

Giornale Italiano di **Farmacoeconomia** e **Farmacoutilizzazione**

**FOCUS SU UTILIZZO, RISCHIO-BENEFICIO E COSTO-EFFICACIA
DEI FARMACI E SULLE POLITICHE SANITARIE**

Rivista ufficiale di:
Servizio di Epidemiologia e Farmacologia Preventiva (SEFAP) e
Società Italiana di Terapia Clinica e Sperimentale (SITECS)

Pubblicazione trimestrale Volume 11 • Numero 4 • Dicembre 2019

RASSEGNA

**Machine learning: la capacità di prevedere applicata
alla ricerca e alla pratica clinica**

ARTICOLI ORIGINALI

**L'impatto dell'inserimento di nuovi ruoli professionali
sulla soddisfazione del paziente diabetico.
Evidenze dal progetto Europeo MUNROS**

SELEZIONE DELLA LETTERATURA

ANGOLO DEL MASTER IN FARMACOVIGILANZA

La pubblicazione di un articolo sul giornale GIFF implica l'impegno degli Autori a rispettare una open access Creative Commons license (CC-BY). Secondo i termini di questa licenza, gli Autori conservano la proprietà dei diritti d'autore dei loro articoli. Tuttavia, la licenza consente a qualsiasi utente di scaricare, stampare, estrarre, riutilizzare, archiviare e distribuire l'articolo, purché sia dato credito agli Autori e alla fonte del lavoro, attribuendo adeguatamente la paternità del materiale, fornendo un link all'originale e indicando se sono state effettuate modifiche.

Editor in Chief

Alberico L. CATAPANO

Dip. di Scienze Farmacologiche e Biomolecolari
Università degli Studi di Milano

Board editoriale

Vincenzo ATELLA

Facoltà di Economia
Università degli Studi di Roma "Tor Vergata"

Claudio BORGHI

Dipartimento di Medicina Clinica e Biotecnologia Applicata
"D. Campanacci" Università degli Studi di Bologna

Ovidio BRIGNOLI

Società Italiana di Medicina Generale (SIMG)

Achille CAPUTI

Dipartimento Clinico Sperimentale Medicina e Farmacologia
Università degli Studi di Messina

Manuela CASULA

Dipartimento di Scienze Farmacologiche e Biomolecolari
Università degli Studi di Milano

Francesco CIPOLLONE

Centro Studi di Medicina e Scienze dell'Invecchiamento
Università degli Studi di Chieti - Pescara

Giovanni CORRAO

Dipartimento di Statistica Socio Demografica
Università degli Studi Milano Bicocca

Romano DANESI

Dipartimento di Medicina Interna
Università degli Studi di Pisa

Gianfranco DE CARLI

GDC - Pharmaservices

Renato FELLIN

Dipartimento di Medicina Clinica e Sperimentale
Università degli Studi di Ferrara

Ettore NOVELLINO

Dipartimento di Chimica Farmaceutica e Tossicologica
Università degli Studi di Napoli Federico II

Francesco ROSSI

Dipartimento di Medicina Sperimentale
Seconda Università degli Studi di Napoli

Elena TRAGNI

Dipartimento di Scienze Farmacologiche e Biomolecolari
Università degli Studi di Milano

Mauro VENEGONI

Centro Regionale di Farmacovigilanza della Lombardia

Giornale Italiano di Farmacoeconomia e Farmacoutilizzazione

**FOCUS SU UTILIZZO, RISCHIO-BENEFICIO E COSTO-EFFICACIA
DEI FARMACI E SULLE POLITICHE SANITARIE**

Rivista ufficiale di:
Servizio di Epidemiologia e Farmacologia Preventiva (SEFAP) e
Società Italiana di Terapia Clinica e Sperimentale (SITeCS)

Pubblicazione trimestrale

Volume 11 • Numero 4 • Dicembre 2019

Sommario

► RASSEGNA

**Machine learning: la capacità di prevedere applicata alla ricerca
e alla pratica clinica**

Ornella Colpani

5

► ARTICOLO ORIGINALE

**L'impatto dell'inserimento di nuovi ruoli professionali sulla soddisfazione
del paziente diabetico. Evidenze dal progetto Europeo MUNROS**

Ruggeri Matteo, Drago Carlo, Damiani Gianfranco, Francesconi Paolo,
Gualdani Elisa, Barletta Valentina

12

► SELEZIONE DELLA LETTERATURA

► **Aspirina nel trattamento e nella prevenzione degli attacchi di emicrania**

► **Sistemi per l'identificazione di prescrizioni potenzialmente inappropriate
ed esiti nei pazienti**

Elena Tragni, Manuela Casula

20

► ANGOLO DEL MASTER IN FARMACOVIGILANZA

24

Cari colleghi,

il quarto numero del GIFF 2019 vi propone alcuni interessanti contributi.

Nella rassegna di Colpani et al. viene affrontato un argomento di particolare attualità quale il *machine learning*, sia dal punto di vista teorico che da quello delle possibili applicazioni in ambito medico, come una migliore stratificazione dei pazienti e diagnosi più accurate in tempi ridotti.

Nel contributo originale di Ruggeri et al., vengono presentati i risultati relativi all'impatto di modelli di cura integrata che prevedano la partecipazione del farmacista o dell'infermiere, accanto al medico, sulla soddisfazione dei pazienti affetti da diabete di tipo 2. L'analisi è stata condotta nell'ambito del progetto MUNROS (*Health Care Reform: the iMPact on practice, oUtcomes and costs of New roles for health pROfeSsionals*), finanziato dall'Unione Europea nel contesto del 7th Framework Program, con l'obiettivo di analizzare nuove figure professionali e nuove competenze e responsabilità nella gestione delle cure.

Nell'angolo dedicato al Master in Farmacovigilanza, Sentinelli et al. presentano il progetto svolto presso PHARMA D&S, in cui sono state analizzate e confrontate due metodiche per la rilevazione di segnali di sicurezza, ovvero l'approccio case by case e l'utilizzo di algoritmi statistici per la generazione del segnale nel database europeo tramite EVDAS (*EudraVigilance Data Analysis System*).

Buona Lettura!

Alberico L. Catapano
Editor in Chief

MACHINE LEARNING: LA CAPACITÀ DI PREVEDERE APPLICATA ALLA RICERCA E ALLA PRATICA CLINICA

Machine learning: the ability to predict applied to research and clinical practice

Ornella Colpani^{1,2}

¹Servizio di Epidemiologia e Farmacologia Preventiva (SEFAP), Dipartimento di Scienze Farmacologiche e Biomolecolari (DiSFeB), Università degli Studi di Milano

²Fondazione Bruno Kessler, Trento

Keywords

Machine learning
Medicine
Big data
Stratification

Abstract

Modern medicine is experiencing nowadays the era of -omic approaches, smartphones and health apps, generating a huge amount of data, the so-called big data. Such context inevitably requires new strategies and tools to store, integrate, and analyze data, and machine learning seems to be a promising framework to address these issues. Machine learning consists of algorithms that allow the machine to learn from data, without being explicitly programmed. The great potential of machine learning algorithms lies in their ability to make prediction or give a more synthetic and effective version of data, through dimensionality reduction and clustering methods. Machine learning can be a key tool in many branches of the biomedical research field, for instance for clustering genome regions according to similarities in epigenomic properties, or in identifying putative miRNA target genes, or to predict drug-target interactions. Nevertheless, the multiplicity of machine learning applications also embraces the clinical practice: patient stratification and prediction of risk scores are crucial for more accurate diagnosis and better therapeutic strategies. However, when training machine learning models, especially when applied to medical field, it is essential to pay attention to both the quality of data and to the effectiveness of the chosen algorithms for the investigated task. Finally, reproducibility and robustness of the learning procedures are crucial for their translation into clinical practice, forcing the FDA in introducing new guidelines in their approval protocols to deal with such brand new scenario.

Introduzione

Sebbene intelligenza artificiale, machine learning e deep learning siano spesso usati come sinonimi, esistono differenze importanti da chiarire. Infatti, l'intelligenza artificiale è un ramo dell'informatica che consente la creazione di sistemi che permettono di dotare le macchine di caratteristiche considerate proprie dell'intelligenza umana. Machine learning e deep learning fanno parte dell'intelligenza artificiale.

Intelligenza artificiale (AI), *machine learning* (ML) e *deep learning* (DL) stanno acquisendo una rilevanza sempre maggiore in ambito medico. Tuttavia, c'è parecchia confusione sull'utilizzo di questi tre termini, che vengono spesso usati impropriamente come sinonimi. È quindi necessario fare un po' di chiarezza su cosa siano esattamente AI, ML e DL e in cosa differiscano.

L'AI si riferisce a un campo dell'informatica dedicato alla creazione di sistemi che svolgono compiti che normalmente richiederebbero l'intelligenza umana. Nell'intelligenza artificiale le macchine completano l'attività in base alle regole e agli algoritmi stabiliti. Il termine "intelligenza artificiale" è quindi generico e comprende anche ML e DL. Il ML rientra nell'AI e si concentra su tutti quegli approcci che permettono alle macchine di imparare dai dati, senza che queste siano programmate esplicitamente. Un aspetto caratterizzante del ML è la dinamicità. Le macchine, infatti, ricevono una serie di dati e sono capaci di apprendere, modificando e migliorando le predizioni man mano che ricevono più informazioni su quello che stanno elaborando. Dunque, con l'apprendimento, gli algoritmi di ML tenderanno di minimizzare gli errori e massimizzare la probabilità che le loro previsioni siano vere [1]. Il DL è un sottogruppo del ML e incorpora modelli e algoritmi computazionali che imitano l'architettura delle reti neurali biologiche nel cervello (reti neurali artificiali, ANN) [1, 2].

Corrispondenza: Ornella Colpani. Servizio di Epidemiologia e Farmacologia Preventiva (SEFAP), Dipartimento di Scienze Farmacologiche e Biomolecolari (DiSFeB), Università degli Studi di Milano, via Balzaretti, 9 – 20133, Milano. E-mail: ornella.colpani@unimi.it

Box 1 Definizioni.

Intelligenza artificiale: campo dell'informatica dedicato alla creazione di sistemi in grado di risolvere problemi e di riprodurre attività proprie dell'intelligenza umana.

Machine learning: applicazione dell'intelligenza artificiale che consente alle macchine di imparare dai dati, senza che queste siano programmate in maniera esplicita.

Deep learning: sottogruppo del *machine learning* basato su modelli e algoritmi computazionali che imitano l'architettura delle reti neurali biologiche nel cervello.

Supervised learning: approccio nel quale alla macchina viene fornito un dataset di esempio (*training set*) per la generazione di diversi algoritmi capaci di correlare le caratteristiche dei dati con un'etichetta (per esempio la relazione tra determinate caratteristiche cliniche con l'etichetta «soggetto sano/malato»). Questi algoritmi vengono poi validati sul *validation set* per poter selezionare il più performante. Infine, l'algoritmo selezionato viene testato su un terzo set di dati (*test set*). La macchina impara quindi a fare previsioni su nuovi dati non etichettati.

Unsupervised learning: approccio utilizzato quando le etichette delle variabili input sono ignote. La macchina impara solo dai pattern nelle caratteristiche dei dati di input. L'obiettivo è quindi quello di fornire una versione sintetica dei dati tramite il loro raggruppamento sulla base di caratteristiche simili o la riduzione delle loro dimensioni.

Reinforcement learning: approccio nel quale la macchina riceve in input un obiettivo da raggiungere: per ogni azione che la porta verso l'obiettivo riceve un premio, per ogni azione che la allontana dallo scopo ottiene una penalità.

Approcci omici: approcci che consentono la produzione di grandi quantità di dati utili per la descrizione e l'interpretazione del sistema biologico studiato. Ne sono esempio la genomica, la metabolomica, la proteomica, la trascrittomica e l'epigenomica.

Epigenomica: studio del set completo di modifiche epigenetiche (modificazioni della cromatina sia a livello del DNA che delle proteine) nel materiale genetico di una cellula, noto come epigenoma. Le modifiche epigenetiche sono estremamente importanti in quanto regolano l'espressione genica.

