

# Considerazioni sui casi di tumore nell'Area della Ricerca di Firenze del Consiglio Nazionale delle Ricerche

Riccardo Benedetti (membro RSU), Manuela Corongiu (RLS e membro RSU)  
Consorzio Laboratorio di Monitoraggio e Modellistica Ambientale

Lorenzo Bigagli (RLS), Giovanni La Penna (RLS)  
Area della Ricerca di Firenze del Consiglio Nazionale delle Ricerche

10 dicembre 2015

## Sommario

Recentemente, tra i dipendenti dell'Area della Ricerca di Firenze del Consiglio Nazionale delle Ricerche (CNR) presso il Polo Scientifico di Sesto Fiorentino si è verificato un numero preoccupante di casi di tumore. L'Azienda Sanitaria di Firenze (ASF) e l'Istituto per lo Studio e la Prevenzione Oncologica (ISPO) di Firenze, con l'interessamento dell'Azienda Regionale per la Protezione Ambientale della Toscana, hanno indagato la questione, concludendo che il numero di casi verificatisi non è statisticamente superiore alla media territoriale e che non è stata riscontrata la presenza di fattori di rischio ambientale.

Questo documento evidenzia svariati aspetti critici, sia nel merito che nel metodo, dell'indagine condotta, con particolare riferimento all'approfondimento statistico.

In conclusione, si ritiene che l'intervento di ASF-ISPO non costituisca un'indagine accurata e che sia stato insufficiente e sbrigativo. Si raccomandano ulteriori approfondimenti e interventi di monitoraggio e prevenzione sul personale e sull'ambiente dell'Area CNR, del Polo Scientifico e della Piana in generale.

## 1 Introduzione

Nel seguito di questa sezione si introducono i fatti essenziali, le definizioni e i dati rilevanti. La sezione 2 descrive l'approfondimento statistico condotto dal gruppo di lavoro ASF-ISPO, discusso nella sezione 3. La sezione 3.1 affronta la questione quantitativamente, ricorrendo a moderni metodi di teoria della probabilità. Infine, la sezione 4 riporta le conclusioni degli autori.

### 1.1 Fatti essenziali

Il 9 agosto 2013, un dipendente dell'Area della Ricerca di Firenze del CNR segnala alle Rappresentanze Sindacali Unitarie (RSU) i numerosi casi di tumore riscontrati tra

i colleghi: 9 casi di tumore al seno e 1 caso di leucemia negli ultimi due anni.

Il 19 agosto 2013, le RSU trasmettono la segnalazione ai Rappresentanti dei Lavoratori per la Sicurezza (RLS).

Il 16 settembre 2013, in una riunione tra Responsabile del Servizio di Prevenzione e Protezione (RSPP), RLS e Medico Competente (MC) dell'Area CNR, si decide di contattare l'Istituto per lo Studio e la Prevenzione Oncologica (ISPO) di Firenze e l'Azienda Regionale per la Protezione Ambientale della Toscana (ARPAT), nonché di monitorare i lavoratori CNR da parte del MC.

Il 18 settembre 2013, l'ARPAT comunica di poter dare risposte solo dietro richiesta diretta di enti pubblici o locali.

Il 17 ottobre 2013, in un incontro di RLS, RSPP e

MC con responsabili dell'ISPO, si decide di contattare l'Azienda Sanitaria Locale (ASL) competente.

Il 29 novembre 2013, avviene l'incontro di RSPP e RLS con responsabili dell'Azienda Sanitaria di Firenze (ASF). L'ASF richiede i dati relativi a mansioni, sesso ed età dei dipendenti a tempo indeterminato dell'Area CNR.

Il 13 dicembre 2013, l'RSPP invia alla ASF i dati richiesti, per l'anno 2013 (v. tabella 1).

Nel gennaio 2014 RLS e RSU del CNR sollecitano svariate volte ASF e ARPAT. I RLS del CNR si rivolgono alla Camera del Lavoro di Sesto Fiorentino (CGIL), il cui segretario Marcello Corti chiede e ottiene un incontro congiunto con il Sindaco, Gianni Gianassi.

Il 4 febbraio 2014, quest'ultimo, in qualità di autorità sanitaria sul territorio, sollecita per scritto ASF a svolgere le dovute indagini. ASF e ARPAT rispondono al Sindaco e al presidente dell'Area CNR che i dati sono in corso di elaborazione da parte della Unità Operativa di Epidemiologia.

Il 17 marzo 2014, in una riunione tra RLS e ASF, si decide di far preparare, a cura del RSPP, le planimetrie dell'Area CNR con la dislocazione precisa della grossa strumentazione e dei locali rifiuti rispetto agli uffici in cui lavorano i dipendenti interessati da patologie. I dati sono inviati al Dipartimento della Prevenzione ASF.

Il 16 aprile 2014, RSU e RLS decidono di informare i lavoratori nel corso di un'assemblea del personale circa tutto l'iter della segnalazione.

Il 16 giugno 2014, si tiene un primo incontro coordinato dal Dipartimento della Prevenzione della ASF, a cui partecipano RSPP, RSU, RLS dell'Area CNR. Sono riportati i risultati delle elaborazioni dei dati e il percorso di monitoraggio che sarà intrapreso di conseguenza.

