

**INDAGINE EPIDEMIOLOGICA  
PRESSO  
LA POPOLAZIONE RESIDENTE A  
FALCONARA MARITTIMA  
E COMUNI LIMITROFI**

**RISULTATI**

**PRIMO RAPPORTO**

A cura di:

Dott.ssa Elisabetta Meneghini  
Dott. Paolo Baili

Dott. Mauro Mariottini (ARPAM, per le parti di competenza)

Direzione  
Dott. Andrea Micheli



SC Epidemiologia Descrittiva e Programmazione Sanitaria  
Fondazione IRCCS "Istituto Nazionale dei Tumori"  
Via Venezian, 1  
20133, Milano  
Italia  
Tel. +39 02 2390 2867  
Fax +39 02 2390 3528  
e\_mail: [andrea.micheli@istitutotumori.mi.it](mailto:andrea.micheli@istitutotumori.mi.it)

Milano 28, novembre 2008

## ***INTRODUZIONE***

L'azienda API è una raffineria (da qui, anche Raffineria) di prodotti petroliferi sita dal 1950 nel comune di Falconara Marittima (da qui, anche Falconara) in provincia di Ancona che per la natura della produzione svolta determina esposizione a sostanze irritanti, tossiche, nocive e indicate o sospettate di potere cancerogeno. Falconara è da tempo oggetto di monitoraggio per la valutazione dei rischi atmosferici connessi alla presenza della Raffineria.

Il rischio in popolazione legato alla presenza della Raffineria è riferibile sia alla possibile esposizione di tipo acuto, sia a rischi d'esposizione di tipo cronico a benzolo e altri poliidrocarburi aromatici di riconosciuta azione oncologica. Gli effetti cronici d'esposizione cronica tendono spesso a comparire dopo un lungo periodo di tempo (sino a 30-40 anni) dall'occorrenza rendendo difficile lo studio della relazione tra esposizione e danno.

Nel luglio 2003 è stata stipulata tra l'Istituto Nazionale dei Tumori per lo Studio e la Cura dei Tumori e l'Agenzia Regionale Sanitaria (ARS) della regione Marche (da qui, anche Committente) una convenzione per uno "Studio di fattibilità per un'indagine epidemiologica presso la Raffineria API di Falconara". Nel marzo 2004 lo studio è arrivato alla conclusione che un'indagine epidemiologica era nelle condizioni analizzate fattibile [1].

A seguito dello studio di fattibilità, la regione Marche ha promulgato una delibera (n° 218 del 15 giugno 2004) in cui venivano approvati i risultati raggiunti e adottati i provvedimenti necessari per l'effettuazione di un'indagine epidemiologica presso la popolazione residente nei dintorni della raffineria API (da qui Indagine). Tale Indagine è stata affidata al Servizio di Epidemiologia Ambientale del dipartimento di Ancona dell'Agenzia Regionale Protezione Ambientale Marche (ARPAM) sotto la direzione scientifica dell'unità di Epidemiologia Descrittiva e Programmazione Sanitaria della Fondazione IRCCS "Istituto Nazionale dei Tumori" (INT).

Il presente primo rapporto descrive sinteticamente le attività svolte e i risultati parzialmente conclusivi a cui l'Indagine è pervenuta.

## 1. IL DISEGNO DELL'INDAGINE

Obiettivo dell'Indagine è la valutazione in popolazione dei rischi di aumentato decesso per tumore del sistema emolinfopoietico in funzione della distanza dalla Raffineria.

Il protocollo dell'Indagine ha definito:

- *l'area di studio*: comuni di Falconara Marittima, Montemarciano, Chiaravalle;
- *il periodo di studio*: dal 1/1/1994 al 31/12/2003;
- *il disegno epidemiologico*: studio di mortalità attuato con la tecnica casi-controlli su base di popolazione.

Per quanto di seguito non specificato si rinvia all'allegato protocollo, che descrive anche la distribuzione dei compiti e delle responsabilità tra i diversi attori coinvolti nell'Indagine [Annesso 1- Protocollo relativo all'indagine epidemiologica presso la popolazione residente a Falconara Marittima e comuni limitrofi].

L'Indagine è stata sviluppata secondo le seguenti fasi:

- 1) individuazione della lista dei *casi*, cioè delle persone decedute per tumore del sistema emolinfopoietico, residenti nell'*area di studio*, nel *periodo di studio*;
- 2) ricostruzione anno per anno del *periodo di studio*, delle popolazioni residenti nell'*area di studio*;
- 3) estrazione di due *controlli* su base di popolazione per ogni *caso*. I *controlli* vennero estratti con metodo casuale dalla popolazione residente nell'*area di studio* avendo come *referimento* il giorno, mese e anno di decesso del *caso* (da qui, anche data di *referimento*). L'estrazione casuale venne condotta entro strati d'appaiamento definiti da genere (es.: un *caso* maschio è stato appaiato ad un *controllo* maschio) ed età (ogni *controllo* è stato estratto dallo strato dei residenti con età pari a +/- 2.5 anni rispetto a quella della data di *referimento* del *caso*);
- 4) effettuazione di interviste a familiari dei *casi* e dei *controlli* (da qui, anche soggetti in studio), attraverso un questionario finalizzato a ricostruire la storia residenziale dei soggetti in studio e verificare l'esistenza di eventuali fattori di confondimento. Impegno venne profuso per garantire una adeguata partecipazione all'indagine (cioè per ridurre il numero di rifiuti alla somministrazione del questionario);
- 5) misurazione dell'esposizione residenziale individuale, definita come la distanza delle residenze dei soggetti in studio dal centroide della Raffineria;
- 6) analisi dei dati raccolti.

Nei paragrafi seguenti verranno sinteticamente descritte le attività sopra-indicate.

## **2. INDIVIDUAZIONE DEI CASI** (attività svolta da ARPAM)

Secondo le procedure in atto, le schede di morte redatte da medici, al momento del decesso, vengono archiviate presso il distretto sanitario del deceduto. Nel caso di decesso al di fuori del comune di residenza, l'ASL del comune di decesso (anche fuori regione) invia copia della scheda di morte all'ASL del comune di residenza, che quindi registra tutti i decessi dei residenti indipendentemente dal comune dove sia avvenuto l'evento. Per individuazione dei *cas*i relativi all'Indagine sono state esaminate tutte le "schede di morte ISTAT" del *periodo di studio* archiviate presso il distretto sanitario Ancona nord a Falconara Marittima.

Sono stati definiti *cas*i, i deceduti nel periodo dal 1/1/1994 al 31/12/2003, residenti alla data di decesso nei comuni di Falconara Marittima, Montemarciano, e Chiaravalle, nella cui "scheda di morte" compariva in almeno una delle tre cause di decesso indicate, la dizione tumore del sistema emolinfopoietico con codice di classificazione ICD-9 200-208 (secondo la IX modica di classificazione del codice internazionale delle malattie [2]). In presenza di cause di decesso dubbie e/o con segnalazione di sintomi e di segni compatibili con patologie tumorali emolinfopoietiche (44 soggetti) si è proceduto alla ricerca ed alla consultazione delle cartelle cliniche riferite agli ultimi ricoveri ospedalieri del deceduto. Le cartelle cliniche su cui permanevano dubbi (4 soggetti) sono state inviate in modo anonimo ad un consulente esterno e indipendente per il giudizio definitivo. La valutazione delle cartelle cliniche ha permesso l'individuazione di 22 casi.

In complesso sono stati individuati 177 casi: 97 residenti a Falconara Marittima, 27 a Montemarciano e 53 a Chiaravalle. La Tabella 1 descrive la frequenza dei casi per sesso e (a) anno di calendario, (b) età e (c) patologia.

## **3. RICOSTRUZIONE DELLE POPOLAZIONI E INDIVIDUAZIONE DEI CONTROLLI**

Il presente paragrafo descrive la fase di raccolta delle liste anagrafiche e ricostruzione delle popolazioni residenti nell'*area di studio* nel *periodo di studio* e le procedure di appaiamento tra *cas*o e i rispettivi *controlli*.

### **3.1 Raccolta e ricostruzione delle liste anagrafiche** (attività svolta da ARPAM)

Gli "uffici di stato civile" dei tre comuni compresi nell'*area di studio* hanno fornito le liste anagrafiche dei propri residenti aggiornate al 2005. Sono stati quindi preparati:

- elenchi con la lista dei residenti nell'*area di studio* per ogni anno dal 1994 al 2003. L'estrazione per anno è stata fatta tenendo conto della data di nascita o di immigrazione e della data di morte o di emigrazione. Ogni lista relativa ad un anno di calendario comprende informazioni relative ai movimenti demografici dei soggetti residenti in quell'anno (per esempio comprende i deceduti in quell'anno di calendario, gli emigrati in quell'anno, gli immigrati in quell'anno, eventuali presenze inferiori alla durata dell'intero anno, ecc). A partire dalle liste annuali, attraverso l'individuazione di univoci codici individuali si è quindi proceduto alla ricostruzione dell'andamento demografico nelle popolazioni in studio
- Per ogni residente (in ogni anno) sono stati resi disponibili:
  - codice individuale univoco per tutte le liste anagrafiche annuali;
  - nome, cognome, data di nascita, sesso;
  - via di residenza (alla data di morte o di emigrazione o al 2005);
  - codici familiari (attribuzione dalle anagrafi comunali di un codice numerico ai soggetti presenti nello stesso stato di famiglia aggiornato al 2005);
  - eventuale data di morte;
  - eventuale data di immigrazione e di emigrazione.

### **3.2 Appaiamento** (attività svolta da INT)

La procedura di appaiamento caso-controllo è stata effettuata secondo i metodi procedurali dello studio case-control nested in the cohort [3-4]. Il campionamento è stato quindi eseguito nella 'popolazione a rischio' [3] e nel rispetto del presupposto di indipendenza dei campioni [4]. Per ogni *caso*, è stata individuata la popolazione a rischio (*set a rischio*), composta dai soggetti dei tre comuni e aventi le caratteristiche richieste per l'appaiamento (l'individuazione di strati di popolazioni per sesso e età - +/- 2.5 anni), essendo in vita al momento della data di *referimento* del decesso del *caso*, dalla quale estrarre 2 controlli con metodo casuale. La procedura ammette che un *caso* il cui evento sia occorso successivamente in senso temporale alla data di occorrenza di un altro *caso*, possa appartenere *al set a rischio* di quest'ultimo e sia quindi eligibile ad essere estratto anche come *controllo*. Per rispettare il principio d'indipendenza statistica tra campioni, l'estrazione di un nuovo *set a rischio* è stata eseguita dopo la reintroduzione in popolazione dei *controlli* già estratti, permettendo così che un nuovo *controllo* potesse anche essere un soggetto già estratto in un precedente *set a rischio*.

Maggiori dettagli relativamente al contenuto di questo paragrafo sono disponibili nell'Annesso 2 - Relazione sull'appaiamento caso-controllo, già inviato al Committente.

**Tabella 1. Frequenza dei casi per (a) anno di calendario, (b) età, (c) patologia e sesso**

(a)	anno	Maschi	Femmine	Totale
	1994	5	6	11
	1995	6	7	13
	1996	11	7	18
	1997	6	14	20
	1998	12	10	22
	1999	8	6	14
	2000	8	9	17
	2001	9	9	18
	2002	13	12	25
	2003	11	8	19
	Totale	89	88	177

(b)	Età, anni	Maschi	Femmine	Totale
	<20	1	1	2
	20 - 39	5	1	6
	40 - 59	11	4	15
	60 - 79	48	43	91
	≥ 80	24	39	63
	Totale	89	88	177

(c)	ICD9	Maschi	Femmine	Totale
	201	2	1	3
	202.1	1	0	1
	202.8	21	27	48
	202.9	6	1	7
	203	14	23	37
	203.1	1	0	1
	203.8	1	0	1
	204	7	3	10
	204.1	13	13	26
	204.9	0	2	2
	205	10	6	16
	205.1	4	2	6
	205.2	1	1	2
	205.9	0	1	1
	207	0	1	1
	207.2	0	1	1
	208	5	4	9
	208.9	3	2	5
	Totale	89	88	177

## 4. INTERVISTE

L'Indagine ha previsto la somministrazione di un questionario (tramite intervista *de visu*) ai parenti dei *casì* e dei *controlli* per registrare, relativamente ai *soggetti* interessati, la storia occupazionale e residenziale, le eventuali patologie pregresse e familiari e l'eventuale esposizione a fattori di confondimento. L'attività ha comportato più sottofasi.

### 4.1 Individuazione dei parenti (attività svolta da ARPAM)

L'individuazione dei parenti dei soggetti arruolati nello studio è stata realizzata utilizzando molteplici fonti:

- selezione dalle anagrafi comunali dei familiari degli arruolati tramite i codici familiari descritti nel paragrafo 2.1;
- ricerca manuale dei recapiti telefonici e delle residenze attraverso gli elenchi telefonici;
- ricerca personale in loco di informazioni per consentire un contatto diretto con i familiari dei soggetti in studio non rintracciati. Per aumentare l'adesione all'Indagine sono state avviate iniziative di divulgazione ed informazione attraverso comunicati stampa, interviste televisive e partecipazione a pubblici eventi e attraverso il coinvolgimento dei medici di medicina generale, dei parroci, e dei comitati cittadini.

Per l'individuazione dei familiari dei *casì* e dei *controlli*, da sottoporre ad intervista diretta, sono stati contattati con priorità i soggetti con il grado di parentela più prossimo (coniuge, figli). Nell'eventualità di un'indisponibilità del soggetto, si è proceduto al contatto personale con familiari di più lontano grado di parentela (fratelli, nipoti). La selezione degli intervistandi è stata effettuata da ARPAM in cieco; non era quindi nota la composizione dei set caso-controlli.

### 4.2 Contatto con gli intervistati (attività svolta da ARPAM)

Individuato il parente da intervistare è stata effettuata una prima telefonata al familiare per fornire indicazioni preliminari sull'Indagine e per preannunciare l'invio presso la sua residenza di una busta contenente il seguente materiale:

- lettera illustrativa e brochure contenenti le finalità dello studio, la modalità di raccolta dati e i riferimenti per ricevere ulteriori informazioni e chiarimenti;
- informativa sulla raccolta e trattamento dei dati personali/sensibili in ottemperanza alla legge sulla privacy (D. Lgs. 196/03);
- lettera per il consenso al trattamento dei dati personali/sensibili da ritirare, firmata, in occasione dell'intervista;
- lettera di presentazione e sollecito alla partecipazione allo studio a firma dei sindaci;

- questionario con le domande da porre all'intervistato.

La disponibilità all'intervista è stata ottenuta mediante una successiva telefonata al familiare dell'arruolato per fissare l'appuntamento c/o la sua abitazione o presso sedi alternative rese disponibili da sindaci, parroci, forze dell'ordine, per effettuare la somministrazione del questionario.

#### **4.3 Questionario e interviste** (attività svolta da ARPAM)

Utilizzando come base un ampio questionario già validato e utilizzato in analoghi studi precedenti è stato predisposto un questionario (sottoinsieme informativo del precedente) volto alla raccolta di informazioni sui *casi* e i *controlli* relativamente a:

- dati anagrafici/stato civile/grado di istruzione;
- esposizione al fumo;
- patologie individuali pregresse;
- patologie familiari;
- esposizioni professionali;
- storia residenziale nei 40 anni antecedenti l'anno di *riferimento*. L'informazione relativamente alla storia residenziale ha riguardato: tipo di abitazione, tipo di riscaldamento, vicinanza dell'abitazione a possibili fonti di inquinamento (elettrodotti, stazioni radio-base, discariche, distributori o depositi carburanti) e a strade interessate a traffico auto veicolare.

Le interviste sono state effettuate da due intervistatori, preventivamente formati, presso il domicilio dei familiari dei *soggetti in studio* o nelle sedi alternative precedentemente indicate.

#### **4.4 Adesione e soggetti in analisi** (attività svolta da INT)

Dei 531 (177 casi e 354 controlli) soggetti individuati dalla procedura di appaiamento caso-controllo come *soggetti in studio*, 493 (il 93%) sono stati rintracciati (165 *casi* e 328 *controlli*). Tra loro hanno accettato di rispondere all'intervista i familiari di 346 (95 *casi* e 251 *controlli*) (da qui, *soggetti in analisi*). Sulla base dell'adesione raggiunta, l'attribuzione dell'esposizione residenziale cumulata ad ogni soggetto (calcolabile dalla storia abitativa fornita nel questionario) è stata possibile per il 70% dei soggetti rintracciati ovvero per il 65% del totale dei *soggetti in studio*.

Sulla base dell'unica informazione disponibile per tutti i soggetti rintracciati, rispondenti e non rispondenti, ovvero la residenza al 2005 o l'ultima residenza per i deceduti (non la storia residenziale, come sarebbe necessario), è stato calcolato che *i casi* con residenza distante dalla raffineria meno di 2 km risultano sovra-rappresentati tra i *soggetti in analisi* per circa il 10%,

mentre i *controlli* con residenza distante più di 2 km dalla Raffineria risultano sotto-rappresentati del 10% circa, con possibilità di sovrastimare il rischio di morte del 20% circa [5]. Maggiori informazioni sull'analisi relativa alla stima dell'*errore di non rispondenza* sono disponibili nell'Annesso 3 - Relazione sui soggetti non rispondenti; l'analisi di cui alla relazione in Annesso 3 è stata condotta prima che altri due *controlli* oltre quelli considerati dessero la loro disponibilità all'intervista.

Dei 251 *controlli* con intervista, solo 166 sono quelli appaiati ad uno dei 95 *cas*i rispondenti e non tutti i 95 set sono completi: 71 set sono composti da 1 *cas*o e 2 *controlli*, e 24 set sono composti da 1 *cas*o e 1 *controllo*. Quindi i soggetti utilizzabili per l'analisi statistica sono circa il 50% dei 531 inizialmente individuati.

## **5. ESPOSIZIONE INDIVIDUALE**

### **5.1 Misurazione dell'esposizione residenziale individuale** (attività svolta da ARPAM)

La misurazione dell'esposizione individuale è stata effettuata per fasi:

1. individuazione del centroide dell'area dello stabilimento API mediante software GIS "Mapinfo Pro" v. 8 e calcolo sul poligono dell'area;
2. rilevazione manuale delle coordinate geografiche mediante GPS (Global Positioning System) di tutte le residenze abitate dai *casi* e dai *controlli* nei 40 anni precedenti la data di riferimento nei comuni di Falconara Marittima, Montemarciano e Chiaravalle;
3. misurazione delle distanze di tutte le residenze (rilevate in fase 2) dal centroide della Raffineria. Per tutte le residenze che non sono state rintracciate (perché l'intervistato non ricordava il numero civico della residenza del familiare o per l'impossibilità di rintracciare l'abitazione per effettuare la georeferenziazione) si è provveduto a loro georeferenziazione mediante software per la ricerca automatica delle coordinate geografiche (CSH Geolocate).

### **5.2 Raccolta ad-hoc su eventuali ulteriori confondenti** (attività in corso da parte di ARPAM)

Per eventuali futuri approfondimenti dell'attuale Indagine è stata preventivata una raccolta *ad hoc* su possibili confondenti presenti nell'*area di studio* relativamente a:

- distributori e depositi di carburante;
- aree destinate al trattamento e/o stoccaggio di rifiuti urbani o pericolosi;
- emittenti CEM RF, stazioni radio base per la telefonia mobile e per le trasmissioni radiotelevisive;
- linee elettriche ad alta tensione;
- sorgenti industriali CEM ELF rilevanti per intensità di emissione e collocazione in prossimità di aree residenziali.