High-throughput imaging: flusso di operazioni automatizzate volte a testare gli effetti di agenti perturbanti sul fenotipo cellulare. Tale flusso prevede l'incubazione delle cellule con l'agente perturbante, l'acquisizione di immagini tramite microscopi automatizzati e l'analisi delle immagini.

Machine learning: concetti chiave

Nel machine learning la macchina apprende dai dati. Esistono due tipi principali di apprendimento: quello supervisionato e quello non supervisionato.

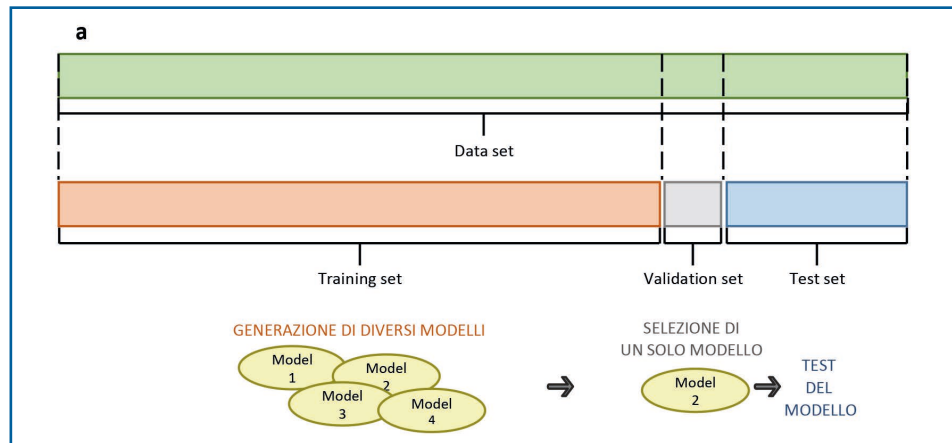
Il ML è la disciplina scientifica che si concentra su come i computer apprendono dai dati. È il risultato dell'incontro tra la statistica (che cerca di individuare relazioni dai dati) e l'informatica ed ha lo scopo di costruire modelli statistici da set di dati molto vasti. Esistono diverse categorie di apprendimento: qui menzioneremo quello supervisionato (*supervised learning*) e quello non supervisionato (*unsupervised learning*). Non parleremo invece dell'apprendimento per rinforzo (*reinforcement learning*), del quale è possibile trovare una breve spiegazione nel Box1.

Supervised learning

Nel *supervised learning* si parte da un set di dati che viene diviso in tre parti: *training set*, *validation set* e *test set* (**Figura 1**). Per cominciare, viene fornito alla macchina un dataset di esempio (*training set*) nel quale i dati sono etichettati. In altre parole, gli esempi sono composti da una coppia di dati contenenti il dato originale e il risultato atteso (variabile di risposta). Se pensassimo all'ambito clinico, questo significherebbe fornire alla macchina una dataset composto da casi clinici (che includono le caratteristiche di ogni paziente, per esempio analisi biochimiche, parametri vitali, ecc.) in cui l'esito per ogni paziente (sano o malato) è noto. Il compito della macchina è trovare la funzione che modelli la relazione tra i due, in modo tale da saper fare previsioni su nuovi dati in cui l'esito (nell'esempio precedente, sano/malato) è sconosciuto. Più formalmente, questi metodi servono ad individuare una funzione che predica la variabile di risposta (y) da un vettore di caratteristiche x che contiene M variabili input, in modo tale che $f(x)=y$. Se la variabile di risposta è di tipo categorico si parla di classificazione, nel caso di variabili di risposta continue (per esempio il costo di una casa sulla base delle sue caratteristiche) si parla invece di regressione. Esistono diversi algoritmi di *learning* in grado di creare funzioni che eseguano classificazioni o regressioni: *support vector machine*, alberi decisionali, *random forest* e *boosting* [3, 4]. I modelli/

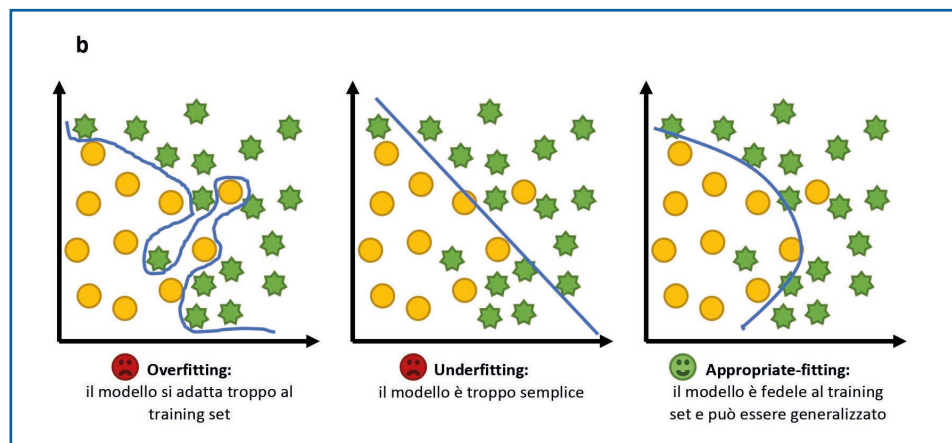
le funzioni generati durante il *training* vengono poi validati sul *validation set*. Questo permette di selezionare il modello più performante, per esempio in termini di accuratezza predittiva. L'ultimo passaggio consiste nel valutare la performance del modello selezionato su un nuovo set di dati, il *test set*.

Figura 1 *Supervised learning*: divisione del dataset in *training set*, *validation set* e *test set*.



Questi passaggi sono fondamentali per garantire di non incorrere nel cosiddetto *overfitting*, cioè la situazione in cui la funzione generata sia troppo sensibile alle caratteristiche particolari del *training set* piuttosto che alle caratteristiche generali del problema in analisi. In caso di *overfitting*, la funzione avrebbe performance di predizione altissime nel *training set* ma la sua performance risulterebbe molto inferiore sul *test set*. Nella **Figura 2** si può trovare una rappresentazione grafica di questo concetto [4, 5].

Figura 2 Rappresentazione schematica dei concetti di *overfitting*, *underfitting* e *appropriate-fitting*. Con bias/variance dilemma si intende la ricerca di un compromesso per ottenere un modello che non si adatti eccessivamente al training set e che non sia troppo semplice.



Unsupervised learning

Contrariamente a quanto visto per il *supervised learning*, gli approcci di *unsupervised learning* sono utilizzati quando le etichette delle variabili input non sono note e, dunque, questi metodi imparano solo i pattern delle caratteristiche dei dati di input. Gli obiettivi dell'*unsupervised learning* sono quindi: capire la distribuzione dei dati, il raggruppamento dei dati sulla base di caratteristiche simili o la riduzione delle loro dimensioni, al fine di fornirne una versione più sintetica ed efficace [6, 7]. Un'importante considerazione riguardo i metodi *unsupervised* è che misurare quanto l'algoritmo sia performante non è sempre semplice, poiché le prestazioni sono spesso soggettive e *domain-specific* [7].

I metodi *unsupervised* più comunemente utilizzati includono il *clustering* e la riduzione della dimensionalità dei dati [6]. Il *clustering* consiste nel raggruppamento dei dati in modo tale che dati appartenenti allo stesso cluster condividano caratteristiche simili, mentre dati facenti parte di cluster differenti siano diversi secondo una data

metrica. Esistono numerosi approcci di *clustering*, tra i quali alcuni dei più comuni sono il *k-means clustering* e il *clustering* gerarchico (Figure 3 e 4). Il *k-means clustering* divide i dati in k gruppi basando la partizione sulla creazione di un centroide per ogni cluster, restituendo come output una funzione che assegna ogni dato ad uno dei cluster [8]. Il *clustering* gerarchico mira alla formazione di una gerarchia di cluster: gli algoritmi stimano quindi le partizioni dei dati con una procedura di ottimizzazione sequenziale. Esistono due diverse strategie: quella agglomerativa (*bottom-up*) e quella divisiva (*top-down*). L'approccio agglomerativo parte da una serie di cluster separati che vengono sequenzialmente appaiati sulla base della similarità, l'approccio divisivo procede invece nella direzione opposta [9].

La riduzione della dimensionalità (*dimensionality reduction*) consiste nella riduzione della complessità dei dati, mantenendone il più possibile la struttura, e si avvale di strumenti sia classici, come l'analisi delle componenti principali (*principal component analysis*, PCA) e la decomposizione ai valori singolari (*singular value decomposition*, SVD), che più moderni, quali la *uniform manifold approximation and projection* (UMAP) e la *topological data analysis* (TDA).

Gli approcci di *unsupervised learning* sono frequentemente utilizzati per pre-processare i dati, comprimendoli prima di fornirli a reti neurali o ad altri algoritmi di *supervised learning* come dati di input [7].

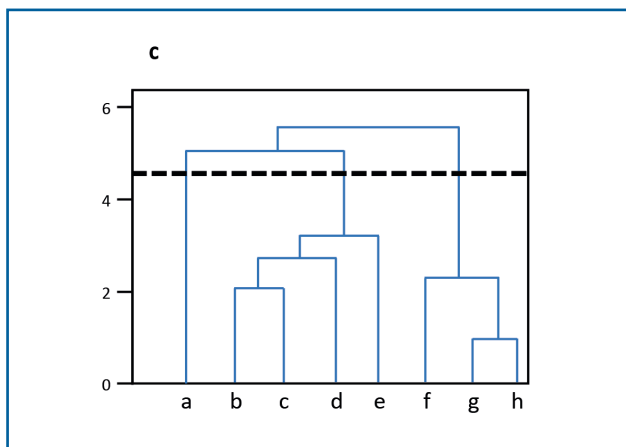


Figura 3 Clustering gerarchico.

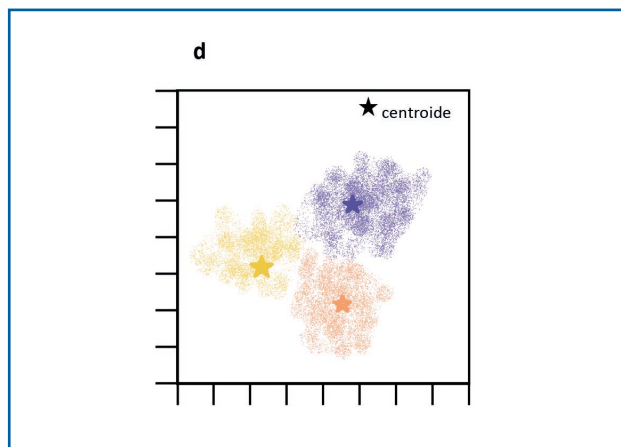


Figura 4 k-means clustering.

Machine learning e ricerca

L'avvento degli approcci omici ha generato una moltitudine di dati che devono essere integrati ed analizzati. Il machine learning, grazie ad approcci di supervised e unsupervised learning, si sta rivelando molto utile per rispondere a queste necessità.

La ricerca scientifica ha compiuto passi da gigante e volge ora lo sguardo verso la generazione e l'integrazione di un'enorme quantità di dati che permetterà un approccio olistico alla comprensione del funzionamento degli organismi viventi e dei meccanismi patologici. L'avvento delle moderne tecniche di sequenziamento a basso costo e la spettrometria di massa sono esempi emblematici della mole di dati che siamo ormai in grado di produrre; basti pensare che progetti come il *1000 Genomes* (progetto nato con lo scopo di fornire una descrizione completa delle varianti genetiche più comuni tramite sequenziamento *whole-genome* [10]) porteranno dati sulla scala dei petabyte (10^{15} byte). Non solo, lo sviluppo del sequenziamento di terza generazione e l'utilizzo di tecniche di single cell sequencing non potranno che incrementare ulteriormente la produzione dei cosiddetti *big data* [11]. In questo contesto è quindi fondamentale trovare la soluzione ai problemi di *storage* e di analisi che i *big data* portano inevitabilmente con loro. Come gestire quantità di dati così rilevanti? Come integrarli con dati provenienti per esempio da tecniche di *imaging*? Come analizzarli? Il ML è sicuramente un'arma potente per ovviare a questi problemi.