Il 19 settembre 2014, si tiene un secondo incontro, presso il Dipartimento della Prevenzione ASF.

Nell'autunno 2014 il gruppo di lavoro ASF-ISPO svolge un'indagine epidemiologica comprensiva della somministrazione di un questionario ad alcune dipendenti, di un approfondimento statistico (v. sezione 2), di un approfondimento ambientale outdoor (richiesto ad ARPAT) e di misure di campionamento di alcuni inquinanti indoor.

Il 20 ottobre 2014, si svolge un incontro con tutto il personale CNR, in cui sono presentati i risultati preliminari dell'indagine. Nonostante la convocazione, ARPAT non si presenta all'incontro. Il Direttore del Dipartimento di Prevenzione ASL 10, Dott. Giuseppe Petrioli, comunica

che è stata riconosciuta l'esistenza di un cluster, ossia un numero di casi superiore alle attese, statisticamente significativo al 99%, di tumori al seno nelle dipendenti CNR dell'Area, dovuto a cause al momento ignote. La Dott.ssa Lucia Miligi, dell'Unità Operativa di Epidemiologia dell'ISPO, mostra una lettera di ARPAT ad ASF dell'aprile 2014, in cui ARPAT dichiara che, sulla base dei propri dati, non ci sono contaminazioni ambientali nella zona. Quindi riporta che è stato condotto uno studio di coorte sui tumori alla mammella, considerando un campione di 90 dipendenti CNR su un periodo di 10 anni, periodo medio di latenza della patologia. Proiettando l'incidenza media del registro dei tumori per la zona Prato-Firenze, si attendevano 1,4 casi di tumore, invece dei 6 riscontrati nel campione.

Il 23-25 ottobre 2014, la stampa locale riprende la notizia con un certo rilievo.

Tra novembre e dicembre 2014, circa 300 afferenti al Polo Scientifico (di cui circa 250 al CNR) firmano una petizione rivolta ai direttori dell'Area CNR affinché promuovano una campagna di misure ambientali e un'indagine sanitaria/epidemiologica, in coordinamento con l'Università di Firenze. Analoghe petizioni circolano presso altre istituzioni del Polo Scientifico.

Il 17 novembre 2014, vengono eseguite le misurazioni degli inquinanti indoor, installando per 24 ore una trentina di campionatori passivi all'interno e all'esterno della struttura CNR. Inoltre sono eseguite alcune misure del campo elettromagnetico al piano terra dell'edificio F.

Il 21 gennaio 2015, dopo varie sollecitazioni da parte di personale del Polo, il gruppo di lavoro ASF-ISPO rilascia una relazione finale, in cui si conclude che: *"l'aggregazione di casi che si è verificato [sic] al CNR non è statisticamente superiore all'atteso e gli approfondimenti condotti non hanno documentato la presenza di fattori di rischio insiti nel luogo di lavoro che lo possano spiegare."* [1, p. 2]

L'11 febbraio 2015, il Laboratorio di Monitoraggio e Modellistica Ambientale (LaMMA) invia all'ASF alcune osservazioni in merito a dubbi sollevati dalla relazione, in particolare riguardo alla metodologia utilizzata.

Il 2 marzo 2015, la ASF risponde con alcune integrazioni alla relazione, che tuttavia non chiariscono completamente i dubbi sollevati.

Il 15 giugno 2015, durante un'assemblea del personale CNR a cui partecipa anche personale del LaMMA, si

discutono le criticità della relazione ASF-ISPO e i risultati di un'analisi probabilistica dell'incidenza del carcinoma mammario sulle dipendenti CNR, che risulterebbero soggette a un rischio più alto rispetto alla media della popolazione di Firenze e Prato, con una probabilità intorno al 95%. Si dà mandato ai nuovi RLS di sollecitare ulteriori approfondimenti e interventi di monitoraggio e prevenzione.

## 1.2 Definizioni

**Tasso d'incidenza** (*incidence rate*, indicato con  $\mu$ ) è la frazione di nuovi casi  $m$  che insorgono nel tempo  $t$  in una popolazione di  $n$  individui inizialmente sani:

$$\mu \stackrel{\text{def}}{=} \frac{m}{nt} \quad (1)$$

Tipicamente  $\mu$  si esprime in numero di nuovi casi su centomila persone all'anno.

**Tasso Standardizzato Indiretto** (*Standardized Incidence Ratio*, indicato con SIR) è il rapporto percentuale tra il tasso d'incidenza  $\mu$  di un certo gruppo e quello  $\mu_0$  relativo a una popolazione di riferimento:

$$\text{SIR} \stackrel{\text{def}}{=} 100 \frac{\mu}{\mu_0} \quad (2)$$

Un valore di SIR pari a 100 indica che non c'è differenza tra il tasso d'incidenza osservato nel gruppo e quello di riferimento; un valore inferiore a 100 indica una minore incidenza della patologia nel gruppo; uno superiore a 100 ne indica una maggiore.