## 6. ANALISI STATISTICA (attività svolta da INT)

L'analisi è stata indirizzata a valutare l'associazione tra rischio di morte per tumore del sistema emolinfopoietico ed esposizione residenziale legata alla presenza della raffineria API. Per la valutazione del rischio si è proceduto alla produzione di stime degli odds ratio (OR) di esposizione (e dei relativi intervalli di confidenza al 95%) mediante l'applicazione di modelli di regressione logistica condizionata che includono la variabile di esposizione e i potenziali confondenti (variabili di appaiamento: età e sesso).

L'Annesso 4 – Programma di analisi statistica -, presenta i metodi di calcolo dell'esposizione residenziale - utilizzati in analoghi studi, e motiva le scelte di strategia d'analisi operate nella presente Indagine: l'obiettivo della strategia utilizzata attraverso il calcolo della storia residenziale cumulata è di acquisire la maggior informazione disponibile per *ogni soggetto in analisi*, ciò all'interno di una finestra temporale che si assume adeguata ai minori tempi di latenza tra esposizione e occorrenza di tumori del sistema emolinfopoietico (si è quindi assunto che i casi deceduti nel periodo 1994 - 2003 siano stati ipoteticamente determinati dai 20 anni precedenti l'anno di riferimento con esclusione degli ultimi 5 anni, assunti come distanza media tra occorrenza della malattia e decesso a causa della malattia. L'Annesso 5 – Calcolo dell'esposizione residenziale cumulata (e dell'esposizione occupazionale cumulata), descrive in dettaglio come è stato ricostruito per ogni soggetto il periodo di 15 anni, assunto come periodo di esposizione. Per ogni *soggetto in analisi* l'esposizione residenziale cumulata è stata calcolata come media pesata dell'inverso delle distanze abitazione-raffineria (misurate attraverso georeferenziazione) relativamente alle case abitate nel periodo di esposizione (15 anni), dove il peso attribuito ad ogni abitazione è pari al numero di anni lì vissuti, durante il periodo di esposizione.

Caratteristiche delle abitazioni (zona urbana, abitazione a piano terra, riscaldamento a legna/carbone, coabitanti fumatori, campi coltivati con fitofarmaci, traffico a meno di 100 metri) e presenza di strutture esponenti a meno di 200 m (distributori di carburante, tralicci per linee elettriche, centrali elettriche, antenne CEM, industrie tessili/calzaturiere, industrie metalmeccaniche, lavanderie, discariche, cave/miniere) sono state considerate potenziali confondenti ed è stato calcolato un indicatore-punteggio pari alla frazione temporale del corrispondente periodo di esposizione.

Per ogni *soggetto in analisi* è stata ricostruita la storia lavorativa all'interno del periodo di esposizione; è stato quindi calcolato un punteggio di rischio occupazionale pari alla frazione del periodo di esposizione con occupazioni a rischio di tumori del sistema emolinfopoietico secondo

due diversi metodi: la matrice professione-esposizione MATline [6] e la matrice OCCAM [7]. La condizione di fumatore o ex-fumatore o non fumatore al termine del periodo di esposizione è stata derivata dalle informazioni relative fornite dai questionari, dove erano anche riportate le eventuali età d'inizio e di fine esposizione a tabacco. Infine il grado d'istruzione, quale proxy di status sociale, e lo stato civile sono stati considerati in analisi come potenziali confondenti. Nelle analisi, indipendenti, relative alle donne e agli uomini, la scelta delle covariate da introdurre nei modelli di regressione ha tenuto conto non solo dell'effetto sulla stima dei parametri ma anche della necessità di limitarne il numero così da adeguarlo al numero dei set rispettivamente analizzabili [8].

### **6.1 Potenza dello studio**

Novantatre tra i 95 set in studio disponevano dell'informazione relativa all'esposizione residenziale cumulata per almeno 1 *caso* ed 1 *controllo* appaiato. Assumendo una prevalenza di esposizione nei *controlli* pari a 0.33 (III terzile del punteggio di esposizione), il minimo OR rilevabile con potenza  $1 - \beta = 0.80$  ed errore di primo tipo  $\alpha = 0.05$  in test ad una coda - considerando anche che non tutti i set sono composti di 1 *caso* e 2 *controlli* - risulta compreso tra 1.91 e 2.10. Nel maschi, 49 set erano disponibili per l'analisi logistica condizionata, associati ad un OR compreso tra 2.43 e 2.78, nelle donne i set erano 44 per un OR compreso tra 2.55 e 2.94 [9].

Le analisi sono state eseguite mediante utilizzo di software STATA, versione 9.0 (Stata Corporation, College Station, TX, USA, 2005).

## 7. RISULTATI

### 7.1 Caratteristiche dei soggetti, delle loro abitazioni e rischio occupazionale

Le caratteristiche dei soggetti analizzati, 95 *casi* e 166 *controlli* appaiati, e la loro associazione con lo status caso controllo sono presentate in Tabella 2. I *casi* hanno un'età media di  $74.2 \pm 14.1$  anni; per i *controlli* la media appaiata (valore medio delle medie entro set) è di  $74.2 \pm 13.8$  anni.

**Tabella 2. Caratteristiche dei soggetti in studio**

	Casi		Controlli		OR <sup>a</sup>	IC 95%
	N	%	N	%		
Età <sup>b</sup>						
0 - 19	1	1.0	2	1.2		
20 - 39	3	3.2	5	3.0		
40 - 59	7	7.4	12	7.2	-	-
60 - 79	52	54.7	85	51.2		
≥ 80	32	33.7	62	37.4		
Sesso <sup>b</sup>						
Maschi	49	51.6	87	52.4		
Femmine	46	48.4	79	47.6	-	-
Istruzione						
Nessuna – lic. media inferiore	79	83.2	144	86.8	1	-
≥ licenza media superiore	16	16.8	22	13.2	1.38	0.65 – 2.92
Stato civile						
Non coniugato	44	46.3	72	43.4	1	-
Coniugato <sup>c</sup>	51	53.7	94	56.6	0.79	0.42 – 1.47
Fumo						
Non fumatore	51	53.7	93	56.0	1	-
Ex fumatore	29	30.5	43	25.9	1.43	0.66 – 3.09
Fumatore	15	15.8	30	18.1	0.98	0.41 – 2.38
Intensità di fumo <sup>d</sup>						
I terzile: 0	51	53.7	95	57.2	1	-
II terzile: >0, ≤58.4	13	13.7	16	9.7	1.68	0.66 – 4.25
III terzile: >58.4	31	32.6	55	33.1	1.24	0.56 – 2.77
Genitori/fratelli/sorelle con tumore del sistema emolinfopoietico						
No	71	86.6	144	97.3	1	-
Sì	11	13.4	4	2.7	3.81	1.18 – 12.3

<sup>a</sup> OR stimato con regressione logistica condizionata, variabili di appaiamento: età e sesso. <sup>b</sup> Variabile di appaiamento. <sup>c</sup> Non coniugato: celibe/nubile, separato/divorziato, o vedovo. <sup>d</sup> Intensità di fumo = anni di fumo x Ln(media giornaliera di sigarette); terzili calcolati sulla distribuzione dei controlli.

Alla fine del periodo di esposizione, la frequenza di non fumatori, è simile tra *casi* e *controlli* (54% e 56 % rispettivamente); gli ex-fumatori sono invece di poco più frequenti nei *casi* che nei *controlli* (31% vs. 26%), mentre i fumatori sono similmente distribuiti. Rispetto ai non fumatori, l'OR di essere *caso* è 1.68 (IC 95%: 0.66 – 4.25) e 1.24 (IC 95%: 0.56 – 2.77), rispettivamente per il secondo e il terzo terzile della distribuzione dell'indicatore intensità di fumo (calcolato come anni di fumo moltiplicato LN(media giornaliera di sigarette)) (p trend = 0.638).

I soggetti con familiari (genitori o fratelli/sorelle) che hanno avuto un tumore del sistema emolinfopoietico sono significativamente più frequenti nei *casi* che nei *controlli*, OR = 3.81 (IC 95%: 1.18 – 12.3). Tuttavia questa variabile non può essere considerata un potenziale confondente perchè, anche qualora essa risultasse associata all'esposizione residenziale cumulativa, il tumore insorto nei familiari potrebbe esso stesso dipendere dall'esposizione residenziale: il suo significato dovrebbe essere ulteriormente indagato in una successiva indagine. Le Tabelle 3 e 4 presentano la distribuzione di frequenza, nei *casi* e nei *controlli*, degli anni trascorsi in abitazioni con caratteristiche associabili a fattori di rischio per l'insorgenza di tumori del sistema emolinfopoietico, o poste nelle vicinanze di strutture a rischio.

La tabella 5 mostra l'associazione tra lo status caso/controllo e la frazione del periodo di esposizione vissuta in abitazioni con caratteristiche a rischio; l'OR in questo caso stima il rapporto tra gli odds di essere *caso* per incremento unitario della frazione di periodo. Considerato che alcune caratteristiche e strutture a rischio sono risultate scarsamente presenti nel territorio indagato, l'analisi si è concentrata solo su quelle variabili che presentavano una numerosità sufficiente per produrre stime attendibili.

I *controlli*, più frequentemente dei *casi*, non hanno avuto coabitanti fumatori per tutto il periodo di esposizione (49% vs. 42%, Tabella 3), viceversa i *casi*, più frequentemente dei *controlli*, hanno coabitato per più di 10 anni con fumatori (52% vs. 46%, Tabella 3); l'OR di essere *caso* per incremento unitario della frazione del periodo di esposizione a fumo passivo è 1.37 (IC 95%: 0.82 – 2.30, p = 0.233) (Tabella 5).

I *casi* che non hanno mai abitato nei pressi di un distributore di benzina sono 53% contro 64% dei *controlli* (Tabella 4); l'OR per aumento unitario di frazione di periodo esposta è 1.44 (IC 95%: 0.83 – 2.49, p = 0.190). I *casi* sono più frequenti dei *controlli* nella categoria che non ha mai abitato vicino a centrali elettriche (73% vs. 66%, Tabella 4) e meno frequenti nella categoria che ha abitato più di 10 anni vicino a centrali (19% vs. 29%, Tabella 4) o vicino ad antenne CEM (13% vs. 20%, Tabella 4); gli OR di esposizione risultano 0.66 (IC 95%: 0.34 - 1.28, p = 0.210) per l'esposizione ad centrali elettriche e 0.51 (IC 95%: 0.22 - 1.19, p = 0.226) per l'esposizione ad antenne CEM.

**Tabella 3. Caratteristiche delle case abitate nel periodo di esposizione**

<i>Caratteristiche dell'abitazione</i>	anni	Casi		Controlli	
		N	%	N	%
Zona urbana	0	8	8.6	13	7.9
	1 – 9	2	2.1	10	6.1
	10 – 15	83	89.3	142	86.0
Unifamiliare	0	66	71.0	119	72.1
	1 – 9	2	2.1	9	5.5
	10 – 15	25	26.9	37	22.4
Piano terra	0	69	74.2	115	69.7
	1 – 9	3	3.2	11	6.7
	10 – 15	21	22.6	39	23.6
Coabitanti fumatori	0	39	41.9	81	49.4
	1 – 9	6	6.5	7	4.3
	10 – 15	48	51.6	76	46.3
Riscaldamento autonomo a legna o carbone	0	80	89.9	144	92.3
	1 – 9	2	2.2	8	5.1
	10 – 15	7	7.9	4	2.6
Aree coltivate nei pressi	0	75	80.7	132	80.0
	1 – 9	0	0.0	9	5.5
	10 – 15	18	19.3	24	14.5
Regolare uso di fitofarmaci <sup>a</sup>	0	81	95.3	140	97.9
	1 – 9	0	0.0	1	0.7
	10 – 15	4	4.7	2	1.4
Traffico a meno di 100 m	0	32	33.7	57	34.7
	1 – 9	4	4.2	11	6.7
	10 – 15	51	53.7	96	58.5

<sup>a</sup> Fungicidi, pesticidi, insetticidi o erbicidi.

**Tabella 4. Strutture presenti a meno di 200 metri dalle case abitate nel periodo di esposizione**

	Anni	Casi		Controlli	
		N	%	N	%
Deposito di carburante	0	87	93.6	159	96.4
	1 – 15	6	6.4	6	3.6
Distributore di carburante	0	49	52.7	106	64.2
	1 – 9	7	7.5	8	4.9
	10 – 15	37	39.8	51	30.9
Antenne CEM	0	74	85.1	121	78.6
	1 – 9	2	2.3	2	1.3
	10 – 15	11	12.6	31	20.1
Tralicci ELF	0	70	78.6	118	73.7
	1 – 9	3	3.4	7	4.4
	10 – 15	16	18.0	35	21.9
Centrale elettrica	0	65	73.0	107	66.4
	1 – 9	7	7.9	8	5.0
	10 – 15	17	19.1	46	28.6
Industria tessile o calzature	0	89	95.7	156	95.2
	10 – 15	4	4.3	8	4.8
Industria metalmeccanica	0	88	94.6	156	95.1
	1 – 15	5	5.4	8	4.9
Carrozzeria	0	76	83.5	136	82.4
	1 – 9	3	3.3	5	3.0
	10 – 15	12	13.2	24	14.6
Lavanderia	0	57	61.3	106	65.4
	1 – 9	9	9.7	10	6.2
	10 – 15	27	29.0	46	28.4
Raffineria di petrolio	0	91	97.8	157	95.2
	10 – 15	2	2.2	8	4.8
Discarica	0	92	89.9	164	99.4
	10 – 15	1	1.1	1	0.6
Cava o miniera	0	93	100.0	164	99.4
	1 – 15	0	0.0	1	0.6

**Tabella 5. Associazione tra status caso o controllo e frazione di periodo vissuta in abitazioni con caratteristiche a rischio: odds ratio (OR) e intervallo di confidenza al 95% (IC 95%)**

<i>Caratteristiche abitazione</i>	Set <sup>a</sup>	OR <sup>b</sup>	IC 95%	p trend <sup>c</sup>
Zona urbana	93	1.25	0.52 – 3.00	0.615
Unifamiliare	93	1.23	0.68 – 2.23	0.486
Piano terra	93	0.92	0.49 – 1.72	0.800
Coabitanti fumatori	92	1.37	0.82 – 2.30	0.233
Aree coltivate nei pressi	93	1.18	0.60 – 2.33	0.630
Traffico a meno di 100 m	93	0.92	0.53 – 1.60	0.765
<i>Strutture a meno di 200 m</i>				
Distributore carburante	93	1.44	0.83 – 2.49	0.190
Antenne CEM	86	0.51	0.22 – 1.19	0.226
Tralicci ELF	89	0.80	0.40 – 1.60	0.520
Centrale elettrica	89	0.66	0.34 – 1.28	0.210
Carrozzeria	91	0.93	0.44 – 1.92	0.851
Lavanderie	92	1.09	0.61 – 1.94	0.772

<sup>a</sup> Numero di set con informazione sulla caratteristica. <sup>b</sup> Odds ratio associato ad un incremento unitario della frazione di periodo trascorsa in case con caratteristiche a rischio o nelle vicinanze di strutture a rischio; OR stimato con regressione logistica condizionata. <sup>c</sup> Likelihood-ratio test.

La Tabella 6 riporta la frequenza dei *casi* e dei *controlli* in categorie di anni di lavoro in occupazioni a rischio, secondo i metodi DORS e OCCAM; la Tabella 7 mostra l'associazione con lo status caso/controllo. La percentuale di soggetti che nel periodo di esposizione non ha svolto alcuna occupazione a rischio è leggermente più alta secondo il metodo DORS (86% dei *casi* e 85% dei *controlli*, Tabella 6) che secondo il metodo OCCAM (79% dei *casi* e 79% dei *controlli*, Tabella 6), tuttavia per entrambi i metodi non vi sono differenze tra *casi* e *controlli* e l'OR di esposizione non è significativamente diverso da 1 (Tabella 7).

**Tabella 6. Numero di anni lavorativi con occupazioni a rischio per l'insorgenza di tumori del sistema emolinfopoietico**

	anni	Casi		Controlli	
		N	%	N	%
Esposizione occupazionale (metodo DORS)	0	77	85.5	140	85.4
	1 - 9	5	5.5	12	7.3
	10 - 15	8	9.0	12	7.3
Esposizione occupazionale (metodo OCCAM)	0	71	78.9	129	78.7
	1 - 9	10	11.1	15	9.1
	10 - 15	9	10.0	20	12.2

**Tabella 7. Associazione tra status caso/controllo e frazione di periodo caratterizzata da occupazioni a rischio: odds ratio (OR) e intervallo di confidenza al 95% (IC 95%)**

	Set <sup>a</sup>	OR <sup>b</sup>	IC 95%	p <sup>c</sup>
Esposizione professionale (metodo DORS)	92	0.98	0.38 – 2.53	0.974
Esposizione professionale (metodo OCCAM)	92	0.75	0.32 – 1.76	0.502

<sup>a</sup> Numero di set con informazione sulla caratteristica. <sup>b</sup> Odds ratio associato ad un incremento unitario della frazione di periodo con occupazioni a rischio; OR stimato con regressione logistica condizionata. <sup>c</sup> Likelihood-ratio test

Nell'insieme di tutti i *soggetti in analisi* (uomini e donne), l'analisi univariata relativa alle caratteristiche personali e delle case abitate nel periodo di esposizione e delle occupazioni svolte nel periodo di esposizione non evidenzia significative differenze tra *casi* e *controlli*. Tuttavia alcune deboli associazioni sono emerse relativamente a esposizione a fumo attivo, a fumo passivo e a distributori di benzina che sono più presenti tra i *casi* che tra i *controlli*, mentre l'esposizione a centrali elettriche e ad antenne CEM è più presente tra i *controlli* che tra i *casi*. Analizzando separatamente uomini e donne alcune associazioni divengono più marcate. Negli uomini, l'OR per fumo attivo (fumatore/ex-fumatore vs. non fumatore alla fine del periodo di esposizione) risulta pari a 1.92 (95% IC: 0.76 – 4.82). Nelle donne, l'OR per fumo passivo (coabitanti fumatori vs. nessun coabitante fumatore nel periodo di esposizione) risulta pari a 2.25 (IC 95%: 1.04 – 4.98) e l'OR per vicinanza ad antenne (almeno 1 abitazione vs. nessuna abitazione a meno di 200 m nel periodo di esposizione) risulta pari a 0.30 (IC 95%: 0.09 – 1.06). La variabile che considera gli anni trascorsi in abitazioni vicine ad antenne CEM può essere scarsamente attendibile. Dato che le antenne CEM sono generalmente di recente installazione e non esattamente riferibili al periodo di esposizione, i soggetti potrebbero essere mal classificati rispetto all'esposizione a campi elettromagnetici. In ogni caso, quando introdotta nei modelli di regressione tale variabile non modificava le stime degli OR.

## 7.2 Odds ratio (OR) di esposizione residenziale

Nel presente paragrafo sono presentati i risultati relativi alla stima dell'eventuale eccesso di rischio di morte per tumore emolinfopoietico misurato attraverso lo studio della variabile esposizione residenziale cumulativa. Come stima del rischio relativo, vengono presentati gli OR grezzi e corretti per le variabili che tra tutti i potenziali confondenti analizzati hanno mostrato un qualche effetto sulla stima dell'OR.

Nell'insieme di tutti i *soggetti in analisi* (uomini e donne) si osserva un modesto aumento non statisticamente significativo di mortalità per tumore del sistema emolinfopoietico all'aumentare dell'esposizione residenziale cumulata (Tabella 8). Gli OR grezzi per terzile di esposizione cumulata sono: 1, 1.03, (IC 95%: 0.53 - 2.00) e 1.14 (IC 95%: 0.60 - 2.17) (p trend = 0.683) (Tabella 8). Controllando per fumo attivo, fumo passivo, vicinanza ai distributori di carburante e vicinanza a centrali elettriche gli OR risultano 1, 0.97 (IC 95% 0.48 - 1.97) e 1.25 (IC 95% 0.62 - 2.43) (p trend = 0.525) (Tabella 8); includere le variabili di rischio occupazionale, calcolato sia con metodo DORS che OCCAM, non modifica la stima degli OR.