Uno dei campi in cui il ML si è rivelato di grande utilità è l'epigenomica, cioè lo studio delle modificazioni epigenetiche che contribuiscono alla regolazione dell'espressione genica. I ricercatori stanno raccogliendo numerosi dati tramite approcci omici (per esempio DNase-seq, FAIRE-seq, Hi-ChIP, ecc.) e il ML può essere un valido aiuto nell'interpretazione dei risultati e nell'integrazione di dati ottenuti da diverse metodiche [12].

Gli *epigenomic sequencing assay* si basano sulla frammentazione del DNA e terminano con il sequenziamento dei frammenti che presentano una determinata proprietà (come per esempio il legame ad una specifica proteina di interesse). Per fornire un'immagine completa dello stato epigenomico di ogni parte del genoma, è necessario combinare i risultati di numerosi saggi ed esperimenti. A questo scopo, sono stati sviluppati metodi di annotazione genomica semi-automatica (SAGA) che, basandosi su approcci di *unsupervised learning*, clusterizzano le regioni del genoma sulla base di similarità in termini di proprietà epigenomiche [13]. Lo studio di Malta et al. [14] è un esempio del grande potenziale degli algoritmi di ML. Partendo da dataset molecolari di cellule staminali e della loro progenie, sono stati generati due indici di staminalità utilizzati successivamente per valutare le caratteristiche epigenomiche e trascrittomiche di cellule cancerose (12000 tumori primari di 33 tipi di cancro diversi) con diverso grado di dedifferenziamento. Questo ha permesso di rilevare pattern di eterogeneità intra-tumore e potrebbe contribuire all'individuazione di nuovi target terapeutici [14].

Tramite ML è anche possibile predire quali sequenze di DNA possono costituire un sito di legame per fattori di trascrizione (superando le limitazioni dei saggi comunemente utilizzati in laboratorio), identificare possibili geni target dei microRNA e predire possibili effetti deleteri di polimorfismi [13, 15].

In maniera del tutto simile, il ML trova applicazione anche nella farmacologia computazionale [13]. Il ML si è infatti rivelato molto utile agli approcci di *image-based profiling* per il processo di *drug discovery*. Tali approcci si basano sul *high-throughput imaging*, un flusso di operazioni automatizzate che iniziano con l'incubazione delle cellule con agenti perturbanti (*small molecules*, siRNA, reagenti CRISPR/Cas9 [crRNA:trcrRNA], vettori plasmidici), a cui seguono l'acquisizione di immagini tramite microscopi automatizzati e, ad ultima, l'analisi delle immagini. Il fine di questi approcci consiste nel testare gli effetti degli agenti perturbanti di interesse sul fenotipo cellulare. Il *high throughput imaging* genera grandi raccolte di dati; le dimensioni e la complessità e la natura multiparametrica di tali dataset richiedono quindi strategie di analisi automatizzata. Per queste ragioni molti studi di *profiling* si affidano ad approcci di *unsupervised* e *supervised learning* per valutare cambiamenti fenotipici di singole cellule o di popolazioni cellulari. Per esempio, il *clustering* di profili *image-based*, per l'identificazione di molecole che condividono meccanismi di azione, si basa proprio su approcci di *unsupervised learning* [16, 17].

Machine learning e pratica clinica

Il machine learning sta acquisendo rilevanza anche in campo medico. Infatti, metodi di unsupervised learning permettono di stratificare i pazienti mentre supervised learning è possibile individuare nuovi fattori di rischio per diverse patologie, migliorando notevolmente l'accuratezza diagnostica.

Il ML potrebbe rivelarsi molto promettente anche nell'ambito della pratica clinica, dove potrebbe essere di supporto all'attività decisionale dei medici [18]. Per esempio, algoritmi di ML potrebbero generare una stima del rischio di un paziente per uno specifico outcome, aiutando il medico a impostare terapie più consapevoli [19]. Nondimeno, approcci di ML potranno essere usati nella precision medicine per identificare i pazienti che potrebbero rispondere meglio ad una determinata terapia, indirizzando il medico verso la strategia terapeutica migliore [20]. Non solo, studi comparativi hanno mostrato che tramite ML è possibile identificare nuovi predittori di rischio, migliorando l'accuratezza diagnostica quando paragonati a metodi tradizionali [3, 19]. Per esempio, in uno studio prospettico di coorte condotto in UK [21], quattro algoritmi di ML (*random forest*, *logistic regression*, *gradient boosting machines*, *neural networks*) sono stati paragonati all'algoritmo raccomandato dall'*American Heart Association/American College of Cardiology* (ACC/AHA) per la predizione di un primo evento cardiovascolare entro 10 anni. I risultati sono molto interessanti. Il diabete, classicamente considerato di grande rilevanza per il rischio cardiovascolare, non era presente tra i fattori di rischio principali in nessuno dei quattro algoritmi di ML testati. Erano invece presenti come fattori di rischio altre condizioni quali COPD, malattie mentali severe, prescrizione di corticosteroidi orali, alti livelli dei trigliceridi. In aggiunta, i quattro algoritmi, in particolare le reti neurali, hanno mostrato capacità predittive superiori a quelle del modello ACC/AHA [21]. Un altro studio di particolare interesse è quello di Esteva et al. [22]. Tramite tecniche di deep learning è stato possibile classificare correttamente lesioni cutanee a partire da immagini cliniche, con performance paragonabili a quelle di dermatologi esperti. Questo metodo è implementabile su dispositivi mobili e ren-

denderebbe quindi accessibile ed economico tale sistema diagnostico.

Un altro campo medico in cui il ML si sta facendo strada è il *medical imaging*, grazie alla capacità degli algoritmi di riconoscere pattern che vanno oltre la percezione umana, superando le performance dei radiologi. Il ML sarà dunque un valido aiuto per i medici che potranno trovare un supporto nell'interpretazione delle immagini riducendo i tempi richiesti per una diagnosi [23-24]. Ne è esempio il lavoro di Schoepf et al., che ha mostrato l'utilità di una *computer- aided detection* (CAD) per la rilevazione di emboli polmonari segmentari e subsegmentari [25].

Gli algoritmi di ML potrebbero anche essere una risposta promettente all'annoso problema della variazione fenotipica. Lo studio dell'associazione tra genotipo e fenotipo si basa tipicamente sulla dicotomizzazione dei partecipanti in due categorie, caso o controllo, in funzione di determinati criteri clinici. Tuttavia questa suddivisione rischia di essere eccessivamente semplicistica poiché non tiene in considerazione l'eterogeneità fenotipica tra i soggetti. Questo si traduce in una riduzione del potere statistico e dell'abilità di individuare una vera associazione tra una malattia e un locus genico. Approcci di *unsupervised learning*, forti delle grandi raccolte di dati ora disponibili, permetterebbero di definire profili fenotipici in maniera più fine o di stratificare i pazienti in sottogruppi più omogenei, superando le limitazioni dovute alla variazione fenotipica [26].

Gli approcci di ML sembrano quindi essere destinati ad entrare a far parte della pratica clinica, dove contribuiranno, tra le altre cose, alla stratificazione dei pazienti, all'identificazione dei soggetti più responsivi a determinate terapie e alla definizione di nuovi indici di rischio.

Machine learning: criticità

L'applicazione del machine learning alla pratica clinica richiede la messa a punto di un percorso di approvazione che si adatti alle particolari caratteristiche dei software ML-based.

Le prestazioni del software dovrebbero essere robuste in tutta la popolazione di pazienti target, la sua implementazione dovrebbe garantire un uso corretto e le sue analisi o previsioni dovrebbero fornire un contesto che aiuti l'interpretazione [27].

Una prima importante considerazione sull'utilizzo degli approcci ML nella ricerca e, ancora di più, nella pratica clinica riguarda la qualità degli algoritmi che è dipendente da quella dei dati utilizzati in fase di *training*. Infatti, non è escluso incorrere in bias statistici. Questo può essere conseguenza per esempio di campioni subottimali, misure errate nei predittori ed eterogeneità degli effetti. Si pensi ad un algoritmo generato da dati ottenuti solo da soggetti caucasici: è intuitivo pensare che lo stesso algoritmo avrebbe capacità predittive ridotte se applicato a soggetti asiatici o africani. Dunque, bisogna sempre tener presente che un algoritmo potrebbe avere performance ridotte se applicato per esempio a soggetti appartenenti alle minoranze etniche (fairness) [28]. È quindi fondamentale mitigare questo tipo di bias, per esempio preferendo dati provenienti da studi randomizzati invece che da studi osservazionali, dati non dipendenti dal giudizio del medico, o utilizzando strumenti volti a valutare il rischio di bias nell'algoritmo [29]. I modelli di machine learning sono generalmente più performanti quando il training set è più ampio: un nodo fondamentale sarà quindi il bilanciamento tra le normative sulla privacy e il bisogno di accesso a dati quanto più vasti e diversificati [30].

Un secondo punto riguarda l'approvazione dei software *machine learning-based* concepiti per la pratica clinica. Quando messi a punto per l'ambito medico, tali software vengono definiti come *medical device* nell'ambito della normativa *Food, Drug, and Cosmetic Act*. Tra il 2017 e il 2018 la FDA ha approvato 14 *AI/ML-based software* come *device*. Nell'aprile 2019, la FDA ha annunciato una revisione degli approcci regolatori per i software ML: ciò si è reso necessario a causa di alcune caratteristiche intrinseche di questi software, che rendono la loro approvazione complicata rispetto ai *medical device* tradizionali. Infatti, gli algoritmi ML sono per loro stessa natura interattivi. Ciò significa che un algoritmo, col passare del tempo, potrebbe comportarsi in maniera differente rispetto al momento della sua approvazione. Inoltre, bisogna tenere in considerazione che la maggior parte dei software (11 su 14) è stata approvata tramite il 510(k) *pathway*, che si basa sulla dimostrazione dell'equivalenza sostanziale tra il nuovo software e quello già in uso. Questo creerebbe una catena di *device* sostanzialmente equivalenti tra loro ma che col passare degli anni risulterebbero essere in realtà molto

diversi dal *device* originale. Infine, è necessario un percorso di approvazione che testi le effettive efficacia e sicurezza dei *ML-based software* e non solo la loro equivalenza [31]. A tal scopo bisogna ricordare che, sebbene alcuni software siano stati approvati dalla FDA, questi tool presentano molteplici criticità anche per quanto riguarda la loro riproducibilità che sta subendo una vera e crisi [32] e ciò richiede un approccio moderno per il superamento di questi limiti [33].

Conclusioni

In conclusione, il ML ricoprirà un ruolo sempre più preponderante sia nella ricerca che nella pratica clinica. Le applicazioni sono molteplici e la comunità scientifica sta guardando con grande interesse ai progressi di questi approcci così potenti e innovativi. La raccolta di grandi quantità di dati, generati tramite approcci omici ma anche dall'esplosione dell'utilizzo di dispositivi mobili, *app* e condivisione di dati on-line, sta aprendo la strada ad una medicina caratterizzata da diagnosi più accurate [34], migliori stratificazioni dei pazienti, tempi di diagnosi ridotti. Proprio perché il ML si sta rivelando estremamente potente, è di cruciale importanza stabilire una regolamentazione adeguata alla particolare natura di questi approcci, così da garantirne l'efficacia e la sicurezza.