**Casi attesi** (indicato con  $\bar{m}$ ) è il numero di casi di cui si prevede l'insorgenza in un gruppo di  $\bar{n}$  individui durante un periodo di tempo  $\bar{t}$ . Si può stimare dal tasso d'incidenza o dal SIR:

$$\bar{m} = \mu \bar{n} \bar{t} = \frac{\text{SIR}}{100} \mu_0 \bar{n} \bar{t} = \frac{\text{SIR}}{100} \bar{m}_0 \quad (3)$$

Se il gruppo è composto da  $k$  classi distinte, con popolazioni  $\bar{n}_1, \bar{n}_2, \dots, \bar{n}_k$  e tassi d'incidenza  $\mu_1, \mu_2, \dots, \mu_k$ , costanti nel periodo di tempo  $\bar{t}$ , i casi attesi sono dati dalla somma dei casi attesi per ciascuna classe:

$$\bar{m} = \sum_{i=1}^k \mu_i \bar{n}_i \bar{t} \quad (4)$$

## 1.3 Dati

Come accertato dal MC, tra le circa 120 dipendenti CNR si sono verificati 8 casi di tumore al seno, tutti diagnosticati tra marzo 2012 e agosto 2014, di cui:

- 6 casi nell'Area CNR di Sesto Fiorentino (3 casi nel 2012, 2 nel 2013, 1 nel 2014);
- 2 casi presso la sede di Via Caproni, a Peretola, dell'istituto IBIMET, a circa 1.500 metri dall'Area CNR di Sesto Fiorentino.

Dei 6 casi nell'Area CNR, 3 hanno colpito ricercatrici e 3 personale amministrativo, con età media alla diagnosi di 50 anni (range 42-56)<sup>1</sup> e con la seguente localizzazione:

- 4 casi (2 ricercatrici, 2 amministrative) nell'edificio E;
- 1 caso (ricercatrice) nell'edificio D;
- 1 caso (amministrativa) nell'edificio C.

La tabella 1 riporta i tassi d'incidenza età-specifici annuali del tumore alla mammella, per centomila persone, nell'area di Firenze e Prato:<sup>2</sup>

Tabella 1: Tassi d'incidenza.

Fascia d'età	$\mu_0$
20-69 anni	157,1
20-29 anni	4,8
30-39 anni	49,9
40-49 anni	202,4
50-59 anni	279,6
60-69 anni	313,7

La figura 1 riporta la distribuzione dei dipendenti CNR per l'anno 2013.

<sup>1</sup>Dalla relazione ASF-ISPO, i 6 casi risulterebbero: 3 nella fascia d'età 40-49 anni; 2 nella fascia d'età 50-59 anni; 1 nella fascia d'età 60-69 anni. Si noti che questa distribuzione non sembra compatibile con gli altri dati. Comunque questo non inficia i risultati e le conclusioni che seguono.

<sup>2</sup>Secondo quanto dichiarato da ISPO, i tassi in tabella sono quelli del Registro dei Tumori della Regione Toscana per Firenze Prato nel periodo 2000-2005. [2] Da una nostra verifica, dal Registro risulterebbero valori diversi per le fasce d'età: 40-49 anni ( $\mu = 202,45$ ); 50-59 anni ( $\mu = 279,45$ ); 60-69 anni ( $\mu = 315,1$ ); e soprattutto per la fascia complessiva 20-69 anni ( $\mu = 170,34$ ). Comunque queste differenze non inficiano i risultati e le conclusioni che seguono.

Anno 2013		Lavoratori del CNR					
Classe età	Ricerca		Tecnica		Amministrazione		Totale Femmine per classe età
	M	F	M	F	M	F	
21-30 anni	0	0	1	0	0	0	1
31-40 anni	18	14	9	0	3	3	17
41-50 anni	50	21	14	5	4	15	41
51-60 anni	42	21	16	6	1	9	36
61-70 anni	11	4	3	0	0	2	6
<b>TOTALE</b>	<b>121</b>	<b>60</b>	<b>43</b>	<b>11</b>	<b>8</b>	<b>30</b>	<b>101</b>

Figura 1: Dipendenti dell'Area CNR nel 2013, divisi per sesso, settore d'impiego e classe d'età.

## 2 Approfondimento statistico di ASF-ISPO

L'analisi epidemiologica effettuata dal gruppo di lavoro ASF-ISPO si basa sulla stima del SIR attraverso il metodo statistico della *massima verosimiglianza*. [3]

Come criterio di significatività della stima, il gruppo di lavoro ASF-ISPO determina un certo intervallo di confidenza e verifica se tale intervallo include il valore 100. In tal caso, il SIR è *“convenzionalmente considerato come “statisticamente non significativo”: non è cioè possibile affermare con sufficiente sicurezza che i due valori, quello in esame e quello della popolazione di riferimento, siano realmente diversi; l'eccesso o il difetto di incidenza osservati potrebbero essere solo un effetto del caso”*. [1, nota 3, p. 18]

L'analisi del gruppo di lavoro ASF-ISPO procede come segue:

1. Individuazione della popolazione campione. L'analisi si limita alla popolazione dell'Area CNR di Sesto Fiorentino, escludendo i casi di tumore verificatisi presso la sede IBIMET a Peretola. Sono considerati due gruppi, tra i dipendenti al 2013 (v. tabella 1):
  - A. Le 90 dipendenti donne del settore ricerca e amministrazione, dove si sono registrati i sei casi di tumore;
  - B. Tutte le 101 dipendenti donne dell'Area CNR, includendo anche il settore tecnico.