Per le donne, all'aumentare dell'esposizione residenziale cumulata, risulta un più marcato aumento di mortalità non statisticamente significativo; gli OR grezzi per terzile di esposizione residenziale cumulata sono: 1, 1.32 (IC 95%: 0.55 - 3.17) e 1.25 (IC 95%: 0.51 - 3.06) (p trend = 0.624) (Tabella 9). Controllando per fumo attivo, fumo passivo, vicinanza a distributori di carburante, vicinanza a centrali elettriche e rischio occupazionale (metodo DORS), gli OR divengono: 1, 2.06 (IC 95%: 0.71 - 5.98) e 1.87 (IC 95%: 0.64 - 5.49) (p trend = 0.277) (Tabella 9). Si raggiunge la significatività statistica controllando anche per la vicinanza ai tralicci (almeno 1 abitazione vs. nessuna abitazione a meno di 200 metri nel periodo di esposizione); nel sottogruppo delle 105 donne (40 casi e 65 controlli appaiati) con tutte le informazioni per questa analisi, gli OR grezzi sono: 1, 1.68 (IC 95%: 0.68 - 4.15) e 1.40 (IC 95%: 0.53 - 3.69) (p trend = 0.469); gli OR controllati per fumo attivo, fumo passivo, vicinanza a distributori di carburante, vicinanza a centrali elettriche, rischio occupazionale (metodo DORS) e vicinanza a tralicci risultano: 1, 3.32 (IC 95%: 1.02 - 10.8) e 2.44 (IC 95%: 0.69 - 8.61) (p trend = 0.114).

Negli uomini non si evidenzia nessun effetto (Tabella 10); gli OR grezzi, per terzile di esposizione cumulata, sono: 1, 0.77 (IC 95%: 0.29 - 2.07) e 0.99 (IC 95%: 0.41 - 2.40) (p trend = 1.000) (Tabella 10); nessuno dei potenziali confondenti, fumo, rischio occupazionale, caratteristiche delle abitazioni modifica sostanzialmente le stime OR grezze.

**Tabella 8. Odds ratio di tumore del sistema emolinfopoietico per terzili di esposizione residenziale cumulata. Uomini e donne. Analisi basata su 89 casi e 149 controlli appaiati**

	Terzili di esposizione residenziale cumulata					P trend*
	I	II	III			
OR grezzo	1	1.03 (0.53 - 2.00)	1.14 (0.60 - 2.17)			0.683
OR controllato per Distributori carburante <sup>¶</sup> , Centrali elettriche <sup>  </sup>	1	1.03 (0.52 - 2.05)	1.26 (0.63 - 2.53)			0.506
OR controllato per Fumo attivo <sup>†</sup> , Fumo passivo <sup>§</sup>	1	0.94 (0.47 - 1.88)	1.11 (0.58 - 2.13)			0.736
OR controllato per Fumo attivo <sup>†</sup> , Fumo passivo <sup>§</sup> , Centr. el. <sup>  </sup> , Distr. carb. <sup>¶</sup>	1	0.97 (0.48 - 1.97)	1.25 (0.62 - 2.43)			0.525

\* Log likelihood ratio test. <sup>†</sup> Fumo attivo: status di fumatore o ex-fumatore vs. non fumatore alla fine del periodo di esposizione. <sup>§</sup> Fumo passivo: frazione del periodo di esposizione trascorsa con coabitanti fumatori. <sup>¶</sup> Frazione del periodo di esposizione trascorsa in abitazioni a meno di 200 m da distributori di carburante. <sup>||</sup> Frazione del periodo di esposizione trascorsa in abitazioni a meno di 200 m da centrali elettriche.

**Tabella 9. Odds ratio di tumore del sistema emolinfopoietico per terzili di esposizione residenziale cumulata. Donne. Analisi basata su 44 casi e 73 controlli appaiati.**

	Terzili di esposizione residenziale cumulata					p trend*
	I	II	III			
OR grezzo	1	1.32 (0.55 - 3.17)	1.25 (0.51 - 3.06)			0.624
OR controllato per Fumo attivo <sup>†</sup> , Fumo passivo <sup>§</sup>	1	1.63 (0.62 - 4.31)	1.34 (0.51 - 3.53)			0.562
OR controllato per Distributori carburante <sup>¶</sup> , Centrali elettriche <sup>  </sup>	1	1.42 (0.56 - 3.55)	1.42 (0.54 - 3.72)			0.484
OR controllato per Fumo attivo <sup>†</sup> , Fumo passivo <sup>§</sup> , Centr. el. <sup>  </sup> , Distr. carb. <sup>¶</sup>	1	2.06 (0.71 - 5.98)	1.84 (0.63 - 5.40)			0.289
OR controllato per Fumo attivo <sup>†</sup> , Fumo passivo <sup>§</sup> , Centr. el. <sup>  </sup> , Distr. carb. <sup>¶</sup> , Rischio occupazionale (metodo DORS) <sup>§§</sup>	1	2.06 (0.71 - 5.98)	1.87 (0.64 - 5.49)			0.277

\* Log likelihood ratio test. <sup>†</sup> Fumo attivo: status di fumatore o ex-fumatore vs. non fumatore alla fine del periodo di esposizione. <sup>§</sup> Fumo passivo: almeno 1 abitazione vs. nessuna abitazione con presenza di fumatori nel periodo di esposizione. <sup>¶</sup> Distributori di carburante: almeno 1 abitazione vs. nessuna abitazione a meno di 200 metri dalla struttura nel periodo di esposizione. <sup>||</sup> Centrale elettrica: almeno 1 abitazione vs. nessuna abitazione a meno di 200 metri dalla struttura nel periodo di esposizione. <sup>§§</sup> Rischio occupazionale – metodo DORS: almeno 1 occupazione a rischio vs. nessuna occupazione a rischio di tumore del sistema emolinfopoietico secondo il metodo DORS.

**Tabella 10. Odds ratio (95% IC) di tumore del sistema emolinfopoietico per terzi di esposizione residenziale cumulata. Uomini. Analisi basata su 46 casi e 79 controlli appaiati.**

	Terzili di esposizione residenziale cumulata					P trend*
	I	II	III	III	P trend*	
OR grezzo	1	0.77 (0.29 - 2.07)	0.99 (0.41 - 2.40)			1.000
OR controllato per Distributori carburante <sup>¶</sup> , Centrali elettriche <sup>  </sup>	1	0.70 (0.25 - 1.29)	0.97 (0.37 - 2.54)			0.973
OR controllato per Fumo attivo <sup>†</sup> , Fumo passivo <sup>§</sup>	1	0.64 (0.22 - 1.84)	1.02 (0.41 - 2.54)			0.947
OR controllato per Fumo attivo <sup>†</sup> , Fumo passivo <sup>§</sup> , Centr. el. <sup>  </sup> , Distr. carb. <sup>¶</sup>	1	0.60 (0.21 - 1.75)	1.01 (0.38 - 2.72)			0.972

\* Log likelihood ratio test. <sup>†</sup> Fumo attivo: status di fumatore o ex-fumatore vs. non fumatore alla fine del periodo di esposizione. <sup>§</sup> Fumo passivo: frazione del periodo di esposizione trascorsa con coabitanti fumatori. <sup>¶</sup> Frazione del periodo di esposizione trascorsa in abitazioni a meno di 200 m da distributori di carburante. <sup>||</sup> Frazione del periodo di esposizione trascorsa in abitazioni a meno di 200 m da centrali elettriche.

## 8. OSSERVAZIONI SUI RISULTATI

### 8.1 Il diverso rischio residenziale tra donne e uomini

L'analisi mostra un diverso comportamento tra donne e uomini: mentre il contingente femminile mostra un eccesso di rischio di mortalità (sebbene non significativo nella gran parte delle analisi), non è così per il contingente maschile.

Il fenomeno può essere attribuito alla diversa esposizione che l'indicatore residenziale determina in ragione dello stato occupazionale: si osserva infatti una pronunciata diversa frequenza di coloro che hanno trascorso parte del periodo di esposizione come casalinga/pensionato nei due sessi (Tabella 11). Le donne che hanno vissuto almeno 5 anni (1/3 del periodo di esposizione) come casalinga/pensionata sono 36% contro 20% degli uomini; le donne che hanno vissuto almeno 10 anni (2/3 del periodo di esposizione) come casalinga/pensionata sono 30% contro 13% degli uomini. E' quindi plausibile che il diverso comportamento dell'indicatore residenziale nelle donne rispetto agli uomini dipenda dal fatto che mediamente le donne, in ragione del loro stato occupazionale, abbiamo trascorso più tempo a casa di quanto sia avvenuto negli uomini.

**Tabella 11. Numero di anni, all'interno del periodo di esposizione, vissuti come pensionato o casalinga**

Anni	Maschi	Femmine	Totale
0	103	80	183
2	2	0	2
3	4	0	4
5	2	1	3
6	1	1	2
7	1	1	2
8	4	2	6
9	1	2	3
10	1	3	4
11	2	0	2
12	4	5	9
13	1	0	1
14	1	1	2
15	9	29	38
Totale	136	125	261

Restringendo l'analisi ai 134 soggetti, donne e uomini insieme, che hanno trascorso almeno 5 anni come pensionati o casalinghe o non occupati (53 *casì* e 81 *controlli* appaiati), gli OR grezzi per terzile di esposizione residenziale cumulata sono: 1, 1.18 (IC 95%: 0.48 – 2.86) e 1.44 (IC 95%: 0.61 – 3.41) (p trend = 0.405). Controllando per fumo attivo, fumo passivo, vicinanza a

distributori di carburante e vicinanza a centrali elettriche gli OR divengono: 1, 1.20 (IC 95%: 0.47 – 3.11) e 1.67 (IC 95%: 0.65 – 4.29) (p trend = 0.276). Restringendo ulteriormente l'analisi ai soggetti con almeno 10 anni trascorsi come pensionati o casalinghe o non occupati (37 casi e 51 controlli appaiati), gli OR grezzi sono 1, 1.15 (IC 95%: 0.41 – 3.26) e 1.99 (IC 95%: 0.71 – 5.59) (p trend = 0.187); gli OR controllati per fumo attivo, fumo passivo, vicinanza a distributori di carburante e vicinanza a centrali elettriche sono: 1, 1.05 (IC 95%: 0.32 – 3.44) e 2.96 (IC 95%: 0.79 – 11.1) (p trend = 0.098), raggiungendo così un livello prossimo alla significatività.

L'analisi negli uomini pensionati e non occupati suggerisce che anche nel contingente maschile, disponendo di informazioni sulla permanenza presso le abitazioni, si evidenziano condizioni di esposizioni a rischio, sebbene non significative (per gli uomini pensionati/non occupati per almeno 5 anni risulta un OR pari a 1.23 [IC 95%: 0.29 – 5.21] confrontando il III terzile vs. il I terzile di esposizione residenziale).

## **8.2 Il rischio residenziale per le donne**

La stima dell'OR di esposizione nelle donne risulta statisticamente significativa nei modelli controllati per la variabile che contraddistingue i soggetti che hanno avuto almeno un'abitazione a meno di 200 metri da un traliccio (informazione autoriportata e non misurata). Questa variabile è di difficile interpretazione e richiede qualche commento. (a) Lo scopo di accertare se nelle vicinanze dell'abitazione ci fossero tralicci è quello di controllare un possibile confondimento da vicinanza a linee elettriche. Tuttavia, soggetti con abitazioni lontane da un traliccio ma vicine ad una linea elettrica possono essere erroneamente classificati come non esposti a linee elettriche. Quindi la variabile 'vicinanza al traliccio' può essere scarsamente correlata all'effettiva esposizione a linee elettriche. (b) Il fatto che tale variabile risulti negativamente associata allo status caso-controllo e positivamente correlata al punteggio di esposizione residenziale cumulata  $r = 0.281$  ( $p = 0.014$ ) nei controlli, suggerisce che possa essere una proxy di qualche altra variabile o di un gruppo di altre variabili. Analisi per approfondire il significato di questa variabile e la sua interpretazione sono tuttora in corso. Inoltre si è ritenuto opportuno acquisire la distanza tra le abitazioni e la più vicina linea elettrica (tramite georeferenziazione), che fornisce una più corretta classificazione di una eventuale esposizione a campi magnetici prodotti dalle linee elettriche. L'analisi di questa variabile, da poco resa disponibile, è ancora in corso.

### **8.3 La misura del rischio residenziale**

L'attribuzione di un punteggio di esposizione residenziale dovute alla prossimità alla Raffineria ha tenuto conto dei tempi vissuti nelle abitazioni e delle loro distanze (georeferenziate) dalla Raffineria, ma non della direzione dei venti prevalenti nell'area in studio come inizialmente previsto dal programma di analisi statistica. Questa informazione è in corso di elaborazione da parte del Dipartimento di Energetica dell'Università Politecnica delle Marche, unitamente all'elaborazione di un punteggio di esposizione ad altre sorgenti emissive distinte dalla Raffineria.

### **8.4 La significatività statistica**

Nella presentazione dei risultati si è fatto riferimento all'usuale convenzione interpretativa del concetto di significatività statistica. Essa ci ha portato a considerare come statisticamente significativi solo quei risultati che permettano di rifiutare l'ipotesi di non-associazione tra esposizione residenziale e mortalità con un errore inferiore al 5% avendo come ipotesi alternativa sia la condizione di un eccesso che, all'opposto, di una riduzione di mortalità. Si tratta di una scelta di presentazione molto conservativa, poiché in ragione della nota associazione tra esposizione a benzolo e altri idrocarburi aromatici e rischio di tumore del sistema emolinfopoietico, l'Indagine è stata avviata per saggiare, la sola ipotesi di eccesso di mortalità. In termini operativi, ciò significa che se nella presentazione si fosse adottato un criterio *non conservativo*, avremmo riscontrato che l'analisi eseguita sulle persone casalinghe o pensionate o non occupate (uomini e donne) avrebbe determinato un significativo rischio di eccesso di morte.

## 9. Conclusioni e prospettive

La presente Indagine relativa alle popolazioni dei comuni di Falconara Marittima, Chiaravalle e Montemarciano, circostanti la raffineria API, ha evidenziato nel periodo 1994-2003 un aumento non statisticamente significativo della mortalità per tumore del sistema emolinfopoietico (ICD-9: 200-208) per quanto concerne il contingente femminile.

In considerazione della non ampia frazione di persone che hanno aderito, tra i parenti dei casi o dei controlli casualmente estratti (circa il 50% di quelli inizialmente individuati), l'Indagine, mentre suggerisce dal punto di vista delle tendenze un fondato sospetto che nel periodo considerato vi sia stato un aumento della mortalità attesa nelle aree circostanti la Raffineria, non può essere considerata conclusiva dal punto di vista quantitativo perché la mancata adesione può aver determinato una sovra-stima del rischio di morte. Il fenomeno ha coinvolto sistematicamente il contingente femminile, il sottogruppo dei soggetti interessati a possibile esposizione residenziale appartenenti al contingente maschile e globalmente, con un dato prossimo alla convenzionale significatività statistica - significativo in analisi che avessero adottato criteri meno conservativi, l'insieme delle persone interessate a possibile lunga esposizione residenziale (donne e uomini casalinghe, pensionati e non occupati per più di 10 anni).

Il sospetto, relativamente ad una ipotesi di eccesso di rischio di mortalità nelle donne, è supportato da una serie di evidenze tutte nella direzione di tale eccesso, ma tutte non in grado di raggiungere la convenzionale significatività statistica, tranne una: in un modello d'analisi, tra i diversi adottati per produrre stime del rischio depurate da possibili fattori di confondimento (altri fattori concomitanti, che associati al rischio di morte possono erroneamente far emergere o nascondere il rischio residenziale), è stato osservato un aumento statisticamente significativo di mortalità, ciò includendo però nel modello una informazione di dubbia interpretazione che richiede ulteriori approfondimenti, incompatibili con i limiti temporali e di mandato assegnati alla presente l'Indagine.

Nell'interesse generale della salute delle popolazioni dei comuni indagati, l'Indagine è stata decisamente utile per supportare una chiara ipotesi di rischio, ma per rafforzare i conseguenti interventi di protezione ambientale e di sanità pubblica (interesse primario dello scrivente), è necessario un supplemento di Indagine con un nuovo mandato che affini l'investigazione con elementi precedentemente convenientemente non considerati. Lo scrivente Istituto Nazionale dei Tumori, a tale proposito, si dichiara disponibile ad accertare le modalità contrattuali per un nuovo mandato.

Si suggerisce che il nuovo mandato, oltre ad investigare

- a. sul significato di quella informazione sopracitata di dubbia interpretazione, possa ulteriormente investigare:
- b. sul rischio residenziale, considerando le informazioni recentemente rese disponibili da ARPAM in indagini ad-hoc e anche dati, ad oggi non ancora disponibili, relativi alle modalità di ricaduta degli inquinanti (immissioni) che tengano conto della direzione prevalente dei venti, della morfologia del territorio esposto e delle variazioni climatiche, così da individuare con più precisione l'area interessata dal sospetto eccesso di rischio;
- c. sulla evoluzione del rischio nel periodo, al fine di accertare se il sospetto eccesso di mortalità sia attribuibile ad eventi ancora in corso o sia attribuibile ad un limitato o ad un passato periodo di tempo;
- d. sul possibile eccesso di rischio in relazione alle diverse forme tumorali interessate dall'Indagine dovuto alle caratteristiche biomolecolari. Si tratta di una proposta estremamente innovativa, basata sul presupposto ampiamente dimostrato in altri sistemi (carcinomi epatocellulari e esposizione a aflatossine e spettro mutazionale del gene TP53), che è in grado di evidenziare la diretta correlazione causa /effetto tra esposizione e mutazione. La proposta richiede l'applicazione di approcci "genome wide" su materiale tumorale conservato in paraffina in *casi* esposti e non-esposti; analisi a costi elevati per la quale, data la rilevanza scientifica, sarebbe possibile ottenere ulteriori supporti da soggetti pubblici e/o privati;
- e. in modo chiarificatore, sull'intero insieme dei *casi* e dei *controlli* considerati, superando tutte le barriere di riservatezza e tutela della privacy, che hanno contribuito a limitare la partecipazione dei cittadini all'Indagine. Nell'interesse più alto della salute pubblica, ora evidenziato dai risultati della presente Indagine e confermando i doveri di riservatezza su dati individuali per chi opera in sanità, si raccomanda che il nuovo mandato indichi la possibilità di (1) ricostruire la storia residenziale dei soggetti in studio attraverso i dati disponibili alle anagrafi e non solo attraverso le informazioni raccolte mediante questionari, (2) acquisire tali informazioni anche su coloro per i quali i parenti hanno rifiutato l'intervista.

In conclusione, il presente Primo Rapporto rende disponibili informazioni sulle attività svolte e sui risultati della complessa indagine epidemiologica condotta sulla popolazione residente a Falconara Marittima e comuni limitrofi in proficua collaborazione con i tecnici locali e grazie

alla sperimentata professionalità degli operatori del Servizio di Epidemiologia Ambientale del dipartimento di Ancona dell'Agenzia Regionale Protezione Ambientale Marche (ARPAM). L'Indagine è stata opportunamente promossa a seguito della richiesta di cittadini e per il consenso delle amministrazioni locali e della Regione Marche, che ne ha anche supportato lo sviluppo. In relazione, a precedenti accordi scritti e verbali lo scrivente si riserva di recapitare un ulteriore Rapporto Finale entro e non oltre il 31 gennaio 2009, che consideri anche elementi che dovessero scaturire nel frattempo. Entro tale data e all'interno del presente mandato, lo scrivente si rende disponibile ad un incontro con i committenti, nella modalità che Essi riterranno opportuno, e ad una presentazione pubblica relativamente all'Indagine qui riportata.