Bibliografia

- [1] Jakhar D, Kaur I. Artificial intelligence, machine learning and deep learning: definitions and differences. Clin. Exp. Dermatol, 2019; ced. 14029.
- [2] LeCun Y, Bengio Y, Hinton G. Deep learning. Nature, 2015; 521: 436-444.
- [3] Deo RC. Machine Learning in Medicine», Circulation, 2015; 132: 1920-1930.
- [4] Schrider DR, Kern AD. Supervised Machine Learning for Population Genetics: A New Paradigm. Trends Genet. TIG, 2018; 34: 301-312.
- [5] Angermueller C, Pärnamaa T, Parts L, Stegle O. Deep learning for computational biology. Mol. Syst. Biol, 2016; 12: 878.
- [6] Camacho DM, Collins KM, Powers RK, Costello JC, Collins JJ. Next-Generation Machine Learning for Biological Networks. Cell, 2018; 173: 1581-1592.
- [7] Maini V, Sabri S. Machine Learning for Humans. pag. 97.
- [8] Brusco M, Steinley D. Psychometrics: Combinatorial Data Analysis», in International Encyclopedia of the Social & Behavioral Sciences (Second Edition), J. D. Wright, A. c. di Oxford: Elsevier, 2015; 431-435.
- [9] Kimes PK, Liu Y, Neil Hayes D, Marron JS. Statistical significance for hierarchical clustering. Biometrics, 2017; 73: 811-821.
- [10] 1000 Genomes Project Consortium et al. A global reference for human genetic variation. Nature, 2015; 526: 68-74.
- [11] Schadt EE, Linderman MD, Sorenson J, Lee L, Nolan GP. Computational solutions to large-scale data management and analysis. Nat. Rev. Genet, 2010; 11: 647-657.
- [12] Xu C, Jackson SA. «Machine learning and complex biological data», Genome Biol, 2019; 20: 76, s13059-019-1689-0.
- [13] Zitnik M, Nguyen F, Wang B, Leskovec J, Goldenberg A, Hoffman MM. Machine learning for integrating data in biology and medicine: Principles, practice, and opportunities. Inf. Fusion, 2019; 50: 71-91.
- [14] Malta TM, et al. Machine Learning Identifies Stemness Features Associated with Oncogenic Dedifferentiation. Cell, 2018; 173: 338-354.e15.
- [15] Ching T, et al. Opportunities and obstacles for deep learning in biology and medicine. J. R. Soc. Interface, 2018; 15: 20170387.
- [16] Pegoraro G, Misteli T. High-throughput imaging for the discovery of cellular mechanisms of disease. Trends Genet. TIG, 2017; 33: 604-615.
- [17] Scheeder C, Heigwer F, Boutros M. Machine learning and image-based profiling in drug discovery. Curr. Opin. Syst. Biol., 2018; 10: 43-52.
- [18] Ascent of machine learning in medicine. Nat. Mater., 2019, 18:407.
- [19] Peterson ED. Machine Learning, Predictive Analytics, and Clinical Practice: Can the Past Inform the Present?. JAMA, 2019.
- [20] Guthrie NL, et al. Achieving Rapid Blood Pressure Control With Digital Therapeutics: Retrospective Cohort and Machine Learning Study. JMIR Cardio, 2019; 3: e13030.
- [21] Weng S., Reps J, Kai J, Garibaldi JM, Qureshi N. Can machine-learning improve cardiovascular risk prediction using routine clinical data?. PLoS One, 2017; 12: e0174944.
- [22] Esteva A., et al. Dermatologist-level classification of skin cancer with deep neural networks. Nat., 2017, 542: 115-118.
- [23] Erickson BJ, Korfiatis P, Akkus Z, Kline TL. Machine Learning for Medical Imaging. RadioGraphics, 2017; 37: 505-515.
- [24] Topol EJ., High-performance medicine: the convergence of human and artificial intelligence. Nat. Med., 2019, 25: 44-56
- [25] Schoepf UJ, Schneider AC, Das M, Wood SA, Cheema JI, Costello P. Pulmonary embolism: computer-aided detection at multidetector row spiral computed tomography. J. Thorac. Imaging, 2007; 22: 319-323.
- [26] Basile AO, Ritchie MD. Informatics and machine learning to define the phenotype. Expert Rev. Mol. Diagn., 2018; 18: 219-226.
- [27] Towards trustable machine learning. Nat. Biomed. Eng., 2018, 2: 709-710.
- [28] Kelly CJ., et al. Key challenges for delivering clinical impact with artificial intelligence. BMC Med., 2019, 17, 195.
- [29] R. B. Parikh, S. Teeple, e A. S. Navathe. Addressing Bias in Artificial Intelligence in Health Care. JAMA, 2019.
- [30] Rajkomar A., et al. Machine Learning in Medicine. N. Engl. J. Med., 2019; 380:1347-1358.
- [31] Hwang TJ, Kesselheim AS, Vokinger KN. Lifecycle Regulation of Artificial Intelligence - and Machine Learning-Based Software Devices in Medicine. JAMA, 2019.
- [32] Huston M. Artificial intelligence faces reproducibility crisis. Science, 2018, 359:725-726.
- [33] Beam AL., et al., Challenges to the Reproducibility of Machine Learning Models in Health Care. 2020, Published online
- [34] Pani L. L'innovazione sostenibile. pag. 148.

L'IMPATTO DELL'INSERIMENTO DI NUOVI RUOLI PROFESSIONALI SULLA SODDISFAZIONE DEL PAZIENTE DIABETICO. EVIDENZE DAL PROGETTO EUROPEO MUNROS

The impact on patients satisfaction of the introduction of new health professional roles. Evidences from the MUNROS EU-funded project

Matteo Ruggeri^{1,2,3}, Carlo Drago⁴, Gianfranco Damiani^{3,5}, Paolo Francesconi⁶, Elisa Gualdani⁶, Valentina Barletta^{6,7}

¹Centro Nazionale per l'Health Technology Assessment, Istituto Superiore di Sanità, Roma

²St. Camillus International University of Health Sciences, Roma

³Università Cattolica del Sacro Cuore di Roma

⁴Università Niccolò Cusano, Roma

⁵Fondazione Policlinico Universitario Agostino Gemelli IRCCS, Roma

⁶ARS Toscana

⁷Azienda USL Toscana Centro SOS DSPO San Giuseppe, Empoli

Keywords

New professional roles
Integrated care
Type 2 diabetes

Abstract

In this work we will present results relating to the impact of integrated care models on the satisfaction of patients with type 2 diabetes in the Italian context. In particular, attention will be focused on the introduction, within the care pathway of the diabetic patient, of additional professional roles compared to the medical doctors, such as: nurse, pharmacist or other professions (e.g. podiatrist, psychologist). The data were extracted from those collected in the MUNROS project, funded by the European Union in the context of the 7th Framework Program.

Through an analysis of multiple correspondences, an overall satisfaction indicator was constructed. Secondly, through the propensity score matching, we investigated about the variables that had a significant impact on the probability of being treated by Nurses or Pharmacists. This allowed to test whether the inclusion of the pharmacist or nurse in addition to the medical doctor had a statistically significant impact on the patient's satisfaction.

The results of our study show how patients followed by nursing staff with respect to educational support activities, reported a satisfaction indicator of 0.8 against 0.7 of patients who had received assistance only from the doctor. This improvement was statistically significant with a value of $p < 0.05$. In the remaining comparisons, an improvement in satisfaction was observed but the differences were not statistically significant.

Our work shows that through the redesign of integrated healthcare according to the logic of the involvement of non-medical health professionals, it would be possible to lighten the workload of doctors without decreasing the level of perceived quality of care.

Introduzione

I processi di integrazione delle risorse professionali che sono coinvolte nella gestione delle cronicità, meglio noti come "cure integrate", si trovano sempre di più al centro dell'attenzione dei decisori sia a livello europeo che nazionale. L'idea alla base di tali processi è che la delega di alcune attività di educazione e monitoraggio da parte del personale medico verso altre figure professionali, come gli infermieri e i farmacisti, possa rappresentare un'occasione per ridurre i costi lasciando sostanzialmente invariati i livelli di qualità delle cure percepiti dai pazienti [1-7].

Corrispondenza: Matteo Ruggeri. Centro Nazionale per l'HTA, Istituto Superiore di Sanità, via Giano della Bella 34, 00161 Roma. E-mail: matteo.ruggeri@unicatt.it

Numerosi studi sono stati condotti relativamente agli effetti prodotti dall'introduzione di logiche di lavoro in team nella gestione del diabete. Jesmina et al. [8], nella loro analisi del modello canadese, mostrano che la maggior parte degli indicatori di outcome, la qualità e la continuità delle cure, sono tutti positivamente associati alle pratiche di lavoro in team. Tuttavia non vengono evidenziate relazioni significative che mostrino una diminuzione dei tempi di attesa e di accesso alle cure da parte dei pazienti legata ad un aumento dei livelli di coordinamento fra professionisti sanitari. Già nel 1994 un lavoro dell'*Integrated Diabetes Team* pubblicato sul *British Medical Journal* [9] si prefiggeva l'obiettivo di valutare l'approccio della cura integrata per il diabete. I risultati ottenuti mediante un trial pragmatico randomizzato mostravano come dal punto di vista clinico non ci fossero differenze statisticamente significative tra i pazienti trattati convenzionalmente e quelli con cura integrata, ma la soddisfazione dei pazienti insulino-dipendenti aumentava se assegnati al braccio di cura integrata rispetto al braccio tradizionale. In particolare, i pazienti che ricevevano una cura integrata effettuavano più visite e svolgevano esami e test più frequentemente rispetto ai pazienti del gruppo che riceveva cure convenzionali. Inoltre, i pazienti trattati con cura integrata erano più inclini a credere che la loro condizione fosse sotto il controllo dei loro medici. Il vantaggio più comunemente percepito dell'assistenza integrata era l'accessibilità, il risparmio di tempo e la continuità delle cure. Infine, dal punto di vista economico, i costi per i pazienti trattati con cure integrate erano più bassi rispetto ai pazienti che avevano ricevuto cure convenzionali. Negli anni la gestione integrata nel diabete si è rivelata uno strumento fondamentale per perseguire e raggiungere risultati soddisfacenti con particolare riguardo all'efficacia degli interventi, all'efficienza delle cure, della salute e della qualità di vita dei pazienti con patologie di lunga durata. Oggi sono molti gli approcci di cura integrata implementati in diverse realtà. L'approccio sistemico di tipo *Disease Management*, ad esempio, ampiamente diffuso in campo internazionale, mira a realizzare logiche di integrazione e coordinamento fra professionisti sanitari con il coinvolgimento attivo del paziente, volte al miglioramento della qualità delle cure e degli outcome sanitari. Tale approccio trova la sua applicazione con particolare riferimento alla gestione delle cronicità nel *Chronic Care Model*, che costituisce il riferimento organizzativo di tutte le esperienze più significative in Italia e all'estero. Mattke et al. grazie ad una meta-analisi che ha consentito di effettuare una valutazione dei programmi di *Disease Management*, dimostrano che tale approccio ha migliorato i processi clinici di cura (ad esempio l'aderenza alle linee guida basate sull'evidenza), portato ad un migliore controllo delle malattie, ridotto i tassi di ospedalizzazione e aumentato la soddisfazione dei pazienti e la qualità della vita correlata alla salute [10].

La soddisfazione del paziente può essere utilizzata come indicatore della qualità percepita dell'assistenza [11-14]. La soddisfazione del paziente dipende fortemente dal miglioramento percepito nella qualità della vita e quindi dall'efficacia, dalla sicurezza e dall'adeguatezza dei trattamenti. La percezione che i pazienti hanno della qualità delle cure ricevute è inoltre dipendente dalle modalità di interazione con la forza-lavoro in ambito sanitario. Ciò che condiziona fortemente il rapporto tra professioni sanitarie e pazienti è legato principalmente ai tratti personali dei medici (cordialità, cortesia, rispetto), ai servizi ricevuti (pulizia e qualità del cibo) il tempo trascorso con i pazienti e i fattori relativi alle visite (numero di visite, la durata della visita e il modo in cui le informazioni sulla salute vengono comunicate) [11-13]. Ciò che invece porta ad una riduzione della soddisfazione del paziente è legata alle barriere che riducono la qualità percepita dell'assistenza, che includono scarsa comunicazione e preferenza per un particolare genere di fornitore di assistenza sanitaria. Tuttavia, prendere decisioni, a livello di politica o di singolo fornitore, basandosi esclusivamente sulla soddisfazione del paziente, potrebbe portare a effetti negativi sui costi generali dell'assistenza e su altri indicatori di qualità come i tempi di attesa, le prescrizioni eccessive e i rinvii non necessari. È quindi necessario ricercare un punto di equilibrio tra la massimizzazione della soddisfazione del paziente e la fornitura di cure efficaci e sostenibili [13].