2. Determinazione del numero  $\bar{m}_A, \bar{m}_B$  di casi attesi per i due gruppi. Applicando la formula (4) per proiettare i tassi d'incidenza età specifici del registro dei tumori per la zona Firenze Prato (v. tabella 1) su un periodo di 12 anni (2003-2014), ossia dall'inizio delle attività dell'Area CNR di Sesto Fiorentino, si ottiene:

$$\bar{m}_A = 2,209164 \quad (5)$$

$$\bar{m}_B = 2,531916 \quad (6)$$

3. Definizione di uno stimatore per i casi attesi  $\bar{m}$  sulle dipendenti dell'Area CNR. Assumendo che il numero di casi osservabili  $m$ , dato il numero atteso  $\bar{m}$ , segua una distribuzione di Poisson:

$$P(m|\bar{m}) = \frac{\bar{m}^m}{m!} e^{-\bar{m}} \quad (7)$$

La relativa funzione di verosimiglianza per  $\bar{m}$  risulta: [3]

$$L(\bar{m}; m = 6) \propto \bar{m}^6 e^{-\bar{m}} \quad (8)$$

Il massimo della funzione (8) si ha per  $\bar{m} = 6$ , per cui il valore più verosimile del SIR è dato rispettivamente da:

$$SIR_A = 271,60 \quad (9)$$

$$SIR_B = 236,97 \quad (10)$$

Si noti che il tasso di incidenza su entrambi i gruppi risulta molto più elevato (ben più che doppio) di quello sulla popolazione di riferimento.

4. Determinazione dell'intervallo di confidenza  $[IC_{\text{inf}}; IC_{\text{sup}}]$  al 95% di significatività per  $\bar{m}$ . IC è l'insieme di valori entro cui si colloca, con una probabilità del 95%, il valore reale di  $\bar{m}$ . Gli estremi sono approssimabili ponendo  $m = 6$  nelle seguenti espressioni (vedi [4]):

$$IC_{\text{inf}} \cong m - 1,96\sqrt{m} + 1 \cong 2,20 \quad (11)$$

$$IC_{\text{sup}} \cong m + 1,96\sqrt{m} + 2 \cong 12,8 \quad (12)$$

La figura 2 rappresenta il grafico della funzione di massima verosimiglianza e l'intervallo di confidenza al 95% per i casi attesi.

5. Calcolo dell'intervallo di confidenza al 95% del SIR. Dalle formule (3), (5), (6), (11) e (12), si ricava:

$$IC_{SIR_A} = \left[ \frac{IC_{\text{inf}}}{\bar{m}_A} \frac{IC_{\text{sup}}}{\bar{m}_A} \right] = [99,54 \quad 579,45] \quad (13)$$

$$IC_{SIR_B} = \left[ \frac{IC_{\text{inf}}}{\bar{m}_B} \frac{IC_{\text{sup}}}{\bar{m}_B} \right] = [86,85 \quad 505,59] \quad (14)$$

Poiché gli intervalli di confidenza del SIR includono il valore 100, il gruppo ASF-ISPO conclude che i dati non sono significativi di un tasso d'incidenza del tumore al seno superiore alla media per le dipendenti dell'Area CNR di Sesto Fiorentino.

### 3 Discussione

L'approfondimento statistico effettuato dal gruppo di lavoro ASF-ISPO, dettagliato nella sezione 2, presenta svariati aspetti critici, sia nel merito che nel metodo.

Nel merito, i risultati numerici dell'analisi presentano margini molto esigui, rispetto a limiti di confidenza molto ampi, che riflettono una popolazione numericamente limitata. Ad esempio, per il campione A, l'intervallo di confidenza del SIR (13) include il valore 100 per un'inezia: 0,46, su un'estensione dello stesso intervallo oltre

1.000 volte superiore. Per giunta, l'espressione (11) risulta essere un'approssimazione per difetto: lo scarto esatto è ancora minore ( $< 0,33$ ).<sup>3</sup>

Questo suggerisce una probabile elevata sensibilità del risultato alle variabili del problema. In effetti, le conclusioni si ribaltano considerando un periodo d'osservazione (follow up) di 10 anni (invece di 12), com'era stato inizialmente considerato e riportato nell'incontro del 20 ottobre con il personale CNR.<sup>4</sup>

La giustificazione addotta per l'estensione del follow up a 12 anni è "che il CNR di Sesto Fiorentino è attivo dal 2003 (e i casi si riferiscono a quanti diagnosticati nel periodo 2003-2014)". [1, p. 19] In realtà, tutti i sei casi si sono concentrati tra il marzo del 2012 e l'agosto del 2014, circostanza che appare significativa, ma che tuttavia non è rilevata nell'analisi ASF-ISPO. Inoltre, la gran parte del personale dell'Area CNR vi si è trasferita anche molto dopo il 2003 (l'Area stessa è stata completata nel giugno del 2005) e "la durata media al lavoro in anni presso la sede di Sesto Fiorentino è 8,8 anni per le lavoratrici che si sono ammalate". [1, p. 24]<sup>5</sup> In sostanza, l'estensione del follow up a 12 anni appare una scelta arbitraria e non sufficientemente motivata, soprattutto considerata la sensibilità delle conclusioni rispetto a tale variabile.