## Referenze

1. Baili P, Mariottini M, Meneghini E, Micheli A. Studio di fattibilità per l'avvio di una indagine epidemiologica sulla popolazione residente presso la raffineria API di Falconara Marittima. *Epidemiol Prev.* 31(1 Suppl 2):48-53. 2007
2. World Health Organisation. *International classification of diseases, 1975, 9th edn.* Geneva: WHO, 1977
3. Clayton D, Hills M, eds. *Statistical Models in Epidemiology.* Oxford: Oxford University Press, 1998
4. Yamane T, ed. *Elementary Sampling Theory.* Englewood Cliffs (NJ): Prentice-Hall Inc., 1967
5. Kleinbaum DG, Kupper LL, Morgenstern H, eds. *Epidemiologic Research: Principles and Quantitative Methods.* New York: Van Nostrand Reinhold Company Inc., 1982
6. DoRS, Documentazione Regionale Salute: <http://www.dors.it>
7. Progetto OCCAM - Occupational Cancer Monitoring - Sistema informativo sui tumori di origine professionale: <http://www.occam.it>
8. Peduzzi PN, Concato J, Kemper E, Holford TR, Feinstein A. A simulation study of the number of events per variable in logistic regression analysis. *Journal of Clinical Epidemiology*, 99: 1373-9. 1996
9. Schlesselman J. *Case control studies: design, conduct, analysis.* Oxford: Oxford University Press, 1982

**Annesso 1.**

**PROTOCOLLO RELATIVO  
ALL'INDAGINE EPIDEMIOLOGICA PRESSO  
LA POPOLAZIONE RESIDENTE A  
FALCONARA MARITTIMA E COMUNI LIMITROFI**

A cura di:

Dr. Andrea Micheli

Direttore

Struttura Complessa Epidemiologia Descrittiva e Programmazione Sanitaria

Collaborazione:

Dr. Paolo Baili



Fondazione IRCCS "Istituto Nazionale dei Tumori"

Via Venezian, 1

20133 Milano MI

Tel. 02 2390 2867 - 2869

Fax 02 2668 1292

e\_mail: [micheli@istitutotumori.mi.it](mailto:micheli@istitutotumori.mi.it)

## Indice

1. PREMESSA	30
2. INTRODUZIONE	0
3. DISEGNO DELLO STUDIO	32
3.1 Premessa: la scelta dello studio	32
3.2 Obiettivi	33
3.3 Metodo	33
3.4 L'area in studio	33
3.5 Il periodo di studio	33
3.6 Il tipo di dati	33
3.7 I Soggetti Casi	33
3.8 I Soggetti Controlli	33
3.9 I Soggetti in studio	34
3.10 Altre informazioni: il questionario	34
3.11 Misura di esposizione	34
3.12 Analisi	34
3.13 Varie fasi della raccolta dei dati	35
4. TEMPI E PERSONALE DELL'INDAGINE	15
5. RUOLO INT	15

## 1. PREMESSA

Nel luglio 2003 è stata stipulata tra l'Istituto Nazionale dei Tumori per lo Studio e la Cura dei Tumori (INT) e Agenzia Regionale Sanitaria della regione Marche (ARS) una convenzione per uno “Studio di fattibilità per un'indagine epidemiologica presso la Raffineria API di Falconara”.

Nel marzo 2004 lo studio è arrivato alla conclusione che un'indagine epidemiologica è fattibile.

Per saggiare un eventuale eccesso di rischio nelle popolazioni esposte agli inquinanti emessi dalla raffineria API (da qui, Raffineria) nella Regione Marche sono infatti presenti le condizioni:

- d'evidenza per intraprendere uno studio analitico che coinvolga tutti i casi di decesso dell'ultimo decennio per tumori del sistema emolinfopoietico occorsi nel comune di Falconara Marittima (a cui verranno aggiunti il comune di Montemarciano e di Chiaravalle)
- organizzative, con l'avvertenza però, che si attivino da subito le procedure perché la popolazione e i tecnici a vario titolo interessati siano continuamente coinvolti secondo il metodo del consenso informato e della partecipazione attiva.

Lo studio di fattibilità ha inoltre proposto che l'indagine epidemiologica fosse attuata con la tecnica casi-controlli di popolazione e fosse indirizzata alla stima dei rischi di morte per i tumori del sistema emolinfopoietico, in funzione della distanza dalla Raffineria e in relazione anche ai livelli d'esposizione (dipendenti dai venti, dalle condizioni climatiche ecc) a inquinanti atmosferici come il benzolo (di cui è noto il ruolo eziologico per le leucemie) e ad altri composti.

A seguito dello studio di fattibilità, la regione Marche ha promulgato una delibera (n° 218 del 15 giugno 2004) in cui venivano approvati i risultati raggiunti e adottati i provvedimenti necessari per l'effettuazione di un'indagine epidemiologica presso la popolazione di Falconara Marittima.

Il presente rapporto introduce al protocollo per l'attuazione dell'indagine epidemiologica (da qui Indagine) presentando gli elementi utili al suo avvio. La Fondazione IRCCS “Istituto Nazionale dei Tumori” (INT) attraverso la Struttura Complessa di Epidemiologia Descrittiva e Programmazione Sanitaria assumerà la direzione scientifica dell'Indagine.

## **2. INTRODUZIONE**

L'azienda API è una raffineria (Raffineria) di prodotti petroliferi sita dal 1950 nel comune di Falconara Marittima (Falconara) in provincia di Ancona che per la natura della produzione svolta determina esposizione a sostanze irritanti, tossiche, nocive e indicate o sospettate di potere cancerogeno.

La Raffineria dista alcune centinaia di metri dai quartieri Fiumesino, Villanova, (aree Falconara Acquedotto, Scuola e Antonelli) di Falconara, ma altri quartieri della città possono essere interessati a problemi d'esposizione per la possibile ricaduta di sostanze inquinanti emesse dall'azienda. Falconara è oggetto di monitoraggio per la valutazione dei rischi atmosferici connessi alla presenza della Raffineria.

Il rischio legato alla presenza della Raffineria è riferibile sia alla possibile esposizione di tipo acuto, sia a rischi d'esposizione di tipo cronico a benzolo e altri composti di riconosciuta azione oncologica. Gli effetti associati a sostanze cancerogene tendono spesso a comparire dopo un lungo periodo di tempo (sino a 30-40 anni) dall'esposizione rendendo difficile lo studio della relazione tra esposizione e danno.

I dati di mortalità analizzati nello studio di fattibilità erano relativi al ventennio 1980-2000, anni in cui dovrebbero indicativamente manifestarsi gli effetti d'esposizione del periodo 1950-1970, il periodo in cui la Raffineria ha vissuto la fase di sviluppo.

A Falconara i tassi di mortalità per leucemia e per tumori del sistema emolinfopoietico sono in aumento nell'ultimo ventennio. Un tempo a Falconara i tassi erano su livelli inferiori a quelli della Provincia di Ancona, mentre ora sono su livelli generalmente superiori. D'altra parte, in altre aree della regione si è osservato un fenomeno di incremento del rischio simile a quello osservato a Falconara, di cui sarebbe importante interpretarne le cause. A Falconara, la mortalità per malattie del sistema respiratorio non diminuisce nel corso del tempo, come invece accade, spesso significativamente, nella maggior parte degli altri comuni marchigiani.

Di seguito sono indicate gli elementi principali da includere nel protocollo di studio.

### **3. DISEGNO DELLO STUDIO**

#### **3.1 Premessa: la scelta dello studio**

Per accertare la relazione tra esposizione a benzolo e altri inquinanti (da qui, *esposizione*) dovuta alle emissioni dagli impianti API (da qui, *fattori di rischio in studio*) e le tendenze di mortalità individuate dallo studio di fattibilità è necessario avviare un formale studio epidemiologico (Indagine). La relazione tra esposizione e decessi può essere studiata mediante studi di mortalità (da qui, *studio di mortalità*). La relazione tra esposizione e patologie può essere anche studiata mediante studi d'incidenza, cioè relativi all'occorrenza di nuovi casi nella popolazione. Gli studi d'incidenza su base di popolazione si possono eseguire laddove è presente un Registro Tumori di Popolazione (RTP). In campo oncologico, l'esposizione ad un fattore di rischio è associata direttamente all'incidenza della malattia e solo indirettamente alla mortalità. La mortalità è associata indirettamente al fattore di rischio, dipendendo sia dall'occorrenza che dalla probabilità di sopravvivenza alla malattia in studio. Nel caso delle leucemie o di malattie oncologiche del sistema emolinfopoietico nei decenni scorsi la mortalità poteva essere utilizzata come proxy dell'occorrenza della malattia; dagli anni '90, con la diffusione di terapie efficaci per il trattamento della leucemia infantile, l'incidenza è il miglior parametro epidemiologico per la valutazione della presenza di fattori di rischio per le leucemie infantili. Per le leucemie negli adulti il dato di mortalità può ancora essere utilizzato, ma richiede cautela e prudenza.

La probabilità di morte per leucemia dipenderà infatti: (a) dall'esposizione – causale o concausale - a fattori di rischio (endogeni e esogeni, noti e non noti), (b) dal pattern genetico, (c) dall'interazione tra esposizione e pattern genetico, ma anche, (d) dalla probabilità di sopravvivenza a altre cause di morte rispetto a quella in studio (un caso di leucemia, eventualmente associato a esposizione ai fattori, potrebbe ad esempio morire per incidente stradale e non essere quindi registrato come Caso), e (e) dalla probabilità specifica di sopravvivenza (essa dipende dalla curabilità – la probabilità di essere curati - della malattia).

Assumendo che la probabilità di perdere casi-decesso per i decessi dovuti ad altre cause rispetto a quella in studio, sia simile nella popolazione dell'area esposta che nelle popolazioni dell'area di riferimento, e che la curabilità sia altrettanto simile nelle popolazioni a confronto, l'uso delle statistiche di mortalità potrebbe comportare nel tempo una progressiva riduzione della potenza dello studio (lo stesso fenomeno potrebbe cioè determinare una differenza più grande tra casi incidenti attesi e osservati e una differenza più piccola tra casi decesso attesi e osservati).

D'altra parte:

- a) lo studio dei casi di decesso è più facile da eseguire perché le statistiche di mortalità sono facilmente disponibili,
- b) lo studio d'incidenza è praticamente impossibile per la popolazione adulta per quanto riguarda il passato, ed è alto rischio d'errore anche per gli anni più recenti nelle aree dove non è presente un Registro Tumori di Popolazione (il Registro Tumori di Popolazione nelle Marche copre la provincia di Macerata e non copre l'area in studio).

Con le cautele e precauzioni indicate l'Indagine sarà indirizzata allo studio dei rischi di morte per tumori leucemia e del sistema emolinfopoietico.

### **3.2 Obiettivi**

L'Indagine è quindi indirizzata alla valutazione dei rischi di morte per tumore del sistema emolinfopoietico in funzione della distanza dalla Raffineria, considerando anche i livelli d'esposizione a inquinanti atmosferici nelle diverse zone analizzate (dipendenti dai venti, dalle condizioni climatiche ecc) e altre variabili di aggiustamento che verranno acquisite nel corso dell'Indagine.

### **3.3 Metodo**

L'Indagine sarà attuata con la tecnica dello studio casi-controlli su base di popolazione.

### **3.4 L'area in studio**

L'*area in studio* comprende i comuni di Falconara Marittima, Montemarciano, Chiaravalle: essa includerà l'area direttamente esposta come definita nell'introduzione (saranno definite categorie d'esposizione) e l'area di riferimento definita come un'area esterna, confinante e non direttamente interessata all'immissione d'inquinanti dovuti alla Raffineria.

### **3.5 Il periodo di studio**

Dal punto di vista epidemiologico, tanto più grande è l'arco temporale di studio (espresso in questo caso in anni d'osservazione) tanto maggiore sarà il numero di *eventi* che si sarà accumulato nel periodo *d'esposizione*, e tanto più grande sarà quindi la probabilità di mettere in luce anche un piccolo ecceso di rischio dovuto *all'esposizione*, se esso esiste. L'arco temporale ottimale dovrebbe comprendere gli anni dall'installazione della Raffineria (*anno d'inizio dell'esposizione*), all'anno più prossimo ai giorni nostri per il quale saranno disponibili dati di frequenza degli *eventi* della/e malattia/e d'interesse. Dal punto di visto operativo, nell'arco temporale così definito, vi sono invece anni per i quali i dati sono più facilmente disponibili. La qualità dell'informazione, i tempi, e i costi dello *studio* dipendono dalla disponibilità di dati sugli *eventi*.

In considerazione degli aspetti operativi sopraindicati, si è deciso di analizzare i decessi insorti dal 1/1/1994 al 31/12/2003 in quanto le informazioni necessarie dei decessi dal 1980 al 1993 sono di difficile reperibilità e frammentarie.

Il periodo in studio includerà quindi gli anni dal 1994 al 2003 (*periodo di studio*).

### **3.6 Il tipo di dati**

La mortalità può essere espressa per presenza o residenza. Per convenzione, le statistiche di mortalità di tipo geografico in Italia sono usualmente espresse per residenza. In riferimento a tale convenzione nell'Indagine le popolazioni utilizzate sono le popolazioni residenti in un'area specifica e in un tempo specifico e i decessi sono i decessi occorsi in quelle popolazioni residenti.

### **3.7 I Soggetti Casi**

Primo obiettivo dell'Indagine sarà l'individuazione dei Casi occorsi nell'*area in studio* e nel *periodo in studio*. Sarà quindi costituita la lista dei Casi. Per ogni Caso oltre alle informazioni sulla scheda di morte deve essere noto il sesso, la data di nascita, il numero di anni di residenza (periodo di residenza) nell'*area in studio* e l'anno (mese e giorno) in cui il soggetto in questione è diventato Caso. Si tratterà poi di acquisire, per quanto possibile, la documentazione clinica della persona diventata Caso.

### **3.8 I Soggetti Controlli**

Dai Comuni dell'area in studio e per il periodo di studio, verranno acquisite le liste anagrafiche di tutte le persone residenti al tempo di sopravvenienza dei Casi. Per ogni Caso verranno individuati dalle liste anagrafiche dopo appaiamento per a) anno di calendario della comparsa del Caso, b)

sesto, c) età, d) periodo di residenza e metodo casuale (random), 2 soggetti non-Caso al momento dell'appaiamento che parteciperanno allo studio come Controlli.

### **3.9 I Soggetti in studio**

L'insieme dei soggetti Casi e dei soggetti Controlli costituirà il gruppo di *soggetti in studio*. La banca dati relativa ai *soggetti in studio* sarà accessibile solo ai responsabili dell'Indagine. Per motivi di riservatezza, la parte anagrafica individuale sarà nota solo agli analisti e verrà per quanto possibile tenuta in modo riservato. Ad ogni soggetto appartenente alla lista dei soggetti in studio sarà attribuito in modo random un codice identificativo, e in ogni fase dell'Indagine le informazioni verranno associate al solo codice identificativo e non sarà più nota la condizione individuale di Caso o di Controllo.

### **3.10 Altre informazioni: il questionario**

Di ogni soggetto appartenente alla lista dei soggetti in studio deve essere reso disponibile un set di dati che comprendono:

- dati anagrafici reperibili dall'anagrafe dei comuni coinvolti
- storia residenziale reperibile sempre dall'anagrafe dei comuni. Per ogni soggetto bisogna registrare tutti gli spostamenti residenziali dalla nascita (anche se non è avvenuta nel comune). Se tale raccolta non è possibile bisogna almeno registrare la storia residenziale del soggetto dal momento in cui è apparso nell'anagrafe del comune per la prima volta
- storia lavorativa reperibile tramite questionario da somministrare ad un parente del deceduto (si deve registrare il grado di parentela) o in altro modo.

Il questionario avrà l'obiettivo a) di verificare l'eventuale esposizione ambientale e/o professionale attuale o pregressa a fattori di rischio associati a leucemia, b) di raccogliere eventuali informazioni mediche relative a condizioni patologiche precedenti l'arrivo nella zona a rischio.

### **3.11 Misura di esposizione**

La raccolta di informazioni sulla storia residenziale e lavorativa nel tempo di ogni soggetto coinvolto nell'Indagine deve tradursi in misura di esposizione e in variabili di aggiustamento incluse nei modelli di analisi. Saranno inoltre acquisiti dati relativi ai livelli d'inquinamento ambientale nelle zone d'immissione degli inquinanti.

### **3.12 Analisi**

L'analisi sarà condotta da soggetti che non hanno partecipato direttamente alla raccolta dei dati. Lo studio dell'associazione sarà condotto mediante analisi logistica e altre e diversificati approcci e metodiche, controllando gli effetti di confondimento delle diverse covariate che verranno incluse nei modelli.