In questo lavoro, con riferimento all'Italia, presenteremo i risultati relativi all'impatto dei modelli di cura integrata sulla soddisfazione dei pazienti affetti da diabete di tipo 2. In particolare l'attenzione sarà concentrata sull'introduzione, all'interno del per-

corso del malato diabetico, di ruoli professionali aggiuntivi rispetto al medico, quali infermiere e farmacista [14]. I dati sono stati estratti da quelli più generali raccolti nell'ambito del progetto *MUNROS*, finanziato dall'Unione Europea nel contesto del *7th Framework Program* [15].

Metodi

Il progetto *MUNROS*

Lo studio *MUNROS* (*"Health Care Reform: the iMpaCt on practice, oUtcomes and costs of New roles for health pROfeSsionals"*), di tipo cross-sectional e di natura osservazionale, finanziato nell'ambito del programma FP7, coinvolge nove Paesi (Repubblica Ceca, Inghilterra, Germania, Italia, Olanda, Polonia, Scozia, Norvegia, Turchia) e ha come obiettivo quello di analizzare le nuove figure professionali e le nuove competenze e responsabilità ricoperte da quelle tradizionali, concentrandosi sui costi sanitari, sugli outcome di salute e sulla soddisfazione dei pazienti allo scopo di migliorare l'integrazione delle cure. Uno dei *work packages* del progetto ha previsto la raccolta di dati primari relativi a pazienti diabetici per comprendere come l'introduzione di nuovi ruoli professionali influenzasse indicatori di qualità percepita dai pazienti, fra cui anche il livello di soddisfazione. Il protocollo dello studio è stato approvato dai comitati etici dei rispettivi Paesi. Per quanto attiene l'Italia, l'approvazione è stata concessa dal comitato etico dell'Università Cattolica del Sacro Cuore nel 2013.

Popolazione di studio

I pazienti con diabete di tipo 2 sono stati identificati sia in setting ospedalieri che di cure primarie. I partecipanti sono stati invitati a completare un questionario in cui venivano richiesti dati demografici, stile di vita, esperienze di soddisfazione e assistenza. I Paesi che hanno partecipato al progetto sono stati nove, e sono stati appositamente scelti per evidenziare le differenze nei diversi sistemi di assistenza sanitaria, così da massimizzare la possibilità di osservare un cambiamento di fronte all'introduzione dei nuovi ruoli professionali rispetto a ciò che veniva fatto in precedenza. I Paesi inclusi sono stati: quelli in una fase successiva del processo di transizione da sistemi altamente centralizzati (Repubblica Ceca e Polonia); quelli all'avanguardia dell'innovazione (Paesi Bassi, Scozia e Inghilterra); quelli caratterizzati da sistemi più consolidati e stabili (Germania, Italia e Norvegia) e dalla Turchia, un Paese riformatore di successo nel suo sistema sanitario. Per quanto attiene al diabete di tipo 2, però, non sono stati raccolti questionari per i Paesi Bassi e la Germania.

In questo studio mostreremo i risultati relativi ai pazienti italiani. Costoro sono stati reclutati nella Regione Toscana, che da diversi anni ormai ha applicato il *Chronic Care Model* puntando sul concetto di "medicina di iniziativa", che si basa sulla capacità di coinvolgimento proattivo del paziente mediante attività di supporto educativo e *counseling*. Tale realtà può essere considerata una delle declinazioni più efficaci ed efficienti del *Chronic Care Model* sul territorio italiano.

I partecipanti erano pazienti che frequentavano l'ospedale reclutato o l'unità di assistenza primaria associata e che soddisfacevano i seguenti criteri di inclusione:

- uomo o donna di età uguale o superiore a 21 anni;
- in cura per il diabete di tipo 2;
- capace di comprendere lo scopo dello studio e di completare il questionario;
- con almeno due anni di follow-up.

Intervento

L'intervento che si è voluto testare è consistito nel coinvolgimento di professionisti sanitari diversi dal medico nel modello di presa in carico di pazienti diabetici. In particolare, le professionalità su cui si è concentrata l'attenzione sono state:

- a) infermieri,
- b) farmacisti.

Le attività in cui i professionisti erano coinvolti sono state individuate in:

- a) supporto educativo,
- b) monitoraggio.

Nel primo *set* di attività rientravano: il *counseling*, la fornitura di materiale informa-

tivo, indicazioni sulla corretta autosomministrazione dei farmaci e sui corretti stili di vita. Nel *set* di attività riguardanti il monitoraggio, invece, rientravano: la misurazione della pressione arteriosa, il controllo glicemico, il controllo del peso e del *Body Mass Index* (BMI) e il monitoraggio del piede diabetico.

Struttura del questionario e variabile di risposta

Il questionario era diviso in sei parti:

- Parte 1: relativa all'anagrafica e alla salute percepita dal paziente (tramite il questionario *EuroQoL*).
- Parte 2: riguardava le cure ricevute per la patologia in esame e la figura professionale che era entrata in contatto con il paziente per fornirgli educazione (rispetto agli stili di vita, al corretto utilizzo della terapia, ecc.) e monitoraggio (misurazione della pressione arteriosa, test della glicemia ecc.). Le figure professionali venivano divise in cinque macro-categorie: *staff* medico, *staff* infermieristico, farmacisti, professionisti collegati alla sanità (psicologi, terapisti occupazionali, psicoterapeuti, dietologi...) e altre figure professionali sanitarie.
- Parte 3: il paziente esprimeva una valutazione circa il livello di coordinamento tra i molteplici dottori/specialisti e il loro *staff* di supporto dentro e fuori l'ospedale e valutava su una scala da 1 (estremamente insoddisfatto) a 7 (estremamente soddisfatto) la sua soddisfazione durante l'ultima visita in ospedale o nello studio del medico/specialista in riferimento a: tempi di attesa, fornitura delle cure, continuità della cura, lunghezza della visita, informazioni ricevute, frequenza delle visite e soddisfazione circoscritta all'ultima visita;
- Parte 4: riguardava l'utilizzo dei servizi sanitari a seconda che si trattasse di pazienti ricoverati (in modo pianificato e non) o pazienti che avevano avuto accesso alle cure in ospedale (come pazienti ambulatoriali) negli ultimi tre mesi;
- Parte 5: veniva richiesta al paziente la disponibilità a pagare per una visita ideale, con la supplementare specifica di indicare un valore;
- Parte 6: vi era un elenco di caratteristiche demografiche relative al singolo paziente (peso, istruzione, occupazione, reddito, stili di vita).

Una volta analizzate le informazioni presenti nel dataset relativo al questionario per i pazienti, sotto un'ottica descrittiva, è stata costruita una variabile che potesse essere esplicativa del livello di soddisfazione generale del paziente rispetto all'assistenza ricevuta, così da poter racchiudere, all'interno di una sola, tutti gli aspetti relativi alla soddisfazione nella parte 3 del questionario. Per costruire tale variabile è stata condotta una analisi delle corrispondenze multiple, che è un tipo di analisi fattoriale basato sulla costruzione di nuove variabili "sintetiche" (dette fattori), ottenute da combinazioni lineari delle variabili originarie. Rappresenta uno strumento per lo studio delle relazioni tra un certo numero di caratteri statistici qualitativi, ognuno caratterizzato da m_j modalità ($j=1, \dots, p$). Nel nostro caso le variabili attive utilizzate per costruire l'indice sintetico di soddisfazione erano relative alle risposte date dai pazienti rispetto ai livelli di soddisfazione su otto dimensioni:

- 1) tempi di attesa,
- 2) lunghezza della visita,
- 3) fornitura delle cure,
- 4) continuità della cura,
- 5) lunghezza della visita,
- 6) informazioni ricevute,
- 7) frequenza delle visite,
- 8) soddisfazione circoscritta all'ultima visita effettuata.

Ciascuna di queste dimensioni prevedeva sette livelli (modalità) di soddisfazione (con il livello 1 che rappresentava il minimo). Utilizzando tali informazioni, lo score complessivo di soddisfazione è stato costruito individuando, attraverso l'analisi delle corrispondenze multiple, una graduatoria che consentisse di classificare le dimensioni a partire dalla più rilevante, attribuendo poi un sistema di pesi basato sulle frequenze di risposta relative ai singoli livelli e che consentisse all'indicatore complessivo di soddisfazione di assumere il profilo di variabile continua, con la possibilità di assumere valori fra 0 ed 1 (massima soddisfazione).

Analisi statistica

L'analisi descrittiva si è inizialmente sviluppata su informazioni generali (numero di intervistati, età media tra uomini e donne, fase del trattamento), successivamente invece è stata focalizzata l'attenzione sulle diverse variabili collegate alla soddisfazione del paziente.

In secondo luogo è stato condotto un quasi-esperimento, ovvero uno studio che utilizza una metodologia di randomizzazione *ex post* a seconda delle variabili di interesse presenti in uno studio osservazionale, utilizzando il *propensity score* come strumento per individuare gli accoppiamenti fra gruppo di trattamento e controlli sulla base di alcune caratteristiche di interesse e note. Tale metodo ha consentito di stimare gli effetti sulla soddisfazione dei pazienti diabetici dovuti alla presenza di un professionista sanitario (infermiere o farmacista) coinvolti nel fornire attività educazionali e/o monitoraggio a supporto del medico. In primo luogo sono stati prodotti sei modelli di regressione logit al fine di stimare l'impatto di alcune caratteristiche basali del campione sulla probabilità dei pazienti di ricevere in alternativa:

1. monitoraggio da infermieri;
2. monitoraggio da farmacisti;
3. supporto educativo da infermieri;
4. supporto educativo da farmacisti;
5. monitoraggio e/o supporto educativo da infermieri;
6. monitoraggio e/o supporto educativo da farmacisti.

Le variabili indipendenti incluse nel modello, rappresentative delle caratteristiche di base dei pazienti e necessarie per l'identificazione di criteri di randomizzazione *ex post* che fossero coerenti, sono state:

- età (variabile continua);
- genere (con valore 1 in caso di appartenenza al genere femminile);
- livello di istruzione (variabile dicotomica rispetto al diploma di scuola secondaria)
- presenza di comorbidità associate al diabete di tipo 2 (variabile dicotomica con valore 1 in caso di sussistenza di comorbidità).

In terzo luogo, utilizzando il *propensity score*, i pazienti appartenenti ai sei gruppi di intervento sono stati appaiati con rapporto 1:1 a pazienti aventi le stesse caratteristiche di base ma che avessero ricevuto supporto educativo o monitoraggio o entrambi dai soli medici. L'appaiamento è stato effettuato con reinserimento dei controlli.

I confronti si sono sviluppati osservando le variazioni nella variabile risposta ottenuta nell'analisi delle corrispondenze multiple. La significatività statistica delle differenze è stata infine verificata tramite un t-test appaiato.

Risultati

Baseline

Le risposte sono state ricevute da 289 pazienti.

La **Tabella 1** presenta una descrizione della situazione "pre appaiamento", con i valori medi degli indicatori di soddisfazione ricavati dall'analisi delle corrispondenze multiple, a seconda di alcune caratteristiche basali di cui sono riportate le relative numerosità.

Dalla tabella emerge come le attività inerenti il supporto educazione generino maggiori livelli di soddisfazione. Sia il supporto del medico che quello del farmacista e dell'infermiere sembrano essere associati ad alti livelli di soddisfazione. Inoltre emerge come le donne abbiano un livello di soddisfazione maggiore rispetto agli uomini e che anche i soggetti con un grado di istruzione più elevato (laurea) siano maggiormente soddisfatti, così come i pazienti con età maggiore di 60 anni.

Tabella 1 Livelli di soddisfazione.