Per il calcolo dei casi attesi, sono state utilizzate in (4) le classi d'età delle dipendenti al 2013 (v. figura 1), considerandone le popolazioni come costanti nel periodo. Questo è certamente scorretto, in quanto, banalmente, è impossibile che un individuo permanga per 12 anni in una fascia d'età di dieci! Inoltre, basandosi sull'età delle dipendenti al 2013, quasi al termine del periodo d'osservazione, tale ripartizione introduce un bias verso le fasce d'età più avanzate, caratterizzate da tassi d'incidenza maggiori (v. tabella 1). In sostanza, il modello usato per il calcolo del numero dei casi attesi  $\bar{m}$  appare un'approssimazione grossolana.

Un'approssimazione migliore consiste nel considerare che, nell'arco del periodo di osservazione, tutti gli indivi-

<sup>3</sup>Si veda per esempio <http://statpages.org/confint.html> o <http://www.danielsoper.com/statcalc3/calc.aspx?id=86>

<sup>4</sup>Inoltre l'analisi preliminare presentata il 20 ottobre era limitata al campione A e a un'unica fascia d'età (20-69 anni). Il campione B, più ampio, e i tassi d'incidenza età-specifici sono stati introdotti nella relazione finale.

<sup>5</sup>Si noti l'incongruenza con pagina 17: "tutte lavorano dal 2003 al polo scientifico".

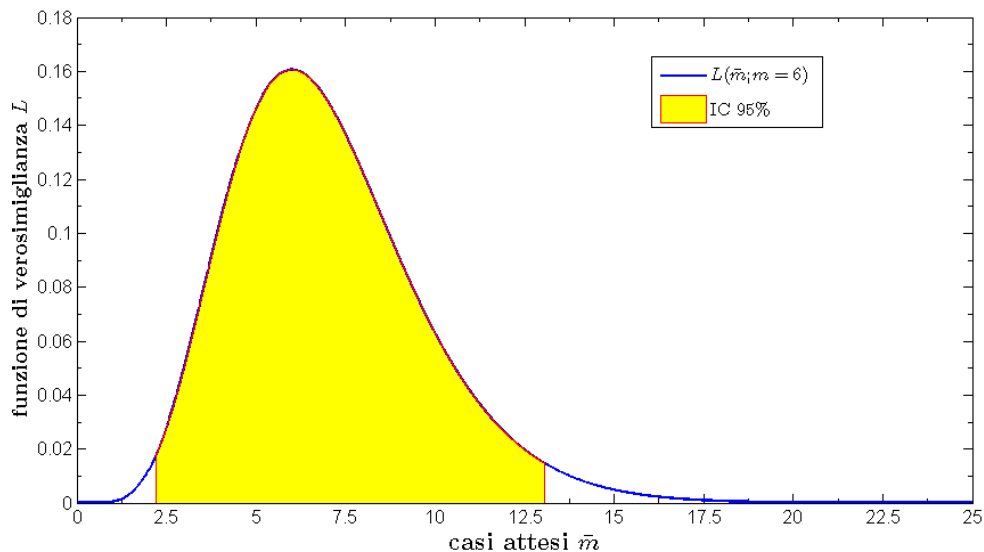


Figura 2: Funzione di massima verosimiglianza per i casi attesi.

dui sono ricaduti in due classi di età. Si nota che è sufficiente attribuire il 28% degli anni-persona a una classe di età inferiore affinché le conclusioni dell'analisi si ribaltino.

Un'approssimazione ancora migliore consisterebbe nell'utilizzo degli effettivi intervalli di permanenza individuale nelle varie classi, basandosi sulle date di ingresso-uscita delle lavoratrici nella struttura. Il gruppo di lavoro ASF-ISPO ammette il limite di non aver considerato "l'informazione completa su tutta la popolazione esposta (chi ha lavorato nel CNR, data inizio, data fine)", [1, p. 18] adducendo che i dati non siano stati forniti dalla sede centrale del CNR. Tuttavia la mancanza di questi dati avrebbe dovuto suggerire di sospendere l'analisi, piuttosto che produrre un'analisi fuorviante.

In definitiva, per quanto premesso e poiché il criterio di significatività adottato si basa su valori di soglia arbitrari (o "convenzionali", per dirla con il gruppo di lavoro ASF-ISPO), le conclusioni tratte appaiono aleatorie e non supportate in modo solido dai dati.

Perplessità ancora maggiori sorgono dal punto di vista metodologico. Di fatto, quello che si è cercato di verificare è se il numero di tumori riscontrato possa essere anche

minimamente compatibile con la media territoriale. Usando metodi statistici, tassi indiretti, verosimiglianze, soglie convenzionali e approssimazioni discutibili, si è concluso che lo è, seppure per un'inezia.

Ora, non si vede come questo possa rispondere e tranquillizzare rispetto alla vera domanda, a cui non si è dato risposta: le dipendenti dell'Area della Ricerca di Firenze del CNR sono esposte a rischi di tumore al seno maggiori del resto della popolazione?

Si poteva affrontare questa questione con lo stesso metodo e lo stesso criterio qualitativo di significatività, stimando il SIR e calcolandone l'intervallo di confidenza al 95%, ma *senza escludere la coda di destra della funzione di distribuzione*. Infatti, SIR molto elevati, ancorché improbabili, sarebbero a maggior ragione significativi e preoccupanti. Si intuisce facilmente dalla figura 2 che anche con questo approccio le conclusioni si sarebbero ribaltate.

