di misurare quanta parte della variabilità della malattia nella popolazione è attribuibile all'area in cui vivono gli individui;⁴⁹

■ **l'approccio individuale** utilizza una misura individuale di esposizione, anche stimata da modello, per esempio in base alla distanza dalle strade ad alto traffico, e quindi viene analizzata con i metodi per gli studi di coorte. Gli studi longitudinali rappresentano l'approccio più valido per stimare l'associazione tra esposizione ambientale e malattia, ma sono costosi e richiedono molto tempo. Per aumentarne l'efficienza questi studi possono essere preceduti da un'indagine preliminare basata sull'approccio a piccole aree;

■ **lo studio di coorte residenziale** (figura 3) è un disegno di studio epidemiologico più adatto, che consente di valutare gli effetti a lungo termine dell'esposizione ambientale e stimare la durata esatta dell'esposizione. Questo tipo di studio richiede l'arruolamento di una coorte di popolazione, di solito su base anagrafica, la stima dell'esposizione ambientale sulla base della residenza di ogni individuo, oltre al controllo di una serie di fattori confondenti, come avviene in tutte le indagini epidemiologiche analitiche.⁵⁰

In prospettiva, nuovi approcci di studio potranno venire dalla storia naturale dei tumori come analisi dei cambiamenti multipli genetici ed epigenetici. Ogni tumore è, difatti, caratterizzato da una peculiare sequenza di cambiamenti genetici ed epigenetici che portano la cellula a deviare dal suo comportamento fisiologico.⁵¹ I primi di questi cambiamenti sembrano avviare la cellula in uno stato pre-canceroso, che può essere presente nello stesso momento in diversi individui della popolazione. Un *cluster* spazio-temporale si può verificare se la stessa popolazione si trova esposta a un fattore di rischio comune che nei sogget-

Figura 3. Metodologia dello studio di coorte residenziale. (Fonte: Ancona 2014).⁵²
Figure 3. Methodology of the residential cohort approach. (Source: Ancona 2014).⁵²



ti già predisposti provoca la progressione verso il tumore. Tale evidenza supporta l'importanza di utilizzare biomarcatori di rischio tumorale che riflettono cambiamenti a livello genetico o epigenetico;⁵³ tuttavia, il loro uso a livello di popolazione va validato, così come le possibili distorsioni nell'interpretazione dei risultati.

Tra gli approcci più innovativi nell'analisi di dati spaziali c'è quello delle partizioni isopopolazione, una metodologia che aumenta la possibilità di interpretare i dati aggregando le unità più piccole in base al principio di una numerosità di popolazione simile.⁵⁴

Questa terza fase fornisce l'opportunità di avviare azioni di sanità pubblica, come programmi di cessazione dell'abitudine al fumo, valutazioni del rischio sanitario, bonifiche ambientali. Azioni che non dovrebbero essere ritardate in attesa dei risultati dell'indagine epidemiologica di tipo eziologico in corso.

CONCLUSIONI

Le strutture sanitarie continuano a ricevere centinaia di segnalazioni di sospetti *cluster* ogni anno, che determinano un notevole investimento di risorse da parte delle istituzioni interessate. In molti casi il *cluster* sospetto non viene confermato e l'indagine è in grado di sostenere che nella popolazione in studio il rischio di malattia non è superiore a quello delle aree di riferimento.

Nei casi, invece, in cui l'eccesso di rischio viene confermato, la conduzione dell'indagine è spesso complessa e l'approccio in fasi (proposto oltre vent'anni fa dai CDC)^{3,4} può essere ancora utile per disporre di criteri condivisi dalla comunità scientifica, che consentono di decidere quali siano gli elementi a favore del proseguimento o dell'interruzione dell'indagine.

Per la scelta di quale debba essere il livello di approfondimento ottimale dell'indagine di *cluster* (per esempio, devono essere valutati in modo approfondito tutti i fattori di rischio sospet-

ti presenti nell'area?) occorre tener conto di due aspetti rilevanti. Il primo riguarda l'allarme generato dal sospetto *cluster* nella popolazione. In genere, la mancanza di risposte tempestive ed esaurienti da parte delle autorità sanitarie può essere percepita come inadeguatezza e incapacità a gestire l'emergenza e può avere l'effetto di aumentare l'allarme. L'altro aspetto riguarda la decisione di investire risorse per condurre studi che hanno una bassa probabilità di essere informativi; infatti, gli studi condotti fino ad oggi in questo ambito mostrano gli importanti limiti metodologici dell'indagine di *cluster* che rendono bassa la probabilità di produrre nuova conoscenza sull'eziologia della malattia in studio. Anche se, in alcuni casi, l'analisi di *cluster* può essere utile a formulare ipotesi eziologiche.^{2,10}

Uno degli aspetti che non andrebbe trascurato è che l'indagine di *cluster* di tumori in una popolazione, anche quando limitata alla fase 1 o ad alcuni aspetti della fase 2, può essere un'opportunità per informare la popolazione sui fattori di rischio individuali e per potenziare interventi di prevenzione e di screening nella popolazione.