CARATTERISTICHE	SODDISFAZIONE	
	supporto educativo (n= 243)*	monitoraggio (n = 187)*
Totale	0,74	0,75
Caratteristiche del paziente		
Uomini (n=157)	0,72	0,75
Donne (n=132)	0,75	0,77
Scuola media (n = 143)	0,62	0,52
Scuola superiore (n = 102)	0,64	0,67
Laurea (n = 44)	0,74	0,82
Età < 60 (n =162)	0,61	0,58
Età > 60 (n =127)	0,85	0,82
Figura professionale coinvolta		
Farmacista (n =62)	0,82	0,86
Infermiere (n =194)	0,87	0,90
Medico (n = 231)	0,86	0,85
Altro professionista (n =15)	0,74	0,72

*numerosità totale dei pazienti che hanno ricevuto monitoraggio e/o supporto educativo indipendente dalla figura professionale.

Analisi multivariata

La **Tabella 2** mostra i risultati dell'analisi multivariata condotta utilizzando il metodo logit. Il risultati mostrano che le variabili basali con più alto impatto sull'aumento di probabilità di ricevere monitoraggio e/o educazione da parte sia del farmacista che dell'infermiere sono la presenza di comorbidità e l'età. La variabile di genere (in questo caso il sesso femminile) ha un impatto generalmente negativo ma non significativo. Il solo caso in cui il sesso femminile ha un impatto positivo e significativo ($p < 0,1$) è sulla probabilità di ricevere monitoraggio dal farmacista.

La **Tabella 3** mostra invece i risultati dei quasi esperimenti rispetto ai livelli di soddisfazione dei pazienti che ricevano monitoraggio e/o educazione dai farmacisti o dagli infermieri in confronto con i medici. In generale i risultati sembrano mostrare un aumento dei livelli di soddisfazione dei pazienti che vedono coinvolti nel loro percorso infermieri e farmacisti. Tuttavia, l'unico caso in cui il confronto è da considerarsi significativo è quello in cui i pazienti ricevano educazione dall'infermiere.

Tabella 2 Analisi multivariata (n =231).

Variabili	Farmacista			Infermiere		
	Educazione	Monitoraggio	Educazione e/o Monitoraggio	Educazione	Monitoraggio	Educazione e/o Monitoraggio
Età	6,71*	3,27*	11,51**	2,46**	6,71*	7,42**
Comorbidità	12,28**	12,84**	14,8**	6,73**	9,50**	8,83**
Istruzione secondaria	5,51	7,72	4,21	7,12	10,03	2,76
Genere femminile	-3,91	2,52*	-8,01	-2,41	-2,38	-2,01

* $p < 0,10$; ** $p < 0,05$

Tabella 3 Risultati dei confronti fra livelli di soddisfazione utilizzando l'appaiamento per *propensity score* (numero di osservazioni).

Attività	Farmacista+Medico vs Medico		Infermiere+Medico vs Medico	
Educazione	0,76 (n=32)	0,69 (n=32)	0,80** (n=34)	0,70 (n=34)
Monitoraggio	0,83 (n=24)	0,81 (n=24)	0,74 (n= 19)	0,78 (n =19)
Educazione e/o monitoraggio	0,85 (n=18)	0,82 (n=118)	0,87 (n=21)	0,83 (n=21)

* $p < 0,10$; ** $p < 0,05$.

Discussione

Il nostro studio dimostra come la percezione del paziente da un punto di vista di soddisfazione del servizio percepito sia strettamente legato alle figure professionali sanitarie con cui viene a contatto. In particolare, su un campione di pazienti italiani presi in carico dalla rete di cure integrate della Regione Toscana, abbiamo dimostrato come il supporto educativo fornito dall'infermiere, che affianca il medico nella gestione di pazienti diabetici, aumenta i livelli di soddisfazione. Questi risultati non sono confermati nel caso del coinvolgimento dell'infermiere in attività di monitoraggio. Il nostro studio ha preso in considerazione anche l'inserimento del farmacista all'interno del percorso del paziente. Tuttavia, i risultati non sono significativi. Uno dei motivi potrebbe essere la bassa numerosità di pazienti che hanno ricevuto supporto educativo o monitoraggio da queste figure professionali. Ulteriori studi potranno esplorare il ruolo del farmacista in modo più robusto.

I nostri risultati si integrano con alcune tendenze già riscontrabili in letteratura che, sebbene siano state prodotte con metodi diversi ed abbiano investigato aspetti relativi all'introduzione di nuovi ruoli professionali differenti dalla soddisfazione dei pazienti, possono essere ricondotte ad un unico razionale di riferimento. Secondo Ouwens et al. [8] la figura dell'infermiere contribuisce a guidare i cambiamenti che stanno avvenendo nell'assistenza sanitaria, partecipando a pieno titolo al team interprofessionale. Questi cambiamenti richiedono un set nuovo o migliorato di conoscenze, abilità e attitudini in materia di benessere e assistenza della popolazione con una rinnovata attenzione per la centralizzazione del paziente, il coordinamento delle cure, l'analisi dei dati e il miglioramento della qualità. Inoltre, anche con riferimento al tipo di attività che vengono messe in atto, i risultati del nostro studio tendono a confermare precedenti evidenze. In uno studio condotto da Wilkes et al. [9] che aveva l'intento di definire il ruolo dell'infermiere in un team multidisciplinare nella gestione di malattie croniche e complesse, i risultati principali mostrano che le attività svolte dagli infermieri possono essere ricondotte a 6 ruoli chiave: consigliere, sostenitore, coordinatore, educatore, membro del team e valutatore. Il rapporto di lavoro dell'infermiere con gli altri professionisti della salute coinvolti nel team agevola questi ultimi alla comprensione dell'approccio della cura integrata, rendendo i team più efficaci. In precedenza, Baker et al. [10] ad esempio, avevano adottato un approccio coordinato per monitorare gli standard di pratica tra gli infermieri specialisti all'interno dei team di valutazione dello screening del seno nel West Midland. I risultati mostravano che la maggior parte delle donne che partecipavano alla valutazione dello screening del seno percepiva l'interazione con un infermiere specialista come altamente benefica per supporto psicologico e informazioni fornite.

Il nostro studio è uno dei pochi ad estrapolare l'impatto del ruolo dell'infermiere sui livelli di soddisfazione dei pazienti diabetici. Il motivo è dovuto alla complessità della patologia che rende difficile l'arruolamento e soprattutto la randomizzazione dei pazienti. Tale problema non è solo relativo alle condizioni cliniche ma anche all'interazione di fattori ambientali, familiari e socio-culturali, che hanno inevitabilmente un impatto sulle modalità di reazione (es.: i livelli di coinvolgimento del paziente) soprattutto in riferimento alle attività di supporto educativo svolte dai professionisti sanitari. L'approccio del quasi-esperimento mediante un processo di randomizzazione fittizia effettuato tramite l'utilizzo del *propensity score* ha consentito di condurre tale analisi in un'ottica comparativa, utilizzando dati raccolti in un contesto non controllato. Il nostro studio utilizza dati provenienti dalla regione Toscana, in cui, grazie all'implementazione del programma di medicina di iniziativa, il ruolo dei professionisti sanitari a supporto del medico è più facilmente osservabile. Tuttavia, tale caso non esaurisce la complessità dello scenario italiano. Anche in altre realtà il ruolo dell'infermiere nell'ottica della *Community Care* ha raggiunto maggiore importanza. In Emilia Romagna, ad esempio, nel 2011, è nata la Casa della Salute: sede di erogazione dell'insieme delle cure primarie rappresentate dai servizi sanitari, socio-sanitari e socio assistenziali rivolti alla popolazione dell'ambito territoriale di riferimento del Nucleo di Cure Primarie. La Casa della Salute, come sistema integrato di servizi, si prende cura delle persone fin dal momento dell'accesso attraverso: accoglienza e orientamento ai servizi; assistenza ambulatoriale per problemi urgenti; possibilità di completare i principali

percorsi diagnostici che non richiedono l'accesso in ospedale; gestione delle patologie croniche; integrazione e valorizzazione delle competenze delle diverse professionalità coinvolte; condivisione dei percorsi assistenziali.

Ulteriori studi, in un contesto assistenziale diverso, potranno approfondire con maggior dettaglio i risultati ottenuti anche con particolare riferimento alle altre figure professionali (ad esempio il farmacista) e ad altri percorsi di presa in carico di malati cronici (BPCO, Cardiovascolare...).

In un contesto di cambiamento degli scenari di presa in carico e, più in generale della gestione delle cronicità, i *decision makers* dovranno affrontare sfide di sostenibilità sempre più complesse. Oltre che all'introduzione di nuove tecnologie, una attenzione sempre maggiore sarà da rivolgere alla costruzione efficace ed efficiente di nuove architetture organizzative. Il settore sanitario, per quanto pervaso da una alta complessità tecnologica, rimane *labor intensive*. Per questo motivo, l'aumento dei livelli di efficienza dei percorsi assistenziali non può che passare da un ripensamento dei ruoli dei professionisti sanitari. Nondimeno, la ristrutturazione di tali percorsi non potrà essere esente dal preservare gli attuali livelli di qualità dell'assistenza. Per questo motivo, lo studio delle relazioni fra introduzione di nuove figure professionali nei percorsi e livelli di soddisfazione appare quanto mai attuale, sia al fine di comprendere se le azioni finora svolte sono da considerarsi riuscite, sia nel caso in cui si debba scegliere quali direzioni far prendere al processo di ristrutturazione dei sistemi di presa in carico.

Bibliografia

- [1] Lewis R, Rosen R, Goodwin N, Dixon J. Where next for integrated care organizations in the English NHS? The King's Fund 2010; London.
- [2] Curry N, Ham C. Clinical and service integration. The route to improved outcomes. The King's Fund 2010; London.
- [3] Ouwers M, Wollersheim H, Hermens R, Grol R. Integrated care programmes for chronically ill patients: a review of systematic reviews; International Journal for Quality in Health Care 2005; 17(2): 141-46.
- [4] Schoenfelder T. Patient Satisfaction: A valid indicator for the quality of primary care, Primary Health Care 2010; 2: e106.
- [5] Jesmina S, Thinda A, Sarma S. Does team-based primary health care improve patients' perception of outcomes? Evidence from the 08 2007 Canadian Survey of Experiences with Primary Health Health Policy 2008; 105: 71-83.
- [6] Diabetes Integrated Care Evaluation Team. Integrated care for diabetes: clinical, psychosocial and economic evaluation; British Medical Journal 1994; 308.
- [7] Mattke S, Seid M, Ma S. Evidence for the effect of disease management: is \$1 billion a year a good investment; American Journal of Managed Care 2007; 13(12): 670-6.
- [8] Ouwers M, Wollersheim H, Hermens R, Hulscher M, Grol R. Integrated care programmes for chronically ill patients: a review of systematic reviews; International Journal for Quality in Health Care 2005; 17(2): 141-46.
- [9] Wilkes L, Cioffi J, Cummings J, Warne B, Grad Cert Aged Care, Harrison K. Clients with chronic conditions: community nurse role in a multi-disciplinary team; Journal of Clinical Nursing 2013; 23(5-6): 844-55.
- [10] Baker J Patient satisfaction with clinical nurse specialists practice. Nursing standard 2013; 27 (37): 41-47.
- [11] Lee T, Choi KC, Chan CW, Chair SY, Chan D, Fung SY, Chan EL. The impact on patient health and service outcomes of introducing nurse consultants: a historically matched controlled study; BMC Health services research 2013; Oct 23; 13: 431.
- [12] Rahmqvist M. Patient satisfaction in relation to age, health status and other background factors: a model for comparisons of care units; International Journal for Quality in Health Care 2001; 13 (5): 385-90.
- [13] Jackson JL, Chamberlin J, Kroenke K. Predictors of patient satisfaction; Social Science & Medicine 2001; 52 (4): 609-620.
- [14] Tinelli, M, Ryan, M, Bond, C. Patients' preferences for an increased pharmacist role in the management of drug therapy International Journal of Pharmacy Practice 2009; 17(5): 275-82.
- [15] Bond C, Bruhn H, De Bont A, Van Exel J, Busse R, Sutton M, Svobodova H. The iMPact on practice, oUtcomes and costs of New roles for health pROfeSsionals: a study protocol for MUNROS; BMJ open 2016; 6(4).