Ancor meglio, si può affrontare la questione quantitativamente, ricorrendo a moderni metodi di teoria della probabilità, ovvero semplicemente valutando la probabilità che il valore del SIR per le dipendenti CNR sia superiore al valore atteso e "normale" di 100, come dettagliato nel seguito.

### 3.1 Analisi probabilistica

Definendo l'ipotesi:

$\mathcal{H}$  = "il tasso d'incidenza  $\mu$  generato dai fattori di rischio a cui sono esposte le dipendenti CNR è maggiore di quello  $\mu_0$  della popolazione con analoga distribuzione d'età nell'area di Firenze e Prato",

si tratta di calcolare la probabilità che  $\mathcal{H}$  sia vera, dato quanto osservato (gli  $m = 6$  casi di tumore sulle  $\bar{n}$  dipendenti CNR nel tempo  $\bar{t}$ ) e noto (i tassi d'incidenza età-specifici  $\mu_0$  sulla popolazione di Firenze e Prato).

Per la regola di marginalizzazione, applicata ai possibili valori di  $\mu$ , e per quella del prodotto, si può scrivere:

$$P(\mathcal{H}|m \bar{n} \bar{t} \mu_0) = \int_0^{+\infty} P(\mathcal{H}|\mu m \bar{n} \bar{t} \mu_0) p(\mu|m \bar{n} \bar{t} \mu_0) d\mu \quad (15)$$

Osservando che:

$$P(\mathcal{H}|\mu m \bar{n} \bar{t} \mu_0) = \begin{cases} 0 & \text{se } \mu \leq \mu_0, \\ 1 & \text{se } \mu > \mu_0. \end{cases} \quad (16)$$

Si può riscrivere:

$$P(\mathcal{H}|m \bar{n} \bar{t} \mu_0) = \int_{\mu_0}^{+\infty} p(\mu|m \bar{n} \bar{t} \mu_0) d\mu \quad (17)$$

Per il teorema di Bayes, la distribuzione di probabilità integranda può essere espressa come:

$$p(\mu|m \bar{n} \bar{t} \mu_0) = \frac{1}{C} p(\mu|\bar{n} \bar{t} \mu_0) P(m|\mu \bar{n} \bar{t} \mu_0) \quad (18)$$

dove  $C$  è una costante di normalizzazione.

Assumendo che l'insorgenza della patologia sia un processo di Poisson, da (3) e (7) si può ricavare:

$$P(m|\mu \bar{n} \bar{t} \mu_0) = P(m|\bar{m}) = \frac{\bar{m}^m}{m!} e^{-\bar{m}} \quad (19)$$

Si può inoltre assumere come *distribuzione a priori* per il tasso d'incidenza  $\mu$  quella che massimizza l'entropia, rispettando il vincolo di valor medio pari a  $\mu_0$ :

$$p(\mu|\bar{n} \bar{t} \mu_0) = p(\mu|\mu_0) = \frac{1}{\mu_0} e^{-\frac{\mu}{\mu_0}} \quad (20)$$

Sostituendo (18)-(20) in (17), si giunge a:

$$P(\mathcal{H}|m \bar{n} \bar{t} \mu_0) = \frac{1}{C} \int_{\mu_0}^{+\infty} \frac{1}{\mu_0} e^{-\frac{\mu}{\mu_0}} \frac{\bar{m}^m}{m!} e^{-\bar{m}} d\mu \quad (21)$$

Sostituendo la variabile d'integrazione con (3) e raccogliendo le costanti, si ha:

$$\begin{aligned} P(\mathcal{H}|m \bar{n} \bar{t} \mu_0) &= \frac{1}{C} \int_{\bar{m}_0}^{+\infty} \frac{1}{\bar{m}_0} e^{-\frac{\bar{m}}{\bar{m}_0}} \frac{\bar{m}^m}{m!} e^{-\bar{m}} d\bar{m} \\ &= \frac{1}{C'} \int_{\bar{m}_0}^{+\infty} \bar{m}^m e^{-\bar{m} \frac{1+\bar{m}_0}{\bar{m}_0}} d\bar{m} \end{aligned} \quad (22)$$

Infine, sostituendo l'argomento dell'esponenziale con  $t$ , l'integrale in (22) è riconducibile alla funzione gamma incompleta:

$$\Gamma(a, x) = \int_x^{+\infty} t^{a-1} e^{-t} dt \quad (23)$$

da cui si ottiene:

$$\begin{aligned} P(\mathcal{H}|m \bar{n} \bar{t} \mu_0) &= \frac{1}{C''} \int_{\bar{m}_0+1}^{+\infty} t^m e^{-t} dt \\ &= \frac{1}{C''} \Gamma(m+1, \bar{m}_0+1) \end{aligned} \quad (24)$$

I ragionamenti di cui sopra possono essere ripetuti per l'ipotesi complementare  $\bar{\mathcal{H}}$ , invertendo la (16), da cui:

$$\begin{aligned} P(\bar{\mathcal{H}}|m \bar{n} \bar{t} \mu_0) &= \frac{1}{C'''} \int_0^{\bar{m}_0+1} t^m e^{-t} dt \end{aligned} \quad (25)$$