Essere in grado di trasmettere in modo chiaro informazioni che derivano da un'approfondita revisione della letteratura scientifica può, inoltre, essere di supporto alla comunicazione di risultati che nella maggior parte dei casi sono giudicati dalla popolazione non esaurienti e non conclusivi. In popolazioni in cui sono presenti fonti di inquinamento ambientale è di fondamentale importanza verificare, attraverso i dati ambientali disponibili, se i livelli di inquinanti presenti possano rappresentare o meno un potenziale rischio per la popolazione; in caso di mancanza di dati, è importante sostenere la richiesta di condurre campagne di misure ambientali.

In conclusione, dalla pubblicazione delle linee guida dei CDC del 1990³ ci sono stati molti cambiamenti nella qualità dei dati, nel livello di tecnologia e nei mezzi di cui si dispone per l'analisi

si di sospetti *cluster*. Le fonti dei dati sono diventate più ricche e l'ampia letteratura prodotta in due decenni di indagini di *cluster* di tumori consente oggi una revisione critica degli approcci utilizzati. Tali indagini continuano, tuttavia, a presentare molti aspetti critici da considerare ogniqualvolta si abbia a che fare con una denuncia di sospetto *cluster*. Anche se la maggior parte dei *cluster* indagati rimane senza una spiegazione certa, si ritiene che uno degli obiettivi importanti sia quello di aumentare le conoscenze degli epidemiologi e di chi si occupa di sanità pubblica su-

gli strumenti disponibili e sui limiti della ricerca epidemiologica riguardo a questo tema. All'allarme per un sospetto *cluster* di casi di tumore si dovrebbe essere sempre capaci di dare una risposta, anche se ciò implica l'ammissione dei limiti delle proprie conoscenze e l'incertezza nell'interpretazione dei risultati disponibili. Gli aspetti legati alla comunicazione dei risultati rappresentano un ulteriore elemento di criticità degli studi di *cluster*.

Conflitti di interesse dichiarati: nessuno.