### **3.13 Varie fasi della raccolta dei dati**

	<b>FASE</b>	<b>UFFICI</b>	<b>RESPONSABILE</b>	<b>TEMPISTICA</b> *	<b>NOTA</b>
1)	Liste popolazioni residenti	Anagrafi	Collaboratore analista	Entro 4° mese	Presso l'analista deve essere creato un archivio e un database in cui inserire di volta in volta i dati raccolti
1)a	Invio lettera richiesta dati ai Sindaci	Anagrafi	Resp. SEA	15/03/2005	
1)b	Sollecito invio dati e richiesta referente	Anagrafi	Resp. SEA	25/03/2005	
1)c	Esame dati disponibili e creazione archivio e database	Anagrafi	Analista	10/04/2005	
1)d	Raccolta ed inserimento dati nel database	Anagrafi	Analista e collaboratore	10/07/2005	
2)	Raccolta certificati decesso per ogni caso (decessi per tumori del sistema emolinfopoietico 1994-2003)	Zona 7 ASUR	Collaboratore analista	Entro 4° mese	Ogni certificato e ogni altro documento va fotocopiato e catalogato in un raccoglitore per ogni soggetto
2)a	Invio lettera richiesta dati alla zona 7 ASUR	Zona 7 ASUR	Resp. SEA	15/03/2005	
2)b	Sollecito invio dati e richiesta referente	Zona 7 ASUR	Resp. SEA	25/03/2005	
2)c	Consultazione diretta denunce cause di morte	Zona 7 ASUR	Collaboratore analista	10/04/2005	Ricerca delle cause oggetto dello studio in tutti i tre campi delle cause di morte
2)d	Esame dati disponibili e creazione archivio e database	Zona 7 ASUR	Analista	30/04/2005	
2)e	Raccolta ed inserimento dati nel database		Analista e collaboratore	30/07/2005	
3)	Controllo tra dati ISTAT e lista casi		INT	15/08/2005	
4)	Lista decessi per età, genere, anno di decesso	Anagrafi	Collaboratore analista	Entro 6° mese	
4)a	Preparazione liste casi per anno, età e sesso		Analista	Entro 6° mese	

4)b	Preparazione liste popolazioni per anno, età e sesso		Analista	Entro 6° mese	
5)	Appaiamento (per ogni caso 2 controlli)	-	Analista, INT	Entro 6° mese	Al termine ha luogo l'assegnamento dei codici casuali
6)	Raccolta dati dalle cartelle cliniche	AUSL, Ospedali	Collaboratore analista	Entro 14° mese	Dalle cartelle cliniche si deve trarre la diagnosi e la data di diagnosi dei casi
6)a	Preparazione ed invio lettera per accesso alle cartelle cliniche	AUSL, Ospedali	Resp. SEA, INT	30/09/2005	
6)b	Preparazione archivio e database per dati clinici		Resp. SEA, Analista	30/10/2005	
6)c	Accesso agli archivi ospedalieri e raccolta informazioni cliniche		Collaboratore analista	Entro 14° mese	
7)	Preparazione e diffusione "interim report"		Resp. SEA, INT	Al 12° mese	
8)	Preparazione lettera ai parenti	-	Resp. SEA, INT	Entro 6° mese	La lettera deve essere firmata dall'assessore e spiegare i motivi dell'intervista e dell'Indagine
9)	Ricerca parenti da contattare (anche al di fuori del Comune di Falconara)	Anagrafi	Collaboratore analista	Entro 8° mese	Bisogna individuare sia indirizzo che numero di telefono
10)	Accordo con sedi INPS, INPDAP e INAIL	INPS-INAIL	Assessore	Entro 7° mese	Accordo per ricostruire la storia lavorativa dei soggetti nei casi di insufficienza dei dati raccolti con i questionari
11)	Preparazione questionario	-	Analista, INT	Entro 10° mese	Il questionario trarrà spunto da questionari già utilizzati. Conterrà quesiti per ricostruire la storia residenziale, occupazionale, sanitaria e esposizionale in genere.
11)a	Esame di questionari già validati e utilizzati dall'INT		Analista, INT	Entro 8° mese	
11)b	Apportare le modifiche riferite alla realtà locale e alle necessità informative dello studio		Analista, INT	Entro 8° mese	
11)c	Sperimentazione del questionario su un campione ridotto di persone per valutarne la somministrabilità (tempi, oggettività e		Analista	Entro 9° mese	

	comprensibilità) dei quesiti				
11)d	Apportare le eventuali correzioni ed aggiustamenti al questionario		Analista, INT	Entro 10° mese	
12)	Scelta e istruzione degli intervistatori	-	Analista, INT	Entro 12° mese	Durante la validazione è importante individuare quali possono essere i documenti d'appoggio fondamentali per la corretta risposta delle domande
12)a	Definizione caratteristiche ed individuazione degli intervistatori		Resp. SEA, INT	Entro 10° mese	Gli intervistatori potranno essere scelti o tra il personale interno all'ARPAM o tra esterni. Sarà verificata la possibilità e l'utilità di utilizzare personale medico igienista data la delicatezza e la riservatezza dei dati da raccogliere.
12)b	Addestramento degli intervistatori		Analista, INT	Entro 11° mese	
12)c	Simulazione di alcune interviste e valutazione dei risultati		Analista, intervistatori	Entro 12° mese	La simulazione dovrà verificare la confrontabilità di interviste effettuate dai due intervistatori alle medesime persone e la capacità di raccogliere informazioni affidabili.
13)	Invio lettere di partecipazione all'Indagine ai <b>parenti</b> da intervistare	Assessorato	Assessore	Entro 12° mese	E' consigliabile mandare anche bozza del questionario in modo tale che gli intervistati non siano all'oscuro delle domande a cui dovranno rispondere indicando anche quali sono i documenti utili alle risposte
14)	Interviste	-	Intervistatore	Entro 18° mese	Le interviste potranno durare anche fino a due ore. E' perciò importante fare decidere l'ora all'intervistato
14)a	Programmare appuntamenti con i parenti con cadenza settimanale		Intervistatore	Entro 18° mese	
14)b	Effettuare almeno 12 interviste a settimana		Intervistatore	Entro 18° mese	
14)c	Inserire i dati nel PC		Intervistatore	Entro 18° mese	

15)	Ricostruzione storia residenziale ed occupazionale per ogni soggetto (caso e controllo)			Entro 20° mese	
15)a	Ricostruzione storia residenziale dei soggetti		Collaboratore analista	Entro 19° mese	Storia residenziale dalla nascita o almeno dal momento in cui il soggetto è apparso nell'anagrafe del comune di Falconara
15)b	Ricostruzione storia lavorativa dei soggetti		Collaboratore analista	Entro 19° mese	Storia lavorativa dalla nascita o almeno dal momento in cui il soggetto è apparso nell'anagrafe del comune di Falconara
15)c	preparare l'archivio ed il database delle storie residenziali e lavorative		Analista, INT	Entro 19° mese	
15)d	approfondire il tema mansioni e pesi da attribuire alle singole esposizioni a rischio		Analista, INT	Entro 19° mese	
15)e	Informatizzazione dei dati residenziali ed occupazionali con le relative codifiche e pesi		Collaboratore analista	Entro 19° mese	
15)f	Ricostruzione storia residenziale in caso di carenza informazioni dal questionario	Anagrafi	Collaboratore analista	Entro 20° mese	Storia residenziale dalla nascita o almeno dal momento in cui il soggetto è apparso nell'anagrafe del comune
15)g	Ricostruzione storia lavorativa dei soggetti in caso di carenza informazioni dal questionario	INPS- INPDAP -INAIL	Collaboratore analista	Entro 20° mese	Storia lavorativa dalla nascita o almeno dal momento in cui il soggetto è apparso nell'anagrafe del comune di Falconara
16)	Misurazione distanze da Raffineria		Analista e collaboratore	Entro 21° mese	Tali misurazioni vanno fatte per ogni luogo (abitazione, posto di lavoro) identificato nella città di Falconara
16)a	Acquisizione toponomastiche dei comuni e supporto tecnico		Analista e collaboratore	Entro 21° mese	
16)b	Definizione e calcolo centroide o perimetro della raffineria		Resp. SEA, Analista, INT	Entro 21° mese	
16)c	Calcolo delle distanze		Collaboratore analista	Entro 21° mese	
17)	Stima ricaduta polveri e misure benzene e IPA		ARPAM	Entro 16° mese	Sarebbe opportuno fare una stima della ricaduta delle polveri in ognuna delle 4 stagioni

17)a	Modellistica delle ricadute delle polveri emesse dalla raffineria o da altre fonti		ARPAM	Entro 16° mese	
17)b	Reperimento di tutti i dati storici disponibili sulle polveri, IPA e benzene		ARPAM	Entro 16° mese	
17)c	Eventuali misure a campione per la verifica della modellistica		ARPAM	Entro 16° mese	
18)	Raccolta dati mancanti	-	Collaboratore analista	Entro 21° mese	Per la raccolta di dati mancanti le fonti possono essere: medico curante, parroco, ex vicini di casa....
19)	Analisi dati		INT, Resp. SEA, Analista	Entro 23° mese	
20)	Preparazione report finale e comunicazione risultati		INT, Resp. SEA, Analista	24° mese	

#### **4. TEMPI E PERSONALE DELL'INDAGINE**

Questo rapporto sarà discusso da tutte le parti interessate, utilizzando il metodo del consenso informato. Sulla base delle osservazioni e delle critiche e dei suggerimenti verrà steso il protocollo dell'Indagine. L'Indagine avrà inizio dopo l'approvazione del protocollo da parte del Comitato Etico dell'INT.

Il tempo previsto per l'Indagine è complessivamente di 24 mesi.

La Regione Marche eseguirà direttamente l'Indagine indicativamente attraverso:

- 1 analista dell'ARPAM al 30% per 2 anni
- 1 collaboratore al 50% per il primo anno per i rapporti con anagrafi e altri enti
- 2 intervistatori per 6 mesi
- 1 segretaria al 10% per la durata dell'Indagine
- 2 tecnici per un totale di 4 mesi per la stima dei livelli di esposizione.

#### **5. RUOLO INT**

La Struttura Complessa di Epidemiologia Descrittiva e Programmazione Sanitaria dell'INT si assume la direzione scientifica dell'indagine e si impegna a svolgere le seguenti attività:

- collaborazione nella stesura del protocollo finale
- consulenza scientifica e coordinamento delle attività svolte dagli enti coinvolti nell'Indagine
- presenza durante gli incontri con la popolazione e gli enti a vario titolo coinvolti.
- attività di analisi dei dati conclusivi dello studio;
- tutte le attività meglio descritte nell'allegato A alla presente convenzione;

Per tali attività viene stimato il seguente budget di spesa:

##### Personale

ricercatore (20 ore x 24 mesi)	8.000 Euro
attività di segreteria (3 ore x 24 mesi)	1.000 Euro
project leader (10 ore x 24 mesi)	6.200 Euro

##### Altre spese

spese per materiale	800 Euro
Spese di trasferta (Milano-Ancona)	2.000 Euro
Overhead	4.000 Euro
Totale	22.000 Euro

## Annesso 2. Relazione sull'appaiamento caso-controllo (1 marzo 2007)

### Introduzione

La procedura di appaiamento caso-controllo è stata scritta secondo i principi dello studio *case-control nested in the cohort*. Il campionamento è stato quindi eseguito nella 'popolazione a rischio' [1] e nel rispetto del presupposto di indipendenza dei campioni [2]. Per ogni caso, è stata individuata la popolazione a rischio di evento al momento dell'evento del caso o 'set a rischio' dalla quale estrarre casualmente 2 controlli. Un caso con evento successivo ad un altro caso appartiene al set a rischio di quest'ultimo ed è eligibile di essere estratto come suo controllo. Per conservare l'indipendenza dei campioni, l'estrazione in un set a rischio è stata eseguita con reintroduzione permettendo che un controllo possa essere estratto anche in altri set a rischio. Per aumentare l'efficienza delle stime statistiche, l'estrazione dei controlli è stata fatta stratificando per sesso e ogni controllo è stato appaiato al caso per età. Dato che in uno studio caso-controllo se un fattore di matching è associato all'esposizione, l'appaiamento per quel fattore introduce un bias e, nel caso in cui il fattore di matching sia associato anche alla malattia, non elimina la possibilità di confondimento [3], in fase di analisi può comunque essere necessario controllare per sesso ed età.

### Definizione dei soggetti in studio

Nel presente studio la popolazione residente nei tre comuni di Falconara Marittima, Chiaravalle e Montemarciano (da qui *area*) costituisce una coorte dinamica.

I casi disponibili per lo studio sono 177 (89 maschi e 88 femmine) soggetti deceduti nel periodo 1994 - 2003 per tumori del sistema linfemopoietico, residenti nell'*area*. Per ogni caso, si è considerato come tempo all'evento la data del suo decesso. La distribuzione per età all'evento differisce tra i due sessi, le femmine risultano mediamente più anziane dei maschi, media e deviazione standard risultano rispettivamente  $76.3 \pm 13.2$  anni e  $70.4 \pm 15.9$  anni ( 5% - 95% percentile: 54.7 – 91.6 anni per le femmine e 39.3 – 89.6 anni per i maschi).

Per ogni caso il set a rischio è stato considerato essere costituito dalla popolazione residente nell'*area* e in vita alla data del decesso del caso, avente stesso sesso ed età uguale all'età del caso  $\pm 2.5$  anni. Il set a rischio non include altri casi con evento nello stesso anno ma include i casi con evento negli anni successivi. Ai fini dell'appaiamento, l'età è stata calcolata come (data di morte - data di nascita)/365.25 per il caso, e come (data di morte del caso - data di nascita)/365.25 per il potenziale controllo ed è quindi una variabile continua. Dato che i casi molto anziani possono avere pochi potenziali controlli si è deciso di raggruppare in un'unica classe i soggetti con età  $\geq 85$  anni. Ad esempio, i soggetti del set a rischio di un caso maschio di 67 anni hanno età compresa tra 64.5 e 69.5 anni, quelli di un caso di 90 anni hanno età  $\geq 85$  anni. Da ogni set a rischio sono stati estratti 2 controlli con campionamento casuale semplice.

### La procedura di appaiamento

La procedura di appaiamento è stata suddivisa in due step.

Step 1. Viene prodotto un file per ogni caso contenente il caso stesso e il suo set a rischio. Ad ogni record del file la procedura associa un numero casuale appartenente all'intervallo (0, 1) ai potenziali controlli e associa 0 al caso; in questo modo ordinando il file in ordine crescente secondo la variabile *random* il caso è sempre il primo record.

Step 2. Nel secondo step da ogni file vengono estratti i primi 3 record in ordine crescente della variabile *random*, cioè il caso e i primi 2 controlli, e accorpati in 2 file distinti per sesso. Viene inoltre associata la variabile *set* che identifica la tripletta 1 CASO + 2 CONTROLLI (set = 1, 2...89 per i maschi e set = 90, 91...177 per le femmine).

Dato che per ogni caso esiste il file dei soggetti eleggibili come controlli ordinati in ordine crescente di *random*, è sempre possibile rimpiazzare un controllo estratto con quello successivo se dovesse essercene bisogno, nel caso di errori nei codici o di soggetti non rintracciabili o altro.

La procedura è stata scritta utilizzando STATA (versione 7.0, 2001, Stata Corporation, College Station, TX, USA).

### Cecità nella raccolta dei dati

Al fine di preservare la cecità sullo status caso/controllo da parte degli operatori che misureranno l'esposizione ed effettueranno le interviste ai parenti dei soggetti in studio, inizialmente l'INT fornirà all'ARPAM solo l'elenco dei codici dei soggetti risultanti dall'appaiamento caso-controllo secondo un ordine casuale, in modo che non sia possibile individuare l'appartenenza al set e lo status caso/controllo del soggetto.

### Esito dell'appaiamento caso-controllo

La tabella 2 mostra la distribuzione dello status caso/controllo per i maschi e per le femmine.

Tabella 2. Sesso e status caso/controllo dei soggetti risultanti dall'appaiamento.

	Maschi	Femmine	Totale
Caso	89	88	177
Controllo	178	176	354
Totale	267	264	531

### Bibliografia

- 1) Clayton D, Hills M, eds. Statistical Models in Epidemiology. Oxford: Oxford University Press, 1998.
- 2) Yamane T, ed. Elementary Sampling Theory. Englewood Cliffs (NJ): Prentice-Hall Inc., 1967.
- 3) Rothman KJ, Greenland Sander, eds. Modern Epidemiology. Philadelphia (PA): Lippincott-Raven Publishers, 1998.

### Annesso 3. Relazione sui soggetti non rispondenti (22 luglio 2008)

#### 1. Introduzione

Il questionario raccoglie informazioni sulla storia abitativa e su possibili strutture o attività a rischio nei pressi delle case effettivamente abitate indipendentemente dalla residenza dei soggetti in studio. L'esposizione cumulata è calcolata come media pesata dell'inverso delle distanze abitazioni-raffineria, con pesi pari ai tempi di durata del domicilio; nel calcolo sono considerate le abitazioni degli ultimi 20 anni (abitate per almeno 1 anno) con esclusione degli ultimi 5 anni, appartenenti quindi ad una finestra temporale di 15 anni.

Dei 531 (177 casi e 354 controlli) soggetti individuati dalla procedura di appaiamento caso-controllo come soggetti in studio, 493 sono stati rintracciati (165 casi e 328 controlli). Hanno accettato di rispondere all'intervista i familiari di 344 (95 casi e 249 controlli) di essi. Quindi le informazioni sull'esposizione sono disponibili per il 70% dei soggetti rintracciati e per il 65% del totale dei soggetti in studio. Nel presente documento, il gruppo di soggetti rintracciati per i quali è stato compilato il questionario, è detto gruppo dei soggetti 'Rispondenti' (344 soggetti); il gruppo dei soggetti per i quali non è stato compilato il questionario, è detto gruppo dei soggetti 'Non Rispondenti' (149 soggetti). Il presente rapporto confronta le caratteristiche dei due gruppi e valuta la possibilità che la stima dell'associazione tra rischio di morte per tumore del sistema emolinfopoietico ed esposizione residenziale sia alterata da una diversa percentuale di rispondenza dei casi e dei controlli entro categorie di esposizione.

#### 2. Metodi

Un confronto tra il gruppo dei 'Rispondenti' e dei 'Non rispondenti' è possibile solo per le caratteristiche disponibili per entrambi i gruppi: sesso, età, anno di riferimento (anno di morte per i casi e anno di morte del relativo caso appaiato per i controlli), status caso-controllo.

La differenza tra proporzioni di 'Rispondenti' e di 'Non Rispondenti' entro gruppi di caratteristiche dei soggetti è saggiata con  $\chi^2$  test. L'odds ratio di rispondenza relativo allo status-caso controllo, definito come

$$\frac{\text{Pr}(\text{rispondere} \mid \text{caso})}{1 - \text{Pr}(\text{rispondere} \mid \text{caso})} / \frac{\text{Pr}(\text{rispondere} \mid \text{controllo})}{1 - \text{Pr}(\text{rispondere} \mid \text{controllo})}$$

è stimato dal rapporto tra gli odds di rispondenza:

$$\frac{\text{N. casi rispondenti}}{\text{N. casi non rispondenti}} / \frac{\text{N. controlli rispondenti}}{\text{N. controlli non rispondenti}}$$

Per i soggetti 'Non rispondenti' non è nota la storia abitativa e non è possibile calcolare un punteggio di esposizione cumulata. Pertanto gli odds di selezione sono stati calcolati relativamente all'unica informazione disponibile per tutti i soggetti rintracciati: la residenza anagrafica al 2005. Il cut-off di 2 km per la distanza tra la residenza al 2005 e il centroide dell'API è stato usato per formare due categorie di esposizione: esposti (soggetti con distanza <2 km) e non esposti (soggetti con distanza  $\geq$  2 km).

L'odds di selezione per i casi è stato calcolato come rapporto tra le frazioni di casi rispondenti esposti e non esposti:

$$\text{odds}_{\text{selezione}}(\text{casi}) = \frac{\text{N. casi esposti rispondenti}}{\text{N. casi esposti rintracciati}} / \frac{\text{N. casi non esposti rispondenti}}{\text{N. casi non esposti rintracciati}}$$

Analogamente, l'odds di selezione per i controlli è calcolato come:

$$\text{odds}_{\text{selezione}}(\text{controlli}) = \frac{\text{N. controlli esposti rispondenti}}{\text{N. controlli esposti rintracciati}} / \frac{\text{N. controlli non esposti rispondenti}}{\text{N. controlli non esposti rintracciati}}$$

L'entità della sovrastima o sottostima (BIAS) è legata al rapporto tra gli odds di selezione come segue:

$$\text{BIAS} = [\text{odds}_{\text{selezione}}(\text{casi}) / \text{odds}_{\text{selezione}}(\text{controlli})] - 1.$$

Tutti i test statistici sono a due code e considerati statisticamente significativi per  $p$  value  $\leq$  0.05. Le analisi sono state eseguite con il software STATA, versione 9.0 (Stata Corporation, College Station, TX, USA, 2005).

#### 3. Caratteristiche dei soggetti non rispondenti

La Tabella 1 presenta le frequenze di rispondenti e di non rispondenti entro categorie di caratteristiche. I due gruppi sono ben bilanciati per quanto riguarda età e anno di riferimento. La proporzione di non rispondenti è maggiore tra gli uomini che tra le donne (55.7% vs 44.3%), con differenza non statisticamente significativa (Tabella 1). La proporzione di rispondenti è significativamente minore nei casi che nei controlli (27.6% vs 47.0,  $p < 0.001$ ) (Tabella 1) con rapporto tra gli odds di rispondenza pari a 0.43 (95% CI = 0.29-0.65). Tale differenza di per sè non comporta distorsioni nella valutazione dell'associazione tra status caso-controllo ed esposizione se essa si verifica in uguale misura entro categorie di esposizione. La distanza tra la residenza al 2005 e il centroide della raffineria API è risultata in media pari a  $4.02 \pm 1.87$  km (range: 0.34-8.14 km).