Osservando che  $P(\mathcal{H}|\cdot) + P(\bar{\mathcal{H}}|\cdot) = 1$ , si può ricavare la costante di normalizzazione:

$$C''' = \int_0^{+\infty} t^m e^{-t} dt = \Gamma(m+1, 0) \quad (26)$$

In definitiva si ha:

$$P(\mathcal{H}|m \bar{n} \bar{t} \mu_0) = \frac{\Gamma(m+1, \bar{m}_0+1)}{\Gamma(m+1, 0)} \quad (27)$$

Poiché, per  $m$  intero, vale:

$$\Gamma(m+1, x) = m! e^{-x} \sum_{k=0}^m \frac{x^k}{k!} \quad (28)$$

si ottiene infine:

$$P(\mathcal{H}|m \bar{n} \bar{t} \mu_0) = e^{-(\bar{m}_0+1)} \sum_{k=0}^m \frac{(\bar{m}_0+1)^k}{k!} \quad (29)$$

Computando la formula con i casi attesi stimati in (5) e (6) (che, come discusso, restano stime inaffidabili), si può concludere che la probabilità che le dipendenti CNR siano esposte a fattori di rischio per il carcinoma mammario superiori a quelli della popolazione generale dell'area di Firenze e Prato è intorno al 95% per il campione A e al 93% per il campione B.

## 4 Conclusioni

Nell'incontro del 20 ottobre, tutti i relatori hanno concordato nel definire “*statisticamente significativo*” il numero di casi di tumore al seno registrati tra le dipendenti CNR. Nella relazione di approfondimento è invece ribadito più volte che tale concentrazione di tumori non è statisticamente significativa.

Un ribaltamento così netto è certamente possibile in linea di principio, ma è altrettanto certo che le motivazioni di conclusioni tanto difformi da quelle iniziali dovrebbero essere chiare e solidamente fondate. Invece, le argomentazioni del gruppo di lavoro ASF-ISPO non sembrano essere sufficientemente supportate dai dati. Anzi, considerate le approssimazioni fatte, l'approfondimento statistico sembrerebbe confermare le affermazioni iniziali.

Il gruppo di lavoro ASF-ISPO estende la stessa analisi ai dipendenti dell'Università di Firenze afferenti al Polo Scientifico. A quanto riportato, il SIR di questo campione più ampio non ha evidenziato anomalie rispetto alla media. Non abbiamo potuto verificare i calcoli, in particolare rispetto al bias discusso nella sezione 3. Comunque osserviamo che, per definizione, un cluster scompare allargando il campione di popolazione considerato.

Valgono inoltre le considerazioni fatte in merito alle criticità metodologiche: la stima della massima verosimiglianza e degli intervalli di confidenza è un metodo logicamente inconsistente e ormai superato, [5] che non può

rispondere al vero interrogativo d'interesse, ossia se e in che misura chi lavora al Polo Scientifico di Sesto Fiorentino sia esposto a un maggior rischio di tumore rispetto al resto della Piana di Prato e Firenze.

Il gruppo di lavoro ASF-ISPO premette che il metodo statistico applicato è solo “*l'approccio che si usa con maggiore frequenza per una prima analisi dei fenomeni di cluster*”. [1, p. 18] Anche gli altri approfondimenti condotti non paiono costituire un'indagine accurata, a cominciare dal questionario epidemiologico, somministrato a cinque delle sei dipendenti affette da tumore, per finire con le misure dirette di tipo chimico e fisico, alla ricerca di eventuali agenti inquinanti indoor.

Le misure di tipo chimico hanno indagato l'eventuale presenza di composti organici volatili (COV), quali benzene, toluene, etilbenzene e xileni. I campionamenti in vari locali dell'Area CNR hanno dato esito negativo, ma si sono limitati a un intervallo temporale di 24 ore, nello stesso giorno (17 novembre 2014). Ciò appare adeguato per rilevare la presenza di una sorgente inquinante a emissione continua, ma non per una con caratteristiche di discontinuità, forse anche più plausibile in un'area dove si svolgono attività di ricerca e non di produzione industriale.

Le misure di tipo fisico hanno riguardato eventuali campi elettromagnetici a bassa frequenza (50Hz), mentre le radiofrequenze sono state stimate da ARPAT con modelli. Le stime di modello risultano in effetti attendibili per sorgenti ben caratterizzate, quali le stazioni radio base (SRB) della telefonia mobile, che infatti sono le uniche che compaiono nella relazione di ARPAT. [1, all. D] ARPAT non fa invece menzione alcuna né degli impianti radio di Monte Morello, che verosimilmente forniscono un contributo al campo totale non inferiore a quello delle SRB, né soprattutto del radar del vicino aeroporto. Sarebbe auspicabile che quest'ultimo fosse oggetto di controlli specifici, a quanto pare mai effettuati, o comunque non menzionati nella relazione.