BIBLIOGRAFIA

- Martland H, Humphries R. Osteogenic sarcoma in dial painters using luminous paint. *Arch Pathol* 1929;7:406-17.
- Goodman M, Naiman JS, Goodman D, LaKind JS. Cancer clusters in the USA: what do the last twenty years of state and federal investigations tell us? *Crit Rev Toxicol* 2012; 42(6):474-90.
- National Center for Environmental Health, CDC, Atlanta, Georgia. Guidelines for investigating clusters of health events. *MMWR Recomm Rep* 1990;39(RR-11):1-23.
- National Center for Environmental Health, CDC, Atlanta, Georgia. Investigating suspected cancer clusters and responding to community concerns: guidelines from CDC and the Council of State and Territorial Epidemiologists. *MMWR Recomm Rep* 2013;62(RR08):1-24.
- Hogan L, Rutherford B, Mitchell V (eds). *Guidelines for the Management of Inquiries Related to Cancer Concerns or Suspected Cancer Clusters*. Baltimore (MD), Maryland Department of Health and Mental Hygiene, 2015. Disponibile all'indirizzo: http://phpa.dhmm.maryland.gov/cancer/SiteAssets/SitePages/Cancer%20Clusters/Maryland%20State%20Guidelines%20on%20Cancer%20Clusters_10.26.2015.pdf
- Surveillance and Assessment Alberta Health and Wellness. *Guidelines for the Investigation of Clusters of Non-Communicable Health Events*. Edmonton (AB) 2011. Disponibile all'indirizzo: <http://www.health.alberta.ca/documents/Investigation-Clusters-Guidelines-2011.pdf>
- Germonneau P, Tillaut Hélène, Gomes Do Espírito Santo E (eds). *Guide méthodologique pour l'évaluation et la prise en charge des agrégats spatio-temporels de maladies non infectieuses*. Institut De Veille Sanitaire 2005. Disponibile all'indirizzo: http://www.invs.sante.fr/publications/2005/guide_ast/guide.pdf
- Queensland Health Guidelines. *Assessment of clusters of non-communicable diseases*. Brisbane 2012. Disponibile all'indirizzo: https://www.health.qld.gov.au/publications/research-reports/population-health/cluster_assessment.pdf
- New Zealand Ministry of Health. *Investigating Clusters of Non-communicable Disease. Guidelines for public health units*. Wellington, Ministry of Health, 2015. Disponibile all'indirizzo: <https://www.health.govt.nz/system/files/documents/publications/investigating-clusters-non-communicable-disease-guidelines-may15-v3.pdf>
- Rothman KJ. A sobering start for the cluster busters' conference. *Am J Epidemiol* 1990; 132(1) Suppl:S6-13.
- Goodman M, LaKind JS, Fagliano JA et al. Cancer cluster investigations: review of the past and proposals for the future. *Int J Environ Res Public Health* 2014;11(2):1479-99.
- Otte KE, Sigsgaard TI, Kjaerulf J. Malignant mesothelioma: clustering in a family producing asbestos cement in their home. *Br J Ind Med* 1990;47(1):10-13.
- Trumbo CW. Public requests for cancer cluster investigations: a survey of state health departments. *Am J Public Health* 2000;90(8):1300-02.
- Nuckols JR, Ward MH, Jarup L. Using geographic information systems for exposure assessment in environmental epidemiology studies. *Environ Health Perspect* 2004;112(9): 1007-15.
- Chang ET, Adami HQ, Bailey WH et al. Validity of geographically modeled environmental exposure estimates. *Crit Rev Toxicol* 2014;44(5):450-66.
- Leyk S, Binder CR, Nuckols JR. Spatial modeling of personalized exposure dynamics: the case of pesticide use in small-scale agricultural production landscapes of the developing world. *Int J Health Geogr* 2009;8:17.
- Buonanno G, Marini S, Morawska L, Fuoco FC. Individual dose and exposure of Italian children to ultrafine particles. *Sci Total Environ* 2012;438:271-77.
- Beyea J. Geographic exposure modeling: a valuable extension of geographic information systems for use in environmental epidemiology. *Environ Health Perspect* 1999;107 Suppl 1:181-90.
- Chaix B, Méline J, Duncan S et al. GPS tracking in neighborhood and health studies: a step forward for environmental exposure assessment, a step backward for causal inference? *Health Place* 2013;21:46-51.
- Wild CP. Complementing the genome with an "exposome": the outstanding challenge of environmental exposure measurement in molecular epidemiology. *Cancer Epidemiol Biomarkers Prev* 2005;14(8):1847-50.
- CORDIS Europe – Community Research and Development Information Service. Projects and results. Health and Environment-wide Associations based on Large population Surveys. 2013. Disponibile all'indirizzo: http://cordis.europa.eu/project/rcn/110918_en.html
- Rare cancers Europe. 2016. Disponibile all'indirizzo: <http://www.rarecancerseurope.org/> (ultimo accesso: 21.03.2016).
- Wild CP, Kleinjans J. Children and increased susceptibility to environmental carcinogens: evidence or empathy? *Cancer Epidemiol Biomarkers Prev* 2003;12(12):1389-94.
- Linet MS, Wacholder S, Zahm SH. Interpreting epidemiologic research: lessons from studies of childhood cancer. *Pediatrics* 2003;112(1 Pt 2):218-32.
- Anderson LM, Diwan BA, Fear NT, Roman E. Critical windows of exposure for children's health: cancer in human epidemiological studies and neoplasms in experimental animal models. *Environ Health Perspect* 2000;108 Suppl 3:573-94.
- Greaves M. Infection, immune responses and the aetiology of childhood leukaemia. *Nat Rev Cancer* 2006;6(3):193-203.
- Figuerola ME, Chen SC, Andersson AK et al. Integrated genetic and epigenetic analysis of childhood acute lymphoblastic leukemia. *J Clin Invest* 2013;123(7):3099-111.
- Elliott P, Wakefield J. Disease clusters: should they be investigated and, if so, when and how? *J R Statist Soc A* 2001;164(1):3-12.
- Elliott P, Savitz DA. Design issues in small-area studies of environment and health. *Environ Health Perspect* 2008;116(8):1098-104.
- Olsen SF, Martuzzi M, Elliott P. Cluster analysis and disease mapping – why, when, and how? A step by step guide. *BMJ* 1996;313(7061):863-66.
- Thun MJ, Sinks T. Understanding cancer clusters. *CA Cancer J Clin* 2004;54(5):273-80.
- Hocking B, Gordon IR, Grain HL, Hatfield GE. Cancer incidence and mortality and proximity to TV towers. *Med J Aust* 1996;165(11-12):601-05.
- Dolk H, Shaddick G, Walls P et al. Cancer incidence near radio and television transmitters in Great Britain. I. Sutton Coldfield transmitter. *Am J Epidemiol* 1997;145(1):1-9.
- Michelozzi P, Capon A, Kirchmayer U et al. Adult and childhood leukemia near a high-power radio station in Rome, Italy. *Am J Epidemiol* 2002;155(12):1096-103.
- Bisanti L, Chiaffarino F, Chierici S et al. *Rapporto su un cluster di leucemie infantili a Milano nel periodo dicembre 2009 – gennaio 2010*. Regione Lombardia, ASL di Milano, 2010. Disponibile all'indirizzo: <http://www.asl.milano.it/user/Default.aspx?SEZ=10&PAG=88&NOT=1680>
- Dipartimento di Epidemiologia del Servizio Sanitario Regionale, ASL Roma 1 (ex ASL Roma E), Regione Lazio. *Valutazione dell'incidenza di leucemia infantile nel Municipio Roma XIII e analisi della frequenza di casi in prossimità di alcune fonti di emissione di campi elettromagnetici (ELF e Radiofrequenze) presenti nell'area*. Rapporto aprile 2013.
- Kulldorff M, Song C, Gregorio D, Samociuk H, DeChello L. Cancer map patterns: are they random or not? *Am J Prev Med* 2006;30(2) Suppl:S37-49.
- Lawson AB. Disease cluster detection: a critique and a Bayesian proposal. *Stat Med* 2006; 25(5):897-916.
- Lawson AB. Commentary: Assessment of chance should be central in investigation of cancer clusters. *Int J Epidemiol* 2013;42(2):448-49; discussion 455-56.
- McNally RJ, Stiller C, Vincent TJ, Murphy MF. Cross-space-time clustering of childhood cancer in Great Britain: evidence for a common aetiology. *Int J Cancer* 2014;134(1):136-43.
- Catelan D, Biggeri A. Multiple testing in disease mapping and descriptive epidemiology. *Geospatial Health* 2010;4(2):219-29.
- Coory M, Jordan S. Assessment of chance should be removed from protocols for investigating cancer clusters. *Int J Epidemiol* 2013;42(2):440-47.
- Elliott SJ, Cole DC, Krueger P, Voorberg N, Wakefield S. The power of perception: health risk attributed to air pollution in an urban industrial neighbourhood. *Risk Anal* 1999;19(4):621-34.
- Greaves M. Molecular genetics, natural history and the demise of childhood leukemia. *Eur J Cancer* 1999;35(2):173-85.
- Greenland S, Sheppard AR, Kaune WT, Poole C, Kelsh MA. A pooled analysis of magnetic fields, wire codes, and childhood leukemia. Childhood Leukemia-EMF Study Group. *Epidemiology* 2000;11(6):624-34.
- Ahlbom A, Day N, Feychting M et al. A pooled analysis of magnetic fields and childhood leukaemia. *Br J Cancer* 2000;83(5):692-98.
- Kulldorff M. A spatial Scan statistic. *Commun Statist Theory Meth* 1997;26(6):1481-96.