Elena Tragni, Manuela Casula

Servizio di Epidemiologia e Farmacologia Preventiva (SEFAP), Università degli Studi di Milano

Aspirina nel trattamento e nella prevenzione degli attacchi di emicrania

ASPIRIN IN THE TREATMENT AND PREVENTION OF MIGRAINE HEADACHES: POSSIBLE ADDITIONAL CLINICAL OPTIONS FOR PRIMARY HEALTHCARE PROVIDERS

Biglione B, Gitin A, Gorelick PB, et al.

Am J Med, pubblicato on line il 9 novembre 2019

Introduzione

L'emicrania è tra i disturbi più comuni e potenzialmente debilitanti riscontrati dagli operatori sanitari. Un paziente su 10 in cure primarie presenta mal di testa, e 3 su 4 di questi hanno emicrania. Generalmente, l'emicrania si verifica in circa il 14% della popolazione generale, con una prevalenza più elevata nelle donne (18%) rispetto agli uomini (9%). Nelle donne, la prevalenza è più alta in età fertile.

Per il trattamento degli attacchi di emicrania acuta da lieve a moderata, i medici possono prescrivere farmaci antinfiammatori non steroidei (FANS), più comunemente ibuprofene. Per coloro che non ottengono un sollievo sintomatico, i triptani possono essere un'efficace terapia di seconda linea per attacchi da lievi a moderati, e un trattamento di prima linea per l'emicrania acuta da moderata a grave. Inoltre, sono stati proposti anche ergotamine, farmaci anti-nausea, oppioidi e glucocorticoidi. Nuovi prodotti biologici, come gli anticorpi monoclonali erenumab, fremanezumab e galcanezumab, sono stati recentemente introdotti sul mercato, ma ad un costo considerevole e i loro rapporti beneficio-rischio sono ancora in corso di valutazione in studi randomizzati. In questa revisione, sono raccolte e valutate le evidenze a supporto dell'uso di alte dosi di aspirina nel trattamento dell'emicrania acuta.

Aspirina ad alte dosi nel trattamento dell'emicrania acuta

In diversi studi randomizzati e nelle metanalisi, aspirina in dosi da 900 a 1200 mg è risultata significativamente più efficace del placebo e comparabile in termini di efficacia a 400 mg di ibuprofene e 50 mg di sumatriptan. Aspirina ha mostrato un'efficacia simile a 100 mg di sumatriptan, sebbene quest'ultimo fornisca un sollievo più rapido.

In uno studio randomizzato, in doppio cieco, controllato con placebo, 374 soggetti affetti da emicrania acuta sono stati assegnati a 1000 mg di aspirina effervescente o placebo. L'aspirina ha mostrato benefici clinicamente importanti rispetto all'endpoint primario predefinito di riduzione della gravità del mal di testa da dolore grave o moderato a lieve o nessun dolore 2 ore dopo la somministrazione. In particolare, il 55% dei soggetti trattati con aspirina ha manifestato una riduzione della gravità

del mal di testa rispetto al 36,8% nel gruppo a placebo ($p < 0,001$). Sono stati riportati effetti avversi nell'8,3% dei soggetti assegnati ad aspirina rispetto al 2,9% dei soggetti che assumevano placebo. Questi effetti avversi, tuttavia, erano da lievi a moderati, senza effetti collaterali gastrointestinali o renali o sanguinamenti maggiori. Un altro studio randomizzato in doppio cieco, controllato con placebo ha confrontato direttamente 3 strategie di trattamento versus placebo. In particolare, 312 soggetti sono stati assegnati in modo casuale a 1000 mg di aspirina effervescente o 50 mg di sumatriptan o 400 mg di ibuprofene o placebo. L'endpoint primario predefinito era la riduzione della gravità del mal di testa da dolore moderato o grave a lieve o assente. Aspirina è risultata significativamente più efficace del placebo ($p < 0,0001$) ma non era significativamente diversa da 400 mg di ibuprofene o 50 mg di sumatriptan. Per quanto riguarda gli effetti avversi, questi sono stati riportati dal 4,1% dei soggetti in aspirina rispetto al 5,7%, 6,6% e 4,5%, rispettivamente, dei soggetti in ibuprofene, sumatriptan e placebo. Un soggetto assegnato ad aspirina è incorso in una colica renale, ma questo è stato ritenuto un evento non correlato al farmaco. Al contrario, tuttavia, si è verificato un caso grave di ulcera duodenale perforata in un soggetto assegnato al gruppo di trattamento con ibuprofene ed è stato ritenuto un possibile evento avverso correlato al farmaco.

Rispetto ad altri farmaci prescritti per l'emicrania acuta, aspirina ad alte dosi ha un ulteriore vantaggio nel ridurre i sintomi frequentemente concomitanti di nausea, fotofobia e fonofobia. In uno studio randomizzato su 433 soggetti, 1000 mg di aspirina effervescente hanno prodotto benefici simili a 50 mg di sumatriptan, in particolare, il 43,8% dei pazienti assegnati ad aspirina e il 43,7% assegnato a sumatriptan hanno riportato una remissione completa di nausea, fotofobia e fonofobia entro 2 ore dopo la somministrazione. Entrambe queste percentuali erano significativamente più alte rispetto al 30,9% dei soggetti trattati con placebo che mostravano una remissione completa di questi sintomi ($p < 0,05$). Nelle analisi di efficacia sull'endpoint predefinito, aspirina ha ridotto significativamente la gravità dell'emicrania acuta rispetto al placebo, con il 49,3% dei soggetti trattati con aspirina che hanno riportato una riduzione della gravità del mal di testa da moderata o grave a lieve o nessun dolore, rispetto al 32,9% di soggetti trattati con placebo ($p < 0,05$). Aspirina ha mostrato un'efficacia simile a sumatriptan, nel cui gruppo il 48,8% dei pazienti ha manifestato una riduzione della gravità del mal di testa. Per quanto riguarda la sicurezza, il 4,7% dei soggetti trattati con aspirina ha riportato effetti avversi, rispetto al 6,7% dei soggetti trattati con sumatriptan e al 3,9% dei soggetti trattati con placebo. Alte dosi (100 mg) di sumatriptan hanno dimostrato efficacia nel trat-

tamento dell'emicrania acuta ma sono stati anche associati a frequenze più elevate di effetti avversi, inclusa la ricorrenza del mal di testa.

Metoclopramide, un antiemetico, è talvolta associato ad aspirina ad alto dosaggio per il trattamento dell'emicrania acuta e può aumentarne il beneficio sia nel trattamento del mal di testa sia nella riduzione dei sintomi associati. In uno studio randomizzato in doppio cieco, controllato con placebo, su 358 soggetti, 900 mg di aspirina più 10 mg di metoclopramide sono stati confrontati con 100 mg di sumatriptan e placebo. Per l'endpoint primario predefinito di riduzione della gravità del mal di testa da dolore grave o moderato a lieve o nessun dolore a 2 ore dall'assunzione, non vi era alcuna differenza statisticamente significativa tra la combinazione di aspirina e metoclopramide e sumatriptan ad alte dosi. Sumatriptan ad alte dosi ha fornito un sollievo più rapido del mal di testa rispetto ad aspirina più metoclopramide, ma è stato anche associato a frequenze più alte di recidive di mal di testa dopo 48 ore ed effetti avversi.

In una metanalisi di 13 studi randomizzati su 4222 partecipanti, 900 mg o 1000 mg di aspirina, con e senza 10 mg di metoclopramide, sono stati confrontati con placebo o con comparatori attivi, principalmente 50 mg o 100 mg di sumatriptan. L'endpoint primario predefinito era la riduzione del mal di testa o la remissione completa del dolore in vari intervalli di tempo. Aspirina ad alte dosi è risultata un trattamento efficace per il sollievo dell'emicrania acuta e per il sollievo dei sintomi associati all'emicrania di nausea e fotofobia. Inoltre, non c'erano differenze significative tra aspirina e 50 o 100 mg di sumatriptan. Per l'endpoint primario, aspirina ad alte dosi ha ridotto significativamente (-60%) il mal di testa dopo 1 ora rispetto a 50 mg di sumatriptan. Dopo 2 ore, tuttavia, 100 mg di sumatriptan hanno prodotto una riduzione significativa (-37%) rispetto ad aspirina ad alte dosi associata a metoclopramide. Per quanto riguarda gli effetti avversi, non vi sono state differenze significative tra coloro assegnati ad aspirina ad alte dosi, con o senza metoclopramide, e coloro assegnati a 50 mg di sumatriptan o placebo. I soggetti assegnati a 100 mg di sumatriptan, tuttavia, hanno manifestato un aumento statisticamente significativo del 34% degli effetti avversi correlati al farmaco rispetto a quelli assegnati ad aspirina più metoclopramide.

Gli operatori sanitari dovrebbero anche essere consapevoli del fatto che la *Food and Drug Administration* statunitense ha emesso un *black box* per metoclopramide se usata per periodi di tempo superiori a 12 settimane, a causa del rischio di sviluppare discinesia tardiva. Inoltre, per quanto riguarda gli effetti collaterali, l'evidenza randomizzata su aspirina è molto più rassicurante di quella dei tradizionali FANS non selettivi e degli inibitori selettivi della COX-2. Inoltre, i triptani possono avere conseguenze cardiovascolari avverse, che dovrebbero essere prese in considerazione al momento della prescrizione.

Conclusioni

La totalità delle evidenze, inclusi dati provenienti da studi randomizzati, suggerisce che aspirina ad alte dosi (da 900 a 1300 mg), somministrata all'inizio dei sintomi, è

un'opzione di trattamento efficace e sicura per l'emicrania acuta. Il profilo relativamente favorevole degli effetti collaterali e i costi estremamente bassi rispetto ad altre terapie farmacologiche su prescrizione possono fornire un'ulteriore opzione clinica per gli operatori sanitari nel trattamento dell'emicrania acuta e ricorrente.

Sistemi per l'identificazione di prescrizioni potenzialmente inappropriate ed esiti nei pazienti

COMPARING POTENTIALLY INAPPROPRIATE PRESCRIBING TOOLS AND THEIR ASSOCIATION WITH PATIENT OUTCOMES

Moriarty F, Bennett K, Kenny RA, et al.

J Am Geriatr Soc, pubblicato on line l'1 novembre 2019

Introduzione

La prescrizione potenzialmente inappropriata (*Potentially Inappropriate Prescribing*, PIP) tra gli anziani è diventata un importante problema di salute pubblica in tutto il mondo e si ritiene che abbia un significativo impatto negativo sia clinico che economico. Una PIP può essere la prescrizione di un farmaco senza un'indicazione clinica, o di un farmaco indicato ma per cui i rischi superano i benefici o per cui esiste un'alternativa più sicura. Può includere anche l'omissione di un farmaco clinicamente indicato. Le persone anziane sono particolarmente a rischio di effetti avversi da prescrizione subottimale, a causa di cambiamenti fisiologici dovuti all'invecchiamento e alla probabilità di avere più malattie e farmaci, predisponendoli quindi ad interazioni farmacologiche.

La PIP viene valutata più frequentemente utilizzando misure o criteri di processo espliciti. I criteri espliciti valutano se la prescrizione è conforme a standard accettati e può essere applicata in assenza di giudizio clinico. Di solito sono sviluppati a partire da revisioni pubblicate, opinioni di esperti e tecniche di consenso o *Delphi*. I farmaci da evitare elencati nei criteri di Beers dell'*American Geriatrics Society (AGS)* sono la misura esplicita di processo di PIP più frequentemente utilizzata e validata. I criteri di Beers sono stati sviluppati negli Stati Uniti nel 1991 e aggiornati nel 1997, 2003, 2012, 2015 e 2019. A questi è seguito un numero significativo di misure modificate o nuove, tra cui i criteri *Screening Tool of Older Persons' Potentially Inappropriate Prescriptions (STOPP)* e *Screening Tool to Alert Doctors to the Right Treatment (START)*, gli *Assessing Care of Vulnerable Elders (ACOVE)* e misure specifiche per Paese. Tutti questi sono stati ampiamente utilizzati nella ricerca, sia per valutare la prevalenza di PIP e la loro relazione con outcome economici e clinici, che come misura di esito per interventi volti a migliorare la pratica prescrittiva.