ASF-ISPO ha richiesto ad ARPAT informazioni su altre possibili esposizioni associate, con diversi gradi di evidenza, al tumore alla mammella, quali policlorobifenili (PCB), pesticidi, COV, clorurati, diossina e idrocarburi policiclici aromatici (IPA). La relazione riporta la risposta di ARPAT: “*Non ci sono sorgenti significative di inquinanti chimici né siti contaminanti interferenti con le*



matrici ambientali e risalenti ad attività pregresse”.<sup>6</sup> [1, p. 16]

Infine, eventuali inquinanti dell’aria, provenienti dalla vicina discarica e dall’ancor più vicino aeroporto, sono stati presi in considerazione solo attraverso uno studio di letteratura, che certamente può costituire un punto di partenza per l’analisi, ma ben difficilmente consente di trarre conclusioni plausibili riferibili a un caso specifico.

A questo proposito, si osserva che gli aeroporti sono annoverati come fonti di COV e pesticidi (gli insetti attirano gli uccelli, aumentando il rischio di bird-strike). Da notare che la relazione ASF-ISPO menziona uno studio sull’area dell’aeroporto di Amsterdam-Schiphol dal quale era emerso “*un SIR per i tumori della mammella di 105 (IC95%: 0,99-1,11), che era significativo nel quinquennio 1994-98 (SIR: 109, IC95%: 1,01-1,18)*”. [1, all. F] Seppur gli autori abbiano interpretato questo risultato come “*compatibile con la variazione regionale osservata in Olanda*”, i ben più elevati (e magari anche più significativi) valori di SIR osservati nell’Area CNR consiglierebbero forse qualcosa di più specifico e mirato che non un generico studio di letteratura.

In conclusione, si ritiene che l’intervento di ASF-ISPO per investigare il preoccupante numero di casi di tumore verificatisi negli ultimi anni tra i dipendenti dell’Area CNR di Sesto Fiorentino sia stato insufficiente e troppo sbrigativo.

Al di là dei generici auspici finali di ASF-ISPO<sup>7</sup>, si ritiene che i dati presentati dovrebbero indurre ad attuare un effettivo e realmente efficace monitoraggio dei dipendenti dell’Area CNR, includendo anche gli uomini e tut-

<sup>6</sup>Si noti che, il 12 aprile 2006, con nota assunta al Prot. Gen. con n. 21385, ARPAT ha segnalato all’Ufficio Ambiente del Comune di Sesto Fiorentino la presenza di rifiuti su di un terreno a meno di 10 metri dall’Area CNR. Il Comune di Sesto Fiorentino ha di conseguenza emesso le ordinanze 603 e 828 del 2006, con cui intima la rimozione e lo smaltimento di rifiuti che rappresentano una discarica di terra e materiali provenienti da demolizioni edili frammisti a materiali contenenti amianto, compresi rifiuti misti contenenti sostanze pericolose. La presenza di questa discarica sembra contraddire la risposta di ARPAT ad ASF-ISPO.

<sup>7</sup>“*è auspicabile comunque che, a seguito di questa segnalazione dei lavoratori, e dell’allarme che si è verificato tra di essi, venga attuata una sorveglianza sanitaria dei lavoratori in grado di monitorare patologie e fattori di rischio, grazie anche all’analisi periodica delle informazioni raccolte in occasione delle visite mediche. In questo modo sarà possibile capire se i dati qui esaminati cambieranno nel tempo e se necessitano di ulteriori approfondimenti e/o interventi di prevenzione.*” [1, p. 26]

to il personale precario, nonché dell’ambiente del Polo Scientifico.

Sarebbe inoltre quantomai opportuno mantenere sotto costante controllo anche l’ambiente esterno al Polo Scientifico, al fine di evidenziare eventuali innalzamenti di agenti inquinanti, purtroppo già a livelli preoccupanti in tutta la Piana, i quali potrebbero favorire l’insorgenza di patologie, specie in sinergia con possibili cause indoor.

## Riferimenti bibliografici

- [1] Gruppo di lavoro ASF-ISPO (a cura di). *Approfondimento sui casi di tumore della mammella verificatisi tra i lavoratori del CNR di Sesto Fiorentino*, 21 gennaio 2014. <http://www.area.fi.cnr.it/files/intranet/prevenzioneProtezione/RelazioneFinaleASLTumoriSeno.pdf>
- [2] Associazione Italiana dei Registri Tumori (AIR-TUM). *Tabella d’incidenza età-specifica per 100.000 - Femmine - Firenze Prato (2000-2005) in ITACAN: Tumori in Italia, Versione 2.0, 2014.* <http://itacan.ispo.toscana.it/italian/Table4r.asp?registry=3800301&sort=0&sex=2&type=0&stat=0&pdf=1&period=2005&period=2004&period=2003&period=2002&period=2001&period=2000&submit=Execute>
- [3] Fisher, Ronald Aylmer. *On the mathematical foundations of theoretical statistics*. Philosophical Transactions of the Royal Society of London, Series A, Containing Papers of a Mathematical or Physical Character, 1922, 309-368.
- [4] Ury, Hans K.; Wiggins, Alvin D. *Another shortcut method for calculating the confidence interval of a Poisson variable (or of a standardized mortality ratio)*. American Journal of Epidemiology, 1985, 122.1: 197-198.
- [5] D’Agostini, Giulio. *Bayesian reasoning in data analysis - A critical introduction*. ISBN: 978-981-238-356-3, World Scientific Publishing, Giugno 2003.