**Tabella 1. Caratteristiche dei soggetti rintracciati, rispondenti e non rispondenti**

	Rispondenti		Non rispondenti		$\chi^2$ p*
	N	%	N	%	
<i>Sesso</i>					
Uomini	171	49.7	83	55.7	0.221
Donne	173	50.3	66	44.3	
<i>Età, anni</i>					
<20	5	1.4	2	1.3	0.947
20-39	11	3.2	6	4.0	
40-59	26	7.6	14	9.4	
60-79	175	50.9	74	49.7	
>80	127	36.9	53	35.6	
<i>Anno di riferimento</i>					
1994-1995	51	14.8	13	8.7	0.451
1996-1997	72	20.9	32	21.5	
1998-1999	69	20.1	30	20.1	
2000-2001	68	19.8	33	22.2	
2002-2003	84	24.4	41	27.5	
<i>Status caso/controllo</i>					
Caso	95	27.6	70	47.0	<0.001
Controllo	249	72.4	79	53.0	
<i>Totale</i>	344	100.0	149	100.0	

In Tabella 2 sono riportate le frequenze dei soggetti rintracciati e dei 'Rispondenti' in relazione allo status caso-controllo e alla distanza tra la residenza al 2005 e la raffineria; in Tabella 3 sono riportate le relative probabilità di selezione, nell'ipotesi che tale residenza sia rappresentativa dell'esposizione cumulata. Tra i casi, quelli con distanza minore di 2 km hanno probabilità di selezione (risposta) di poco superiore a quella di casi con distanza maggiore; l'odds di selezione risulta pari a  $0.621/0.556 = 1.10$ ; viceversa tra i controlli la probabilità di rispondere è di poco maggiore per residenze con distanza > 2 km con odds di selezione pari a  $0.710/0.770 = 0.92$  (Tabella 3). L'entità del bias risulta pari a  $\frac{1.10}{0.92} - 1 = 0.20$ ; quindi, nell'ipotesi che la residenza al 2005 sia rappresentativa dell'esposizione cumulata, l'odds ratio di esposizione risulterebbe sovrastimato del 20%.

**Tabella 2. Frequenza dei soggetti rintracciati e dei soggetti rispondenti in categorie di status caso-controllo e di distanza tra la residenza al 2005 e la raffineria API**

	Soggetti rintracciati		Soggetti rispondenti		
	Distanza residenza (al 2005) -API		Distanza residenza (al 2005) -API		
	<2 km	≥ 2 km	<2 km	≥ 2 km	
Casi	29	136	Casi	18	77
Controlli	62	265	Controlli	44	204

Nota. Per un soggetto rintracciato (controllo) l'indirizzo della residenza al 2005 è risultato errato e non correggibile. Pertanto il totale dei soggetti rintracciati in tabella è 492 anzichè 493.

**Tabella 3. Probabilità di selezione in categorie di status caso-controllo e di distanza tra la residenza al 2005 e la raffineria API**

	Distanza residenza (al 2005) - API		Odds di selezione
	<2 km	≥ 2 km	
Casi	0.621 = 18/29	0.556 = 77/136	1.10
Controlli	0.710 = 44/62	0.770 = 204/265	0.92

#### **4. Conclusioni**

I familiari dei casi hanno risposto significativamente meno dei familiari dei controlli. Tuttavia, se la minor rispondenza dei casi si verificasse in ugual misura tra soggetti esposti e tra soggetti non esposti, essa non comporterebbe necessariamente a una distorsione nella stima dell'associazione tra rischio di morte per tumore del sistema emolinfopoietico ed esposizione cumulata residenziale, quest'ultima non nota per tutti i 531 soggetti in studio.

Il risultato di questa analisi si basa sull'ipotesi che la residenza al 2005, unica informazione disponibile per tutti i soggetti in studio, sia rappresentativa dell'esposizione cumulata nei 15 anni corrispondenti alla finestra temporale. In questa ipotesi, i casi esposti risultano sovra-rappresentati del 10% circa, mentre i controlli esposti sono sotto-rappresentati del 10% circa, con possibilità di sovrastimare l'associazione tra rischio di morte per tumore del sistema emolinfopoietico ed esposizione residenziale del 20% circa. Deve comunque essere sottolineato che non essendo possibile stabilire il grado di attinenza tra la residenza al 2005 e l'esposizione cumulata, questa valutazione può essere poco attendibile.

#### **Bibliografia**

- 1) Kleinbaum DG, Kupper LL, Morgenstern H, eds. Epidemiologic Research: Principles and Quantitative Methods. New York: Van Nostrand Reinhold Company Inc., 1982.
- 2) Rothman KJ, Greenland Sander, eds. Modern Epidemiology. Philadelphia (PA): Lippincott-Raven Publishers, 1998.

## Annexo 4 - Programma di analisi statistica (21 luglio 2008)

### Introduzione

Il presente progetto per valutare l'associazione tra rischio di morte per tumore del sistema emolinfopoietico ed esposizione residenziale alle ricadute della raffineria API (nel seguito Studio API), studierà gli odds ratio (OR) di esposizione con i relativi intervalli di confidenza al 95%. La stima sarà prodotta mediante l'applicazione di modelli di regressione logistica condizionata che includeranno la variabile di esposizione e potenziali confondenti. Uno dei principali problemi di analisi in queste indagini è la modalità di calcolo dei livelli di esposizione dei soggetti in studio. Il presente documento discute sulle metodiche disponibili e descrive l'approccio applicato allo studio API.

### 1. Studio dell'esposizione

Il calcolo di un punteggio di esposizione residenziale cumulata si basa sulla variabile 'distanza tra abitazione e centroide della fonte inquinante' (*intensità dell'esposizione*) e sulla variabile 'numero di anni trascorsi nell'abitazione' (*durata dell'esposizione*). Per i soggetti che hanno avuto più di un'abitazione, ad oggi non è disponibile un metodo condiviso per aggregare opportunamente *intensità* e *durata* di esposizione di ogni abitazione. Una ricerca bibliografica su recenti studi epidemiologici, che hanno valutato l'associazione tra rischio morte (o incidenza) per tumore e storia espositiva, ha evidenziato che i metodi adottati sono sostanzialmente riconducibili a quattro approcci (A, B, C, D).

#### 1.1 Approccio (A)

Descrizione Un primo approccio suggerisce di moltiplicare il punteggio assegnato all'*intensità* ( $\mu_i$ ) e la *durata* ( $t_i$ ) di esposizione di ogni residenza e poi di sommare i risultati ottenuti:

$$(1) \quad \sum_i \mu_i \cdot t_i,$$

dove  $\mu_i = 1/d_i$  è l'inverso della distanza della  $i$ -esima abitazione dalla sorgente inquinante. Questo approccio è stato adottato in studi sull'esposizione alle ricadute di impianti petrolchimici [1], a fumo attivo (con  $\mu_i$  pari al numero di sigarette/dì [2] o al numero di pacchetti/dì fumati nel periodo $_i$  [3]) e a fumo passivo (con  $\mu_i$  pari al numero di fumatori presenti nell'ambiente $_i$ ) [4].

Osservazioni La formula (1) assume che *intensità* e *durata* di esposizione abbiano lo stesso effetto sul rischio di morte/incidenza per tumore. Ad esempio, un'abitazione distante 500 m dalla sorgente inquinante e abitata per 10 anni conferisce lo stesso punteggio di esposizione di un'abitazione distante 1 km e abitata per 20 anni.

#### 1.2 Approccio (B)

Descrizione Un secondo approccio considera i tempi di esposizione come pesi per calcolare una media ponderata delle *intensità* di esposizione:

$$(2) \quad \frac{\sum_i \mu_i \cdot t_i}{\sum_i t_i}.$$

Questo metodo è stato applicato per esposizioni residenziali a campi elettromagnetici (con  $\mu$  pari all'intensità del campo misurata nell'abitazione [5-7]) e ad aree industrializzate (con  $\mu$  pari alla distanza dell'abitazione [8]).

Osservazioni L'approccio (B) ha il vantaggio di non fare ipotesi sulla possibile relazione tra *durata* e *intensità* d'esposizione. Tuttavia un limite della formula (2) è che a parità di distanza, soggetti con tempi di esposizione molto diversi possono avere lo stesso punteggio; infatti, il peso di chi ha vissuto 10 dei suoi 20 anni di vita

$\left( \frac{t_i}{\sum_i t_i} = \frac{10}{20} = 0.5 \right)$  è uguale al peso di chi ha vissuto 20 dei suoi 40 anni di vita  $\left( \frac{t_i}{\sum_i t_i} = \frac{20}{40} = 0.5 \right)$ . La formula (2)

è quindi utilizzabile quando il range d'età dei soggetti non è molto esteso, come negli studi sulle leucemie infantili [5-7], o quando la storia espositiva è riferita ad un periodo di tempo (finestra temporale) limitato [8].

#### 1.3 Approccio (C)

Descrizione Altri studi hanno adottato il criterio di semplificare l'informazione sulla storia residenziale considerando una delle due variabili *intensità* o *durata* di esposizione come variabili dicotomiche del tipo sì/no (variabili dummy). L'approccio (C) suggerisce di definire all'interno dell'area in studio una zona di esposizione, tipicamente quella entro un dato raggio dal centroide della sorgente inquinante, e una zona di non esposizione. Con questo metodo la formula (1) diviene la sommatoria dei tempi in cui l'abitazione del soggetto è stata all'interno della zona esposta, e l'esposizione cumulata risulta pari a:

$$(3) \quad \sum_i \mu_i \cdot t_i,$$

dove  $\mu_i = 0$  per abitazioni esterne alla zona esposta e  $\mu_i = 1$  per abitazioni interne alla zona esposta [9].

Osservazioni La formula (3) ha il vantaggio di non fare assunzioni sulla possibile relazione tra *intensità* e *durata* di esposizione e di non avere limiti di applicabilità per quanto riguarda la finestra temporale, ma rappresenta una forte semplificazione per quanto concerne l'*intensità* di esposizione poiché abitazioni poco oltre il limite della zona considerata esposta conferiscono punteggio nullo indipendentemente dai tempi di esposizione.

#### 1.4 Approccio (D)

Descrizione Un altro approccio che opera una semplificazione in presenza di più abitazioni prevede di sceglierne una come abitazione principale (*main residence*) e conteggiare solo l'*intensità* di esposizione relativa ad essa. Baris et al. discutono 3 possibili scelte di *main residence* nell'ambito della valutazione di esposizioni a campi elettromagnetici: (a) abitazione nella quale si è vissuto più a lungo, (b) abitazione attuale e (c) abitazione precedente [7]. Belli et al. nello studiare l'associazione tra rischio di morte per tumore del polmone, della vescica, o tumore emolinfopoiotico e distanza dalla raffineria di Brindisi adottano come *main residence* quella in cui il soggetto ha vissuto più a lungo, con esclusione degli ultimi 10 anni [10].

Osservazioni L'inconveniente di scegliere una *main residence* e trascurare l'esposizione legata ad altre abitazioni è che nella stessa categoria di *intensità* ci possono essere soggetti con *durate* di esposizione molto diverse. Nell'ambito dell'approccio (D), il metodo adottato da Belli et al. rappresenta un compromesso tra abitazione nella quale si è vissuto più a lungo e abitazione precedente che permette di tenere parzialmente conto sia della durata dell'esposizione che del periodo di latenza tra esposizione e morte per tumore.

## 2. Studio API

Il presente programma di analisi seguirà l'approccio (B) congiuntamente all'introduzione di una finestra temporale. Poiché i tumori del sistema emolinfopoiotico hanno minori tempi di latenza dei tumori epiteliali [11-12], l'introduzione della finestra temporale ha lo scopo di utilizzare solo l'informazione relativa all'intervallo di tempo entro il quale si ritiene che l'esposizione della popolazione abbia prodotto i casi in studio. Nel presente programma di analisi si assume che i casi deceduti nel periodo 1994 - 2003 provengano dal periodo di esposizione 1974 - 1998 con finestra temporale pari a 15 anni. Ad esempio, si assume che un caso deceduto il 01.01.1994 provenga dalla popolazione esposta nel periodo 01.01.1974 - 01.01.1989 e che un caso deceduto il 31.12.2003 provenga dalla popolazione esposta nel periodo 31.12.1983 - 31.12.1998. Il metodo (B) risulta equivalente al metodo (A) per i soggetti con età maggiore dell'ampiezza della finestra temporale, mentre, per soggetti con età minore, il punteggio assegnato alla singola abitazione dalla formula (2) aumenta, a parità di distanza e durata, al diminuire dell'età.

Il punteggio di esposizione cumulata sarà calcolato come media pesata dei punteggi di esposizione attribuiti alle abitazioni che il soggetto in studio ha abitato negli ultimi 20 anni (per almeno 1 anno) con esclusione degli ultimi 5 anni. Il punteggio di esposizione per la singola abitazione verrà calcolato sia tenendo conto solo della distanza che tenendo conto della direzione dei venti prevalenti nell'area in studio secondo il modello di dispersione e ricaduta al suolo degli inquinanti AERMOD [13].

### 2.1 Variabili in studio

Le variabili in studio saranno costruite a partire dalle informazioni, raccolte e trasferite su file a cura del Servizio di Epidemiologia Ambientale dell'ARPAM-Dipartimento di Ancona (dott. Mauro Mariottini), contenute nel questionario (Allegato A). Anche se non specificato nel seguito, esse saranno tutte riferite all'età e all'anno di riferimento del soggetto in studio: per i casi età alla data del decesso e anno del decesso e per i controlli età alla data del decesso e anno del decesso del caso appaiato.

Le informazioni sulla storia abitativa sono state raccolte fino ad un massimo di 4 abitazioni negli ultimi 40 anni precedenti l'anno di riferimento, vi sono quindi fino a 4 periodi contraddistinti da una diversa abitazione:  $t_{res-i}$  ( $i=1, 2, 3, 4$ ). A cura del Servizio di Epidemiologia Ambientale, le coordinate geografiche delle abitazioni appartenenti all'*area in studio* (definita amministrativamente e composta dai comuni di Falconara Marittima, Chiaravalle e Montemarciano) sono state rilevate manualmente con GPS (e successivamente verificate su carta digitale) nel caso di esistenza del numero civico. In assenza di questo, con il software di georeferenziazione automatica degli indirizzi, GEOLOCATE, sono state individuate le coordinate geografiche dell'abitazione attribuendole il punto mediano dell'arco stradale sul quale essa è posizionata. Le distanze (in km) delle abitazioni dal centroide dell'impianto API sono state calcolate con il software GIS Mapinfo Professional v.8.0.

Per ogni soggetto verrà calcolato il periodo di riferimento abitativo,  $T_{res} = \sum_{i=1}^4 t_{res-i}$  e la  $i$ -esima frazione del periodo di

riferimento data da  $\tau_{res-i} = \frac{t_{res-i}}{T_{res}}$ , considerando solo le case abitate per almeno 1 anno nei 20 anni precedenti l'anno di

riferimento con esclusione degli ultimi 5 anni.

Lo status occupazionale (casalinga, studente, pensionato, occupato, non occupato) nei 40 anni precedenti l'anno di riferimento è stato raccolto per un massimo di 8 periodi ( $t_{occ-i}$ ) di almeno 1 anno; per i periodi in cui il soggetto è stato occupato, si è registrato il luogo e il tipo di attività. Per ogni soggetto sarà calcolato il periodo riferimento occupazionale  $T_{occ} = \sum_{i=1}^8 t_{occ-i}$ , e la sua i-esima frazione  $\tau_{occ-i} = \frac{t_{occ-i}}{T_{occ}}$ , considerando le occupazioni di almeno 1 anno nei 20 anni precedenti l'anno di riferimento con esclusione degli ultimi 5.

Le variabili in studio di tipo numerico verranno studiate sia come variabili continue che come variabili categoriche. Per le variabili continue, il metodo dei polinomi frazionali sarà usato per scegliere la forma funzionale che meglio si adatta all'ipotesi di linearità sulla scala logit. Per la forma categorica si useranno come cut-point i terzili della distribuzione nei controlli. Per variabili  $X$  con distribuzione fortemente asimmetrica e con un'alta frequenza di valori nulli (come potrebbero risultare variabili quali l'esposizione cumulata al fumo) si userà la seguente categorizzazione:  $X_{cat} = 0$  se  $X = 0$ ,  $X_{cat} = 1$  se  $0 < X \leq$  mediana (distribuzione nei controlli),  $X_{cat} = 2$  se  $X >$  mediana (distribuzione nei controlli).

Tutti i test statistici saranno a due code e considerati statisticamente significativi per  $p \text{ value} \leq 0.05$ . Le analisi saranno eseguite con il software STATA, versione 9.0 (Stata Corporation, College Station, TX, USA, 2005).

## 2.2 Calcolo dell'esposizione

La variabile d'esposizione primaria è:

$$EXP_{res} \text{ (continua o categorica)} = \frac{\sum_i \mu_{res-i} \cdot t_{res-i}}{\sum_i t_{res-i}},$$

dove

$$\mu_{res-i} = \begin{cases} \frac{1}{d_{res-i}} & \text{se l'abitazione} \in \text{area in studio} \\ 0 & \text{se l'abitazione} \notin \text{area in studio} \end{cases}$$

Seguendo l'approccio che tiene conto solo delle distanze,  $d_{res-i}$  è la distanza tra abitazione e centroide dello stabilimento API. Nell'analisi che terrà conto della ricaduta delle polveri,  $\mu_{res-i}$  sarà il punteggio di esposizione stimato dal modello di dispersione e ricaduta al suolo degli inquinanti AERMOD.

Verrà comunque considerata anche la modalità

$$\mu_{res-i} = \begin{cases} 1 & \text{se } d_{res-i} \leq 2 \text{ km} \\ 0 & \text{se } d_{res-i} > 2 \text{ km} \end{cases},$$

perchè anche se opera una forte semplificazione sulla stima dell'intensità di esposizione consente di non fare assunzioni sulla possibile relazione tra intensità e durata di esposizione, e può essere in grado di rilevare un rischio di morte per tumore del sistema emolinfopoietico da esposizione residenziale nel caso in cui esso sia sufficientemente elevato.

## 2.3 Covariate

Le covariate saranno studiate sia singolarmente che attraverso la procedura di selezione stepwise con  $p_E = 0.15$  e  $p_R = 0.20$ . In ogni caso, indipendentemente dall'esito della selezione, variabili ritenute importanti potranno essere incluse nel modello finale.

Età e sesso saranno incluse nel modello di regressione logistica condizionata come variabili di matching. Di seguito sono elencate le altre variabili studiate come potenziali confondenti dell'associazione.

### 2.3.1 Caratteristiche personali

*Età* (continua). Sebbene l'età sia una delle variabili di matching, può essere necessario introdurla ulteriormente nel modello come covariata

*StatoCivile2* (nominale) = 1 se coniugato  
2 se celibe/nubile OR vedovo OR separato/divorziato

*Istruzione* (nominale) = 1 se nessun titolo OR licenza elementare OR licenza media inferiore  
OR altro titolo inferiore alla licenza media superiore  
2 se licenza media superiore OR diploma universitario/laurea breve OR  
laurea vecchio ord./specialistica.

