

Titolo:

“Grey Models” per la predizione delle variabili epidemiologiche.

Autori:

Maurizio Zarcone, Giuseppe Carruba, Vitale Miceli, Rosalba Staiti, Adele Traina.

Registro Tumori specializzato della Mammella della provincia di Palermo, Dipartimento di Oncologia, ARNAS Civico, Palermo.

Obiettivi:

In un periodo di profonda crisi economica, l'importanza di dover ottimizzare le poche risorse disponibili coinvolge anche la sanità pubblica, facendo emergere la necessità di prevedere, nel modo più accurato possibile, l'andamento delle variabili epidemiologiche, come quelle oncologiche, anche nel breve periodo.

In questo lavoro proponiamo l'applicazione di uno strumento matematico, elaborato in Cina negli anni '80, per prevedere l'andamento dei principali indicatori (incidenza e mortalità) utilizzati nell'epidemiologia dei tumori.

Materiali e metodi:

La teoria dei “sistemi grigi” (GST, Grey System Theory), che utilizza i cosiddetti modelli a “scatola grigia” (grey-box models), nata nell'ambito della teoria dei sistemi di controllo in ingegneria, negli ultimi 30 anni, è stata applicata con successo anche in diversi ambiti della ricerca dei fenomeni naturali e della fisica, dalla biologia alla meteorologia, dall'agricoltura ai terremoti, dall'inquinamento dei corpi idrici al percorso degli uragani, e così via. La necessità di applicazioni pratiche di questa teoria, caratterizzata dalla incompletezza dell'informazione sul sistema oggetto di studio, ha portato progressivamente alla realizzazione di nuovi metodi per affrontare determinati problemi: Grey Relational Analysis (GRA) per l'analisi delle interrelazioni, Grey Modeling (GM) per la modellazione e la predizione, Grey Decision-Making (GDM) e Grey Systems and Control (GSC) per il controllo e la valutazione dei dati.

Uno dei vantaggi principali nell'utilizzo dei “modelli grigi” (GM) è quello di poter fare previsioni di breve periodo con un esiguo numero di rilevazioni; per esattezza, è sufficiente una serie storica di soli quattro dati ideale, dunque, per quei territori coperti da registri di recente istituzione. Inoltre, negli ultimi trent'anni, sono stati sviluppati numerosi “modelli ibridi”, con lo scopo di ridurre gli errori delle stime nei vari campi di applicazione. Il modello base di questi modelli è indicato con GM(1,1), in cui il primo numero tra parentesi indica che viene adottata un'equazione differenziale del primo ordine e il secondo numero si riferisce al numero di variabili inserite nel modello.

Abbiamo utilizzato i dati pubblicati nella monografia AIRTUM del tumore della mammella in Sicilia (1), per quanto riguarda la mortalità, mentre l'incidenza è stata valutata e prevista per la sola provincia di Palermo utilizzando i dati del Registro Tumori della Mammella della provincia di Palermo. I risultati sono stati confrontati con i modelli statistici più comunemente utilizzati per le proiezioni degli indicatori tradizionali del carico sanitario oncologico.

Per i vari modelli analizzati sono stati valutati tre criteri, per il confronto, ovvero si prende in considerazione il minor valore per: lo scarto quadratico medio (MSE, Mean Square Error), l'errore medio assoluto (AME, Absolute Mean Error) e l'errore medio assoluto percentuale (AMPE, Absolute Mean Percentage Error).

## Risultati:

La Tabella mostra che le stime dei GM rispetto alle stime del software MIAMOD/PIAMOD, nella quasi totalità dei casi (solo l'MSE per le province di Agrigento e Caltanissetta è risultato più alto nei GM), hanno ottenuto i valori minori per i criteri utilizzati.

La Figura mostra la capacità predittiva del modello GM(1,1) sull'incidenza del tumore mammario nella provincia di Palermo, utilizzando solo quattro valori annuali (periodo 1999-2002) per prevedere i tassi di incidenza dei due anni successivi: si ottiene un errore massimo tra le stime di  $-3.9/100000$  (anno 2001) e di  $1/100000$  (per l'anno 2004) tra le predizioni e i dati effettivamente osservati.

## Discussione:

Questo lavoro suggerisce la possibilità di utilizzare efficacemente la GST in campo epidemiologico.

I GM possono essere utilizzati come strumenti "di precisione", in gestione clinica, per la previsione dello sviluppo del carico dei tumori nel breve periodo, a condizione, quindi, di aver a disposizione dati aggiornati sull'incidenza e sulla mortalità.

I GM che sembrano meglio adattarsi ai dati sul tumore della mammella, nella regione Sicilia, sono il modello GM(1,1) e un modello ibrido chiamato M3P-ARGM(1,1), basato sulla combinazione tra il modello di base GM(1,1), un modello ARIMA(1,1) "mediato a 3 punti" e una matrice di transizione di Markov.

- (1) AIRTUM, Registri tumori della Sicilia, OER-Sicilia. Il tumore della mammella nella Regione Sicilia. *Epidemiol Prev* 2009; 33(1-2) suppl 1: 1-96.

Tabella. Confronto dei modelli.

	AG	CL	CT	EN	ME	PA	RG	SR	TP
MIAMOD	MSE: 15.270	MSE: 6.962	MSE: 9.115	MSE: 24.849	MSE: 7.870	MSE: 3.763	MSE: 16.450	MSE: 13.732	MSE: 27.154
	AME: 3.119	AME: 2.326	AME: 2.613	AME: 3.671	AME: 2.048	AME: 1.537	AME: 3.139	AME: 3.105	AME: 4.013
	AMPE: 14.799	AMPE: 10.028	AMPE: 9.107	AMPE: 17.009	AMPE: 7.102	AMPE: 6.043	AMPE: 9.582	AMPE: 12.188	AMPE: 14.625
Grey Model ottimizzato	MSE: 15.798	MSE: 7.035	MSE: 6.398	MSE: 24.635	MSE: 6.595	MSE: 2.072	MSE: 9.623	MSE: 11.857	MSE: 24.035
	AME: 3.072	AME: 2.289	AME: 1.929	AME: 3.612	AME: 1.964	AME: 1.041	AME: 2.102	AME: 2.860	AME: 3.327
	AMPE: 14.651	AMPE: 9.844	AMPE: 6.731	AMPE: 16.631	AMPE: 6.911	AMPE: 3.994	AMPE: 6.427	AMPE: 11.289	AMPE: 11.364

Figura. Stime e predizioni del GM(1,1) sui tassi d'incidenza della provincia di Palermo (dati 1999-2004)