48. Studio epidemiologico sulle neoplasie del sistema emolinfopoietico nel territorio dell'Asl 11. Empoli 2012. Disponibile all'indirizzo: <http://www.usl11.toscana.it/finestranews.php?sezione=PREVNEW&tipo=6&destinatari=&utente=&numero=88>
49. Riva M, Gauvin L, Barnett TA. Toward the next generation of research into small area effects on health: a synthesis of multilevel investigations published since July 1998. *J Epidemiol Community Health* 2007;61(10):853-61.
50. Ancona C, Badaloni C, Mataloni F et al. Mortality and morbidity in a population exposed to multiple sources of air pollution: A retrospective cohort study using air dispersion models. *Environ Res* 2015;137:467-74.
51. Hanahan D, Weinberg RA. Hallmarks of cancer: the next generation. *Cell* 2011;144(5):646-74.
52. Ancona C, Mataloni F, Badaloni C et al. Lo studio di coorte di popolazione in aree a forte pressione ambientale: il programma ERAS Lazio. *Epidemiol Prev* 2014;38(2) Suppl 1: 158-61.
53. Goodman M, Bostick RM, Kucuk O, Jones DP. Clinical trials of antioxidants as cancer prevention agents: past, present, and future. *Free Radic Biol Med* 2011;51(5):1068-84.
54. Bronner A, Morignat E, Gay E et al. Iso-population partition: An innovative epidemiological approach to mapping and analyzing spatially aggregated data. *Prev Vet Med* 2015; 122(3):253-56.