### 2.3.2 Comportamenti a rischio

*Fumo\_Status* (nominale)= 2 se Fumatore  
 1 se Ex fumatore  
 0 se Non fumatore

*Sig-anni* (continua o categorica) =  $(et\grave{a}_{fine} - et\grave{a}_{inizio}) \cdot \ln(Sig\_media)$  se fumatore o ex fumatore  
 0 se non fumatore

dove  $\ln(Sig\_media)$  è il logaritmo della media giornaliera di sigarette fumate (Thurston, CEBP 2005), così attribuita:

Sigarette/dì (c218)	<i>Sig_media</i>	$\ln(Sig\_media)$
1-3	2	0.69
4-10	7	1.95
11-20	15	2.71
21-30	25	3.22
31-40	35	3.56
41-60	50	3.91
oltre 3 pacchetti	60	4.09

### 2.3.3. Patologie, esposizioni a radiazioni ionizzanti, radioterapia

*Mgeneriche* (nominale) = 1 se almeno una delle seguenti malattie: Polmonite OR Cardiopatie OR Diabete OR Epatiti OR Emofilia OR Inf. Vie urinarie OR Menigite OR Tubercolosi OR Asma OR Ulcera OR Allergia OR Anemia OR Mononucleosi OR Gastrite OR Infiam. Ghiandole  
 0 se nessuna delle sopraelencate malattie

*Mrare* (nominale) = 1 se almeno una delle seguenti malattie: Ataxia OR Neurofibromatosi OR Immunodeficienza eredit. OR Sindrome di Down OR Riarrangiamenti cromosomici MLL e TEL-AML 1  
 0 se nessuna delle sopraelencate malattie

*Mspecifiche* (nominale) = 1 se almeno una delle seguenti malattie: Epatite OR HIV OR RNA virus OR HTLV II virus  
 0 se nessuna delle sopraelencate malattie

*CAMaligni* (nominale): 1 = sì, 0 = no

*Chemio* (nominale): 1 = sì, 0 = no

*RXdiagnostica* (nominale): 1 = sì, 0=no

Per ogni *esamej* eseguito a scopo diagnostico

*esame1* radiografia  
*esame2* schermografia  
*esame3* TAC  
*esame4* scintigrafia

$RXesame_j$  (continua o categorica) =  $\sum_i parte_i \cdot n_i$ , dove:

$parte_1 = 1$  se testa o arti,  $parte_2 = 2$  se collo o torace o addome,  $parte_3 = 3$  se tutto il corpo e  $n_i =$  numero di esami fatti alla  $parte_i$ .

$RXdiagCum$  (continua o categorica) =  $\sum_i esame_i \cdot parte_i \cdot n_i$ , dove:

$esame_i = 2$  se schermografia o TAC o scintigrafia,  $esame_i = 1$  se radiografia,  $parte_i = 1$  se testa o arti,  $parte_i = 2$  se collo o torace o addome,  $parte_i = 3$  se tutto il corpo, e  $n_i =$  numero di esami fatti alla  $parte_i$ .

*RXintervento* (nominale): 1 = sì, 0 = no

Per ogni intervento con *esamej* eseguito durante un intervento chirurgico:

<i>esame1</i>	radiografia
<i>esame2</i>	schermografia
<i>esame3</i>	TAC
<i>esame4</i>	scintigrafia

$$RX_{interv_j} \text{ (continua o categorica)} = \sum_i parte_i \cdot n_i ,$$

dove:  $parte_1 = 1$  se testa o arti,  $parte_2 = 2$  se collo o torace o addome,  $parte_3 = 3$  se tutto il corpo e  $n_i$  = numero di esami fatti alla  $parte_i$ .

$$RX_{intervCum} \text{ (continua o categorica)} = \sum_i esame_i \cdot parte_i \cdot n_i ,$$

$esame_i = 2$  se schermografia o TAC o scintigrafia,  $esame_i = 1$  se radiografia,  $parte_i = 1$  se testa o arti,  $parte_i = 2$  se collo o torace o addome,  $parte_i = 3$  se tutto il corpo, e  $n_i$  = numero di esami fatti alla  $parte_i$  durante l'intervento.

$$RX_{tot} \text{ (continua o categorica)} = RX_{intervCum} + RX_{diagCum}$$

Radioterapia (nominale): 1 = sì, 0=no

### 2.3.4 Familiarita'

*CApadre* (ordinale) = 2 se padre con diagnosi di (leucemia OR linfoma OR mieloma)  
 1 se padre con diagnosi di altro tumore  
 0 se nessuna diagnosi di tumore.

*CAmadre* (ordinale) = 2 se madre con diagnosi di (leucemia OR linfoma OR mieloma)  
 1 se madre con diagnosi di altro tumore  
 0 se nessuna diagnosi di tumore.

*CAgenitori* = 0 se entrambi i genitori senza diagnosi di tumore  
 1 se 1 solo genitore con diagnosi di altro tumore  
 2 se entrambi i genitori con diagnosi di altro tumore  
 3 se almeno 1 genitore con diagnosi di (leucemia OR linfoma OR mieloma)

*CAfratelli/sorelle* = 0 se tutti i fratelli/sorelle senza diagnosi di tumore  
 1 se 1 solo fratello/sorella con diagnosi di altro tumore  
 2 se almeno 2 fratelli/sorelle con diagnosi di altro tumore  
 3 se almeno 1 fratello/sorella con diagnosi di (leucemia OR linfoma OR mieloma)

*CAfconsanguinei* = 0 se tutti i consanguinei senza diagnosi di tumore  
 1 se 1 solo consanguineo con diagnosi di altro tumore  
 2 se almeno 2 consanguinei con diagnosi di altro tumore  
 3 se almeno 1 consanguineo con diagnosi di (leucemia OR linfoma OR mieloma)

### 2.3.5 Occupazione

Le informazioni sulla storia occupazionale sono state raccolte fino ad un massimo di 8 diverse occupazioni svolte per almeno 1 anno negli ultimi 40 anni precedenti l'anno di riferimento.

Per assegnare un punteggio di esposizione professionale saranno considerate solo le attività svolte per almeno 1 anno nei 20 anni precedenti l'anno di riferimento con esclusione degli ultimi 5 e due diverse fonti di valutazione dei rischi occupazionali:

-la matrice MATline professione-esposizione sviluppata dal Centro di Documentazione per la Promozione della Salute della Regione Piemonte (DoRS) [14],

-la matrice della letteratura che, per i diversi comparti produttivi, riporta gli articoli scientifici che hanno trovato un'associazione tra esposizione al comparto produttivo e rischio di tumore, sviluppata dal progetto OCCAM (OCcupational CANcer Monitoring) con la collaborazione tra ISPESL (Istituto Superiore per la Prevenzione e la Sicurezza sul Lavoro) e Fondazione IRCCS 'Istituto Nazionale dei Tumori' di Milano [15].

A cura del Servizio di Epidemiologia Ambientale dell'ARPAM-Dipartimento di Ancona (dott. Mauro Mariottini), le attività descritte nel questionario sono state riclassificate secondo le attività economiche incluse nella matrice MATline e secondo i comparti produttivi considerati nel sistema OCCAM.

### 2.3.5.1. Modalità DoRS.

Ad ogni attività economica sarà assegnato il punteggio  $r_{DoRS} = 1$  se nella matrice MATline essa è associata ad una delle sostanze classificate come cancerogeni di classe 1 per il sistema emolinfopoietico, ovvero benzene [Monografie IARC 1974, Vol. 7; 1982 Vol. 29; 1987 Suppl. 7], vinile cloruro [Monografie IARC 1974, Vol. 7; 1979 Vol. 19; 1987, Suppl. 7], etilene ossido [Monografie IARC 1994, Vol 60] e radiazioni x e gamma [Monografie IARC 2000, Vol 75]; diversamente sarà assegnato il punteggio  $r_{DoRS} = 0$ . In particolare  $r_{DoRS} = 0$  per le categorie di casalinga, studente, pensionato e non occupato. Saranno quindi costruite due variabili di esposizione occupazionale:

$ExpDoRS$  (nominale) = 1 se almeno una delle attività ha punteggio  $r_{DoRS} = 1$   
 0 se tutte le attività hanno punteggio  $r_{DoRS} = 0$

$$ExpDoRScum \text{ (numerica o categorica)} = \frac{\sum_i r_{DoRS-i} \cdot t_{occ-i}}{\sum_i t_{occ-i}},$$

dove  $t_{occ-i}$  indica il numero di anni impegnati nella i-esima occupazione.

### 2.3.5.2. Modalità OCCAM.

Ad ogni attività sarà assegnato il punteggio  $r_{OCCAM} = 1$  se essa corrisponde ad un comparto produttivo per il quale la matrice della letteratura OCCAM riporta almeno 4 articoli; diversamente sarà assegnato il punteggio  $r_{OCCAM} = 0$ . In particolare  $r_{OCCAM} = 0$  per le categorie di casalinga, studente, pensionato e non occupato. Saranno quindi costruite due variabili di esposizione occupazionale:

$ExpOCCAM$  (nominale) = 1 se almeno una delle attività ha punteggio  $r_{OCCAM} = 1$   
 0 se tutte le attività hanno punteggio  $r_{OCCAM} = 0$

$$ExpOCCAMcum \text{ (numerica o categorica)} = \frac{\sum_i r_{OCCAM-i} \cdot t_{occ-i}}{\sum_i t_{occ-i}}.$$

## 2.3.6 Caratteristiche dell'abitazione

Le abitazioni prese in considerazione sono quelle abitate per almeno 1 anno nei 20 anni precedenti l'anno di riferimento con esclusione degli ultimi 5.

$Urbana$  (nominale) = 1 se almeno un'abitazione in zona urbana  
 0 se nessuna abitazione in zona urbana

$$UrbanaCum \text{ (continua o categorica)} = \frac{\sum_i Urb_i \cdot t_{res-i}}{\sum_i t_{res-i}},$$

dove  $Urb_i$  è la variabile dummy che contraddistingue le abitazioni in zona urbana

$Unifamiliare$  (nominale) = 1 se almeno un'abitazione unifamiliare  
 0 se nessuna abitazione unifamiliare

$$UnifamCum \text{ (continua o categorica)} = \frac{\sum_i Unifam_i \cdot t_{res-i}}{\sum_i t_{res-i}},$$

dove  $Unifam_i$  è la variabile dummy che contraddistingue le abitazioni unifamiliari

$Pianoterra$  (nominale) = 1 se almeno un'abitazione al pianoterra  
 0 se nessuna abitazione al pianoterra

$$PianoterraCum \text{ (continua o categorica)} = \frac{\sum_i PianoT_i \cdot t_{res-i}}{\sum_i t_{res-i}},$$

dove  $PianoT_i$  è la variabile dummy che contraddistingue le abitazioni al piano terra

$Combustibile$  (nominale) = 1 se almeno un'abitazione con riscaldamento autonomo a legna OR a carbone  
 0 se nessuna abitazione con riscaldamento autonomo a legna OR a carbone

$$CombustibileCum \text{ (continua o categorica)} = \frac{\sum_i Comb_i \cdot t_{res-i}}{\sum_i t_{res-i}},$$

dove  $Comb_i$  è la variabile dummy che indica la presenza di riscaldamento autonomo a legna OR a carbone nell'abitazione  $i$ .

$FumoPassivo$  (nominale) = 1 se almeno un'abitazione con coabitanti fumatori  
 0 se nessuna abitazione con coabitanti fumatori

$$FumoPassivoCum \text{ (continua o categorica)} = \frac{\sum_i FP_i \cdot t_{res-i}}{\sum_i t_{res-i}},$$

dove  $FP_i$  è la variabile dummy che contraddistingue le abitazioni con coabitanti fumatori.

### 2.3.7 Strutture a rischio

Le variabili che indicano la presenza di strutture a rischio nei pressi dell'abitazione (abitate per almeno 1 anno nei 20 anni precedenti l'anno di riferimento, con esclusione degli ultimi 5) saranno studiate:

a) singolarmente come status:

$STR$  (nominale) = 1 se almeno una struttura a rischio per almeno una abitazione  
 0 se nessuna abitazione nei pressi di strutture a rischio.

b) singolarmente, introducendo nel modello una alla volta le seguenti variabili:

$$Str_j \text{ (continua o categorica)} = \frac{\sum_{i=1} x_{j,res-i} \cdot t_{res-i}}{\sum_i t_{res-i}},$$

dove:

$res-i$ , indice di abitazione

$t_{res-i}$ , anni trascorsi nell'abitazione  $i$

$j=1, 2, \dots, 13$  indice di struttura a rischio:

- $j = 1$  deposito carburante <200m
- $j = 2$  distributori carburante <200m
- $j = 3$  antenne CEM <200m
- $j = 4$  tralicci ELF <200m
- $j = 5$  industria tessile/calzaturificio <200m
- $j = 6$  raffineria petrolio <200m
- $j = 7$  lavanderie <200m
- $j = 8$  discarica <200m
- $j = 9$  centrale/istallazione elettrica <200m
- $j = 10$  carrozzeria <200m
- $j = 11$  industria metalmeccanica <200m
- $j = 12$  cava o miniera <200m
- $j = 13$  traffico < 100m

$x_{j,res-i}$ , dummy di presenza della struttura  $j$  nei pressi dell'abitazione  $i$ .

c) cumulativamente senza pesare il tipo di struttura a rischio:

$$StrCumI(\text{continua o categorica}) = \frac{\sum_{j=1}^{13} y_j \sum_{i=1}^{13} x_{j,res-i} \cdot t_{res-i}}{\sum_{i=1}^{13} t_{res-i}}$$

con  $y_j = 1$  per ogni  $j=1,2,\dots,13$ ;

### 2.3.8 Aree coltivate

Aree coltivate si riferisce alla presenza di suoli agricoli nei pressi delle abitazioni abitate per almeno 1 anno nei 20 anni precedenti l'anno di riferimento, con esclusione degli ultimi 5.

*Campi* (nominale) = 1 se almeno un'abitazione vicino a suoli coltivati,  
0 se nessuna abitazione vicino a suoli coltivati

$$CampiCum(\text{continua o categorica}) = \frac{\sum_{i=1}^{13} Colt_{res-i} \cdot t_{res-i}}{\sum_{i=1}^{13} t_{res-i}},$$

dove  $Colt_{res-i}$  è la variabile dummy che indica la presenza di suoli coltivati nei pressi dell'abitazione<sub>i</sub>.

*FitoCampi* (nominale) = 1 se nei pressi di almeno un'abitazione è stato fatto uso di fitofarmaci  
(pesticidi OR fungicidi OR insetticidi OR erbicidi)  
0 se nei pressi di nessuna abitazione è stato fatto uso di fitofarmaci

$$FitoCampiCum(\text{continua o categorica}) = \frac{\sum_{i=1}^{13} fito_{res-i} \cdot t_{res-i}}{\sum_{i=1}^{13} t_{res-i}},$$

dove  $fito_{res-i}$  è la variabile dummy che indica l'uso di fitofarmaci (pesticidi OR fungicidi OR insetticidi OR erbicidi) nei pressi dell'abitazione<sub>i</sub>.

### 2.3.9 Attività agricola

Attività agricola si riferisce ad attività svolta dal soggetto in studio nei pressi di abitazioni abitate per almeno 1 anno nei 20 anni precedenti l'anno di riferimento, con esclusione degli ultimi 5.

*AttAgr* (nominale) = 1 se presso almeno un'abitazione il soggetto ha svolto regolarmente attività  
agricola (anche non professionale),  
0 se presso nessuna abitazione il soggetto ha svolto regolare attività agricola

$$AttAgrCum(\text{continua o categorica}) = \frac{\sum_{i=1}^{13} Agr_{res-i} \cdot t_{res-i}}{\sum_{i=1}^{13} t_{res-i}},$$

dove  $Agr_{res-i}$  è la variabile dummy che indica se il presso l'abitazione<sub>i</sub> il soggetto ha svolto regolarmente attività agricola

*UsoFito* (nominale) = 1 se presso almeno un'abitazione il soggetto ha fatto regolare uso di  
fitofarmaci (pesticidi OR fungicidi OR insetticidi OR erbicidi)  
0 se presso nessuna abitazione il soggetto ha fatto uso di fitofarmaci

$$UsoFitoCum(\text{continua o categorica}) = \frac{\sum_{i=1}^{13} UsoF_{res-i} \cdot t_{res-i}}{\sum_{i=1}^{13} t_{res-i}},$$

dove  $UsoF_{res-i}$  è la variabile dummy che indica se il soggetto ha fatto regolare uso di fitofarmaci (pesticidi OR fungicidi OR insetticidi OR erbicidi) presso l'abitazione<sub>i</sub>.

## Bibliografia

1. Yu CL, Wang SF, Pan PC, Wu MT, Ho CK, Smith TJ, Li Y, Pothier L, Christiani DC; Kaohsiung Leukemia Research Group. Residential exposure to petrochemicals and the risk of leukemia: using geographic information system tools to estimate individual-level residential exposure. *American Journal of Epidemiology* 2006; 164(3):200-207.
2. Hjalgrim H, Ekström-Smedby K, Rostgaard K, Amini RM, Molin D, Hamilton-Dutoit S, Schöllkopf C, Chang ET, Ralfkiaer E, Adami HO, Glimelius B, Melbye M. Cigarette smoking and risk of Hodgkin lymphoma: a population-based case-control study. *Cancer Epidemiol Biomarkers Prev.* 2007 Aug;16(8):1561-6.
3. Lubin JH, Caporaso N, Wichmann HE, Schaffrath-Rosario A, Alavanja MC. Cigarette smoking and lung cancer: modeling effect modification of total exposure and intensity. *Epidemiology* 2007;18(5):639-48.
4. Kasim K, Levallois P, Abdous B, Auger P, Johnson KC, and The Canadian Cancer Registries Epidemiology Research Group. Environmental Tobacco Smoke and Risk of Adult Leukemia. *Epidemiology* 2005;16: 672–680.
5. Kleinerman RA, Linet MS, Hatch EE, et al. Magnetic field exposure assessment in a case-control study of childhood leukemia. *Epidemiology* 1997;5:575–83.
6. Linet MS, Hatch EE, Kleinerman RA, et al. Residential exposure to magnetic fields and acute lymphoblastic leukemia in children. *N Engl J Med* 1997;337:1–7.
7. Baris D, Linet MS, Tarone RE, Kleinerman RA, Hatch EE, Kaune WT, Robison LL, Lubin J, Wacholder S. Residential exposure to magnetic fields: an empirical examination of alternative measurement strategies. *Occup Environ Med* 1999;56:562–566.
8. Fano V, Michelozzi P, Ancona C, Capon A, Forastiere F, Perucci C A. Occupational and environmental exposures and lung cancer in an industrialised area in Italy. *Occup Environ Med* 2004;61:757–763.
9. Edwards R, Pless-Mulloli T, Howel D, Chadwick T, Bhopal R, Harrison R, Gribbin H. Does living near heavy industry cause lung cancer in women? A case-control study using life grid interviews. *Thorax.* 2006;61(12):1076-82.
10. Belli S, Benedetti M, Combal P, Lagravinese D, Martucci V, Martuzzi M, Morleo D, Trinca S, Viviano G. Case-control study on cancer risk associated to residence in the neighbourhood of a petrochemical plant. *European Journal of Epidemiology* 2004; 19: 49–54.
11. Richardson DB. Temporal variation in the association between benzene and Leukemia mortality. *Environ Health Perspect* 2008;116:370-4.
12. Finkelstein MM. Leukemia after exposure to benzene: temporal trends and implications for standards. *Am J Ind Med* 2000;38:1-7.
13. AERMOD: [http://www.epa.gov/scram001/dispersion\\_prefrec.htm#aermod](http://www.epa.gov/scram001/dispersion_prefrec.htm#aermod) .
14. DoRS, Documentazione Regionale Salute: <http://www.dors.it> .
15. Progetto OCCAM - Occupational Cancer Monitoring - Sistema informativo sui tumori di origine professionale: <http://www.occam.it> .

## Annesso 5 – Calcolo dell’esposizione residenziale cumulata e dell’esposizione occupazionale cumulata (19 settembre 2008)

### Introduzione

Il presente documento descrive il calcolo delle variabili di esposizione e dei principali fattori confondenti che verranno impiegate nella stima dell’odds ratio al fine di valutare l’associazione tra esposizione e rischio di tumore del sistema emolinfopoietico per lo studio API. Il documento sintetizza come queste variabili sono state costruite a partire dai dati sulla storia abitativa e occupazionale presenti nel database dello studio (realizzato a cura dell’ARPAM, dott. Mauro Mariottini e dott. Marco Baldini).

### 1. Variabile di esposizione residenziale cumulata.

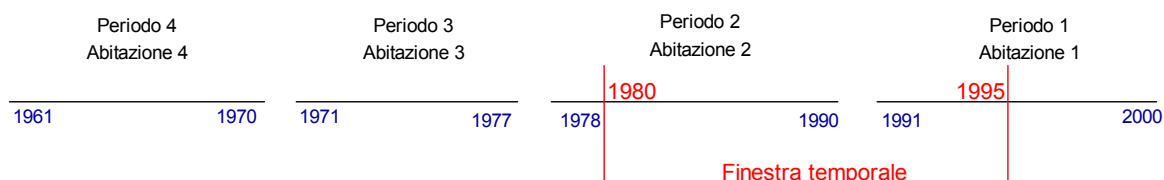
Il Programma di Analisi prevede di calcolare l’esposizione residenziale cumulata come media pesa dei punteggi attribuiti alle singole abitazioni che il soggetto in studio ha abitato entro una finestra temporale di 15 anni, situata nei 20 anni precedenti l’anno di riferimento, con esclusione degli ultimi 5 anni:

$$expRES\_cum \text{ (continua o categorica)} = \frac{\sum_i \mu_{res-i} \cdot t_{res-i}}{\sum_i t_{res-i}}$$

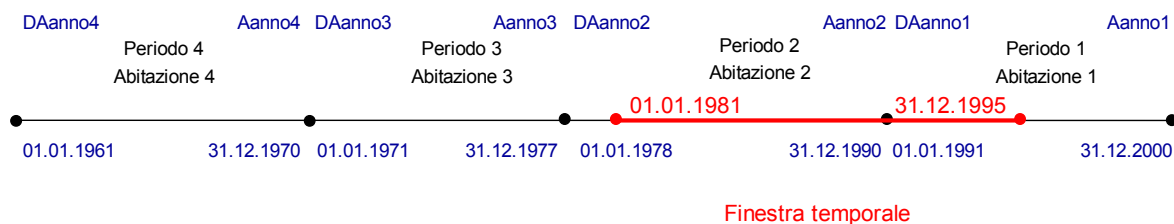
Di seguito è descritto come sono stati calcolati i punteggi di esposizione per la singola abitazione,  $\mu_{res-i}$  e il numero di anni trascorsi nell’abitazione ( $i=1, 2, 3, 4$ ),  $t_{res-i}$ , a partire dalle variabili relative alla storia abitativa presenti nel database dello studio API.

La storia abitativa è stata raccolta per un massimo di 4 abitazioni, partendo da quella dell’anno di riferimento e procedendo a ritroso nel tempo per un massimo di 40 anni. Vi sono quindi un massimo di 4 periodi di tempo contraddistinti da una diversa abitazione ( $periodo_i, i=1,2,3,4$ ). Poichè essi coprono un intervallo di tempo generalmente più esteso della finestra temporale, il tempo di esposizione  $t_{res-i}$  non corrisponde sempre al  $periodo_i$ , ma deve essere calcolato tenendo conto della particolare struttura dei dati di partenza (Figura 1A). Il calcolo di  $\mu_{res-i}$  è semplice e richiede solo di conoscere quali abitazioni appartengono all’area in studio e la loro distanza dal centroide dell’impianto.

**Figura 1A. Rappresentazione grafica della compilazione della storia abitativa per un ipotetico soggetto.**



**Figura 1B. Attribuzione di un asse dei tempi continuo ai dati di Figura 1A**



#### 1.1. Attribuzione di una data di inizio e di una data di fine in corrispondenza di ogni periodo compilato.

Per calcolare i tempi di esposizione sono state ricostruite una data di inizio e una data di fine periodo a partire dall’anno fornito dalle variabili che rappresentano inizio e fine periodo. La storia abitativa è stata raccolta a partire dall’anno di riferimento a ritroso nel tempo, quindi la data di fine del Periodo1 è data da 31.12.anno di riferimento. Per attribuire le altre date, si è assunto che ogni cambio di abitazione si sia verificato il primo giorno dell’anno indicato dall’intervistato come inizio di periodo. Quindi per i dati di Figura 1A, al Periodo1 = 1991 – 2000 si attribuiscono le date 01.01.1991 e 31.12.2000 come inizio e fine periodo, al Periodo2 = 1978-1990 si attribuiscono le date 01.01.1978 e 31.12.1990, e così via in modo da avere una rappresentazione del tempo continua nella scala dei giorni (Figura 1B). Questa assunzione è arbitraria ma non dipende dallo status caso/controllo e non comporta, nel calcolo dei tempi di esposizione, un aumento dell’incertezza legata alla forma dei dati di inizio e fine periodo.

### 1.2 Attribuzione dei limiti della finestra temporale e controllo di coerenza dei periodi.

Le date corrispondenti ai limiti della finestra temporale sono state così calcolate:

$$supFinest = 31.12.(anno di riferimento - 5)$$

$$infFinest = 01.01.(anno di riferimento - 19)$$

per i soggetti con età  $\geq 20$  anni e

$$supFinest = 31.12.(anno di riferimento - 2)$$

$$infFinest = data di nascita$$

Per i 3 soggetti con età  $< 20$  anni.

Per verificare la corrispondenza tra finestra temporale e intervallo di tempo entro il quale sono disponibili le informazioni sull'abitazione sono stati calcolati il minimo delle date di inizio periodo ( $DAmin = \min\{DAanno1, DAanno2, DAanno3, DAanno4\}$ ) e il massimo delle date di fine periodo ( $Amax = \max\{Aanno1, Aanno2, Aanno3, Aanno4\}$ ). Due soggetti sono risultati avere la finestra temporale non totalmente inclusa nel periodo compilato e un soggetto è risultato avere un salto di 5 anni interno alla finestra temporale. Quindi per questi 3 soggetti non è possibile calcolare l'esposizione cumulata. Non risultavano soggetti con sovrapposizione di periodi.

### 1.3 Attribuzione del primo periodo che include l'estremo superiore della finestra temporale e del numero di periodi inclusi nella finestra temporale

Il periodo compilato che include l'estremo superiore della finestra temporale (*perioinizio*) e il numero di periodi da essa coperti (*periodi*) sono calcolabili confrontando le date che delimitano la finestra temporale e le date di inizio e fine periodo. Ad esempio, se

$$(supFinest \geq DAanno1 \text{ AND } supFinest \leq Aanno1) \text{ AND } (infFinest \geq DAanno1 \text{ AND } infFinest \leq Aanno1)$$

$$\rightarrow \text{perioinizio} = 1 \text{ e } \text{periodi} = 1,$$

oppure se

$$(supFinest \geq DAanno3 \text{ AND } supFinest \leq Aanno3) \text{ AND } (infFinest \geq DAanno4 \text{ AND } infFinest < Aanno4)$$

$$\rightarrow \text{perioinizio} = 3 \text{ e } \text{periodi} = 2.$$

Per i dati di Figura 1, *perioinizio* = 1 e *periodi* = 2.

### 1.4 Calcolo dei tempi di esposizione residenziale $t_{res-i}$

Conoscendo il primo periodo incluso nella finestra temporale e il numero di periodi inclusi è possibile calcolare i tempi di esposizione, espressi in anni, come segue.

$$\text{Se } \text{perioinizio}=1 \text{ AND } \text{periodi}=1 \rightarrow t_{res-1} = (supRes - infRes)/365.25$$

$$t_{res-2} = t_{res-3} = t_{res-4} = 0.$$

$$\text{Se } \text{perioinizio}=1 \text{ AND } \text{periodi}=2 \rightarrow t_{res-1} = (supRes - DAanno1)/365.25$$

$$t_{res-2} = (Aanno2 - infRes)/365.25$$

$$t_{res-3} = t_{res-4} = 0$$

... ecc

$$\text{se } \text{perioinizio}=2 \text{ AND } \text{periodi}=1 \rightarrow t_{res-1} = 0$$

$$t_{res-2} = (supRes - infRes)/365.25$$

$$t_{res-3} = t_{res-4} = 0$$

e così via.

Per i dati di Figura 1 risulta  $t_{res-1} = 4.997$  anni,  $t_{res-2} = 9.999$  anni,  $t_{res-3} = t_{res-4} = 0$ .

Il tempo totale di esposizione è dato da  $Tres = t_{res-1} + t_{res-2} + t_{res-3} + t_{res-4}$  e risulta  $Tres = 15$  anni per tutti i soggetti con età  $\geq 20$  anni, e per i minori di 20 anni  $Tres = età - 1$ .

### 1.5 Calcolo del punteggio di esposizione residenziale per ogni abitazione

Il punteggio di esposizione per la singola abitazione è calcolato come

$$\mu_{res-i} = \begin{cases} \frac{1}{d_{res-i}} & \text{se abitazione}_i \in \text{area in studio} \\ 0 & \text{se (abitazione}_i \notin \text{area in studio) OR } (t_{res-i} = 0) \\ \text{missing} & \text{se (abitazione}_i \in \text{area in studio) AND } (d_{res-i} = \text{missing}) \end{cases}$$

cioè  $\mu_i = 1/d_i$  se l'abitazione  $i$  appartiene al comune di Chiaravalle o di Falconara Marittima o di Montemarciano;  $\mu_i = 0$  se l'abitazione  $i$  non appartiene a nessuno dei tre comuni. Inoltre, per evitare che distanze mancanti in periodi esterni alla finestra temporale, producano un'esposizione cumulata mancante, si è posto  $\mu_i = 0$  se  $t_{res-i} = 0$ . Un soggetto risulta avere  $d_{res-i} = \text{missing}$  e quindi la sua esposizione residenziale cumulata sarà missing.

#### 1.6 Calcolo dell'esposizione residenziale cumulata

Vi sono ora tutti gli elementi per calcolare l'esposizione residenziale cumulata come:

$$expRes\_cum = (\mu_1 * t_{res-1} + \mu_2 * t_{res-2} + \mu_3 * t_{res-3} + \mu_4 * t_{res-4}) / Tres.$$

### 4. Esposizione occupazionale cumulata.

Questo capitolo descrive la procedura utilizzata per costruire la variabile di esposizione occupazionale cumulata, calcolata come media pesata dei singoli punteggi di rischio occupazionale:

$$ExpOCCcum \text{ (numerica o categorica)} = \frac{\sum_i r_i \cdot t_{occ-i}}{\sum_i t_{occ-i}}.$$

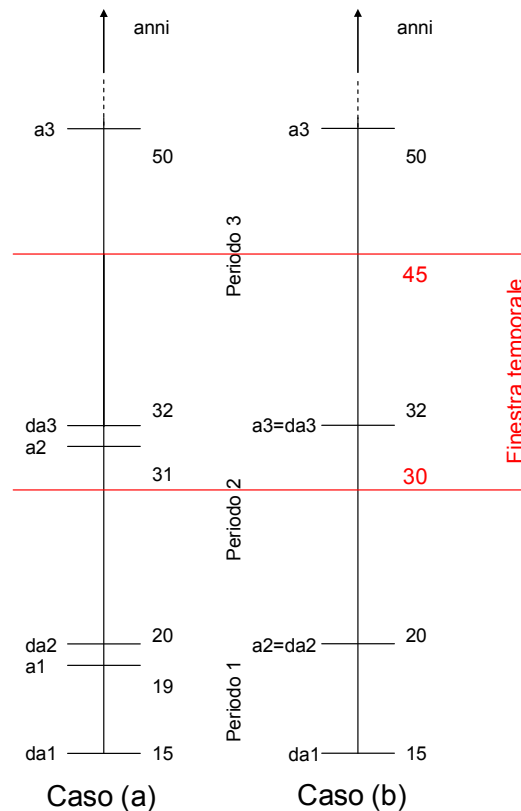
I pesi  $t_{occ-i}$  indicano il numero di anni impegnati nell'occupazione  $i$ , svolta per almeno 1 anno nei 20 anni precedenti l'anno di riferimento con esclusione degli ultimi 5. Il punteggio di esposizione della singola occupazione,  $r_i$ , è stato attribuito, a cura dell'ARPAM (dott. Mariottini) in due diversi modi. Nel primo caso,  $r_i = r_{DoRS-i} = 1$  se nella matrice MATline [1] l'occupazione  $i$  è associata ad una sostanza cancerogena per il sistema emolinfopoietico (benzene, vinil cloruro, etilene ossido, radiazioni x e gamma). Nel secondo caso  $r_i = r_{OCCAM-i} = 1$  se l'occupazione  $i$  corrisponde ad un comparto produttivo per il quale la matrice della letteratura OCCAM [2] riporta almeno 4 articoli. Per i soggetti che nel periodo  $i$ -esimo risultano avere status occupazionale di studente, casalinga, pensionato, o non occupato  $r_{DoRS-i} = r_{OCCAM-i} = 0$ .

Le informazioni sulla storia occupazionale sono state raccolte fino ad un massimo di 8 diverse occupazioni svolte per almeno 1 anno partendo dall'età di 15 anni fino all'età corrispondente all'anno di riferimento. Come per il calcolo dell'esposizione residenziale, si pone il problema di assegnare ad ogni soggetto i tempi di esposizione occupazionale  $t_{occ-i}$ , appartenenti alla stessa finestra temporale introdotta per l'esposizione residenziale, sulla base dei periodi compilati nel questionario. La struttura dei dati raccolti è simile ma non uguale per tutti i soggetti, la Figura 2 illustra i possibili modi di compilazione del questionario.

#### 4.1 Definizione di inizio e di fine periodo compilato e attribuzione dei limiti della finestra temporale.

Al fine di associare alle età di inizio e fine periodo un punto (e non un segmento) sull'asse dei tempi, si assume che ogni cambio di occupazione si verifichi esattamente il giorno del compleanno per tutti i soggetti. In questo modo per i dati di Figura 2, l'intervallo di tempo relativo all'occupazione  $i$  è quello intercorso dal giorno del 15° compleanno fino al giorno precedente il 20°

**Figura 2. Rappresentazione grafica della compilazione della storia occupazionale per due ipotetici soggetti.**



compleanno, ovvero 5 anni, ed è correttamente calcolato come  $periodo_1 = a_1 - da_1$  nel caso (b) e da  $periodo_1 = a_1 + 1 - da_1$  nel caso (a). L'intervallo di tempo relativo alla seconda occupazione è di 12 anni ed è correttamente calcolato come  $periodo_2 = a_2 - da_2$  nel caso (b) e da  $periodo_2 = a_2 + 1 - da_2$  nel caso (a) (dal giorno del 20° compleanno al giorno precedente il 32° compleanno). Per i soggetti con codifica (a), le età di fine periodo sono state ricalcolate come  $a_i = a_{i+1}$ . Infine, per i soggetti con età  $\geq 20$  anni, si assume che i limiti della finestra temporale definiti da

$$eta_{inizio} = et\grave{a} - 20$$

$$eta_{fine} = et\grave{a} - 5$$

corrispondano al giorno del compleanno dell'ultimo 20° anno e dell'ultimo 5° anno prima dell'anno di riferimento (per i dati di Figura 2,  $eta_{inizio} = 30$  e  $eta_{fine} = 45$ ). I minori di 20 anni sono considerati a parte e ad essi è stato attribuito un rischio occupazionale nullo.

#### 4.2 Controlli di coerenza dei periodi

Novi soggetti avevano lacune di pochi anni nella ricostruzione della storia occupazionale ma influenti nel calcolo dell'esposizione cumulata, perché esterne alla finestra temporale. Quattro soggetti risultavano avere occupazione mancante in periodi interni alla finestra temporale e, quindi, per essi non è calcolabile l'esposizione cumulata.

#### 4.3 Attribuzione del primo periodo incluso nella finestra temporale e del numero di periodi ad essa appartenenti.

L'individuazione del primo periodo incluso nella finestra temporale ( $perioinocc$ ) e del numero di periodi occupazionali compilati ad essa appartenenti ( $periodiocc$ ), è analoga a quella seguita per il calcolo dei tempi d'esposizione residenziale.

Se  $(eta_{inizio} \geq da_1 \text{ AND } eta_{inizio} < a_1) \text{ AND } (eta_{fine} \geq da_1 \text{ AND } eta_{fine} \leq a_1)$

→  $perioinocc = 1$ ,  $periodiocc = 1$

Se  $(eta_{inizio} \geq da_1 \text{ AND } eta_{inizio} < a_1) \text{ AND } (eta_{fine} \geq da_2 \text{ AND } eta_{fine} \leq a_2)$

→  $perioinocc = 1$ ,  $periodiocc = 2$

.....

Se  $(eta_{inizio} \geq da_7 \text{ AND } eta_{inizio} < a_7) \text{ AND } (eta_{fine} \geq da_8 \text{ AND } eta_{fine} \leq a_8)$

→  $perioinocc = 7$ ,  $periodiocc = 2$

Se  $(eta_{inizio} \geq da_8 \text{ AND } eta_{inizio} < a_8) \text{ AND } (eta_{fine} \geq da_8 \text{ AND } eta_{fine} \leq a_8)$

→  $perioinocc = 8$ ,  $periodiocc = 1$ .

Per i dati di Figura 2,  $perioinocc = 2$  e  $periodi = 2$ .

#### 4.4 Calcolo dei tempi di esposizione occupazionale

I tempi di esposizione occupazionale,  $t_{occ-i}$ , appartenenti alla finestra temporale sono ora calcolabili come:

se  $perioinocc = 1$  AND  $periodiocc = 1 \rightarrow t_{occ1} = etafine - etainizio$

$$t_{occ2} = t_{occ3} = \dots = t_{occ8} = 0$$

se  $perioinocc = 1$  AND  $periodiocc = 2 \rightarrow t_{occ1} = a1 - etainizio$

$$t_{occ2} = etafine - da2$$

$$t_{occ3} = t_{occ4} = \dots = t_{occ8} = 0$$

.....

se  $perioinocc = 7$  AND  $periodiocc = 2 \rightarrow t_{occ1} = t_{occ2} = \dots = t_{occ6} = 0$

$$t_{occ7} = a7 - etainizio$$

$$t_{occ8} = etafine - da8$$

se  $perioinocc = 8$  AND  $periodiocc = 1 \rightarrow t_{occ1} = t_{occ2} = \dots = t_{occ7} = 0$

$$t_{occ8} = etafine - etainizio .$$

$Tocc = t_{occ1} + t_{occ2} + \dots + t_{occ8}$  è il tempo totale di esposizione occupazionale incluso nella finestra temporale.

Per i dati di Figura 2,  $t_{occ1} = 0$ ,  $t_{occ2} = 2$ ,  $t_{occ3} = 13$ , e  $t_{occ4} = t_{occ5} = t_{occ6} = t_{occ7} = t_{occ8} = 0$ .

#### 4.5 Calcolo dell'esposizione occupazionale cumulata

E' ora possibile calcolare l'esposizione occupazionale cumulata come:

$$expDORS\_cum = (r_{DORS-1} * t_{occ1} + r_{DORS-2} * t_{occ2} + \dots + r_{DORS-8} * t_{occ8}) / Tocc$$

con punteggi  $r_{DORS-1}$  assegnati in base alla matrice MATline e

$$expOCCAM\_cum = (r_{OCCAM-1} * t_{occ1} + r_{OCCAM-2} * t_{occ2} + \dots + r_{OCCAM-8} * t_{occ8}) / Tocc$$

con punteggi  $r_{OCCAM-1}$  assegnati in base alla matrice OCCAM.

#### REFERENZE

- 1) DoRS, Documentazione Regionale Salute: <http://www.dors.it> .
- 2) Progetto OCCAM - Occupational Cancer Monitoring - Sistema informativo sui tumori di origine professionale: <http://www.occam.it> .