La scelta della misura di PIP negli studi è solitamente arbitraria e spesso non viene fornita alcuna motivazione alla base della scelta. A volte vengono utilizzate misure specifiche per Paese (cioè adattate localmente, sulla base dei farmaci comunemente o non comunemente usati). Tuttavia, idealmente, le misure di PIP dovrebbero essere generalizzabili e comparabili tra Paesi. Diversi studi hanno utilizzato più misure di PIP, anche se l'accordo tra queste misure è stato raramente valutato in modo formale. Inoltre, pochi studi hanno effettuato confronti in termini di validità predittiva o associazione con esiti avversi.

Molte delle misure esplicithe di PIP incorporano una vasta gamma di criteri e un'analisi comparativa può aiutare a

identificare classi di farmaci inappropriate o criteri associati ad effetti avversi sulla salute che possono essere applicati nella pratica clinica. Perciò, l'obiettivo di questo studio è stato quello di valutare la concordanza di diverse misure esplicite di PIP (relative sia agli errori di over- o mis-prescrizione che di under-prescrizione od omissione) e di confrontare la loro relazione con gli outcome riportati dai pazienti tra persone anziane residenti in Irlanda.

Metodi

È stato condotto uno studio di coorte prospettico su partecipanti della coorte *The Irish Longitudinal Study on Ageing (TILDA)*. L'esposizione a PIP è stata valutata separatamente sia per le prescrizioni fatte non appropriate che per quelle omesse, utilizzando un sottoinsieme degli indicatori applicabili (criteri di Beers 2012, STOPP versione 1 e ACOVE versione 3 nel primo caso e START versione 1 e ACOVE versione 3 nel secondo caso). Le informazioni erano disponibili per applicare il 69% dei criteri STOPP (45/65), l'81% dei criteri di Beers (42/52) e il 77% degli indicatori ACOVE (17/22) relativi a medicinali inappropriate, mentre il 68% dei criteri START (15/22) e il 34% degli indicatori ACOVE (21/65) relativi alla prescrizione di medicinali indicati. La loro concordanza è stata valutata tramite statistiche κ ed è stata utilizzata la regressione multivariata per determinare le relazioni con le visite al pronto soccorso, le visite dal medico di medicina generale (MMG), la qualità della vita e il declino funzionale (maggiore assistenza necessaria per le attività di vita quotidiana).

Risultati

In totale, sono stati inclusi in questa analisi 1753 partecipanti TILDA di età pari o superiore a 65 anni. La loro età media era di 76 anni (intervallo interquartile 72-80 anni) e il 54,4% era di sesso femminile. Una maggioranza sostanziale (79,7%) riferiva multimorbilità (≥ 2 condizioni croniche diagnosticate dal medico) e i partecipanti assumevano una mediana di 6 farmaci (intervallo interquartile 4-9).

Per i farmaci potenzialmente inappropriate, la prevalenza di criteri relativi a benzodiazepine, anticolinergici, antidepressivi triciclici e coxib/FANS con ipertensione è apparsa più alta nel gruppo con due o più indicatori ACOVE rispetto a quello con due o più STOPP e a quello con due o più criteri di Beers. La prevalenza dell'uso di PPI a dose massima a lungo termine è risultata inferiore nel gruppo con due o più criteri di Beers rispetto ai gruppi con due o più criteri ACOVE o STOPP. La prevalenza di prescrizioni omesse relative a statine, antiaritmici e broncodilatatori per inalazione è risultata inferiore nel gruppo con due o più omissioni ACOVE rispetto al gruppo con due o più omissioni START.

Considerando una esposizione binaria a PIP, le statistiche κ erano 0,28 (IC al 95% 0,24-0,32) per STOPP vs criteri di Beers e 0,22 (0,19-0,26) vs indicatori ACOVE (concordanza lieve), mentre la concordanza era moderata per criteri di Beers vs ACOVE ($\kappa=0,43$; 0,39-0,48). C'era una concordanza lieve tra gli indicatori START e ACOVE (statistica $\kappa=0,34$; 0,31-0,38), e gli indicatori

ACOVE hanno identificato una percentuale maggiore di partecipanti come esposti. La statistica κ per l'esposizione binaria a PIP era 0,51 (0,47-0,55) per gli indicatori START e ACOVE, indicando una concordanza moderata. Confrontando le misure di farmaci inappropriati, tutte e tre (criteri di Beers, STOPP e ACOVE) hanno mostrato un'associazione statisticamente significativa con un aumento del tasso di visite dei MMG. Solo l'esposizione a due o più indicatori ACOVE era associata a una riduzione della qualità della vita (differenza media aggiustata -1,68; da -3,06 a -0,30) e solo due o più criteri di Beers erano associati a un declino funzionale (*odds ratio* [OR] aggiustato 2,27; 1,47-3,52). Per le misure di omissione di farmaci, la probabilità di declino funzionale aumentava con due o più criteri START (OR aggiustato 2,04; 1,24-3,35) o due o più indicatori ACOVE (OR aggiustato 1,89; 1,21 -2,96). Solo due o più indicatori ACOVE erano associati ad un aumento delle visite dal MMG (tasso di incidenza aggiustato 1,13; 1,04-1,24).

Conclusioni

Questo studio ha evidenziato una concordanza lieve-moderata tra le diverse misure di PIP applicate a una coorte di anziani. In generale, c'era un maggiore accordo tra le misure relative a omissioni inadeguate di farmaci (gli indicatori START e ACOVE) rispetto alle misure relative alla prescrizione di farmaci inappropriati (STOPP, criteri di Beers e ACOVE). Inoltre, l'associazione tra PIP ed esiti differiva a seconda della misura di PIP utilizzata per definire l'esposizione. Tutte le misure di farmaci inappropriati erano associate a un aumento del tasso di visite dal MMG, mentre solo i criteri di Beers erano associati al declino funzionale e solo gli indicatori STOPP erano significativamente associati a visite al pronto soccorso. Al contrario, le omissioni erano più comunemente associate a una riduzione della qualità della vita e al declino funzionale e meno all'aumento dell'uso di servizi sanitari. Nessuna singola misura di PIP sembra essere superiore all'altra e, pertanto, la scelta di come misurare la PIP nella ricerca o nella pratica clinica dovrebbe essere considerata alla luce delle circostanze e dei requisiti di ciascun singolo caso. Una revisione sistematica attualmente in corso su tutti gli indicatori espliciti di PIP individuali per paziente fornirà una banca completa degli indicatori esistenti e ciò potrebbe ridurre la necessità di scegliere una singola misura di PIP da utilizzare nella pratica clinica o nella ricerca. Lo sviluppo e la condivisione di standard comuni per rendere operativi gli indicatori nelle cartelle cliniche elettroniche potrà anche migliorare la comparabilità e la coerenza dei risultati di diversi team o contesti di ricerca.



EVDAS (EUDRAVIGILANCE DATA ANALYSIS SYSTEM): NUOVO STRUMENTO PER LA SIGNAL DETECTION EVDAS (Eudravigilance Data Analysis System): new tool for signal detection process

Nicolò Sentinelli, Lucia Biagiotti¹

Progetto di stage svolto presso

¹ PHARMA D&S

Keywords

Signal-detection
case-by-case
EVDAS
EMA
ROR

Abstract

Background In the last couple of years, the European pharmacovigilance system underwent deep and significant changes. In particular, on November 22nd 2017, the European Medicines Agency (EMA) launched a new version of Eudravigilance, with enhanced functionalities, needed to support the significant changes introduced by the European pharmacovigilance legislation in terms of requirements for reporting suspected adverse reactions. Within this new system, EMA enabled Marketing Authorisation Holders (MAH) to access to the EudraVigilance Data Analysis System (EVDAS). EVDAS has been designed to allow users to analyse safety data collected in EudraVigilance, enabling better-informed decisions about the safety profile of medicinal products and, above all, providing signal detection tools.

Aim The aim of this thesis is to compare the results, in terms of safety data, obtained with the classical signal detection (case-by-case) to the new EVDAS signal detection approach, underlying the main differences and trying to give an explanation to them.

Method The safety data about a specific drug (X) owned by the company (Y) has been analysed. The first half of 2018 has been considered. In this period, firstly, a classical signal detection regarding the drug (X) through a case-by-case method has been carried out: describing scientifically all listed and unlisted cases obtained from the EudraVigilance database, analysing safety data obtained by the screening of literature and social media, and merging that information with the analysis of the sales data of X compared to previous years. Regarding the EVDAS signal detection, the platform has been accessed. Then, the time interval for the analysis of data has been selected choosing the ad-hoc option, and the active principle substance (X) has been selected from the provided list. Subsequently, the e-RMR file has been downloaded and the adverse drug reaction (ADR) occurred were collected. At this point, it was possible to obtain the statistic value (calculated as ROR) and the list of possible signals was collected. Last, a comparison between these ADRs (statistically considered as possible signals by EVDAS) and the ADRs founded in the safety database of the company (Y) has been made. For the ADRs, which matched with those obtained from EVDAS, a clinical evaluation to validate the signal has been carried out.

Results and conclusions Considering our study, the comparison between the classical signal detection (case-by-case) and EVDAS doesn't show any particular difference in terms of safety data and results obtained. Although we didn't get new potential safety signals, EVDAS represents a step forward in the present and future of performing a signal detection for pharmaceutical companies.

Introduzione

Il processo di Signal Management

Un segnale di safety è un'informazione riguardo un evento avverso nuovo o noto che può essere causato da un medicinale e richiede ulteriori indagini. L'Agenzia Europea per i Medicinali (EMA), insieme alle autorità regolatorie degli Stati membri e ai titolari dell'autorizzazione all'immissione in commercio (MAH), sono responsabili della rilevazione e della gestione dei segnali di safety. Questi segnali possono essere rilevati da un'ampia gamma di fonti, come segnalazioni spontanee, studi clinici e letteratura scientifica. Ad esempio, il database EudraVigilance è un'importante fonte di informazioni riguardo sospette reazioni avverse e segnali.

L'identificazione di un segnale di safety non implica necessariamente che sia stato uno specifico medicinale a causare l'evento avverso segnalato. Quest'ultimo infatti potrebbe essere dovuto anche ad altre cause, quali patologie concomitanti o altri farmaci o sostanze assunti dal paziente.

Quindi, la valutazione dei segnali di safety serve per stabilire se esiste o meno una relazione causale tra il medicinale e l'evento avverso segnalato, pertanto essa è parte delle attività di routine di farmacovigilanza ed è essenziale per garantire che le autorità regolatorie dispongano delle più aggiornate informazioni sui benefici e sui rischi di un medicinale.

Il processo di gestione del segnale è un insieme di attività svolte per determinare se, sulla base di un esame degli *individual case safety reports* (ICSR), dati aggregati da sistemi o studi di sorveglianza attivi, informazioni sulla letteratura scientifica o altre fonti di dati, ci sono nuovi rischi associati ad un principio attivo o ad un medicinale o se i rischi noti sono cambiati, nonché per stabilire eventuali raccomandazioni, decisioni e comunicazioni correlate.

Il processo di gestione del segnale dell'UE include i seguenti step (**Figura 1**):

- **Signal detection:** attività di ricerca e/o identificazione di segnali utilizzando dati provenienti da qualsiasi fonte.
- **Signal validation:** fase di valutazione dei dati a supporto del segnale rilevato, al fine di verificare che la documentazione disponibile contenga prove sufficienti a dimostrare l'esistenza di una nuova potenziale associazione causale o un nuovo aspetto di un'associazione nota e che pertanto giustifica un'ulteriore analisi più approfondita del segnale.
- **Signal confirmation:** stadio del processo in cui si decide se un segnale convalidato o meno inserito nell'*European Pharmacovigilance Issues Tracking Tool* (EPITT) richiede ulteriori analisi e la prioritizzazione da parte del PRAC (*Pharmacovigilance Risk Assessment Committee*). Ciò dovrebbe essere fatto dal relatore del PRAC o dallo Stato membro (principale) (MS) entro 30 giorni dalla ricezione del segnale convalidato.
- **Signal analysis and prioritization:** stadio in cui il PRAC determina se un segnale già confermato richiede un'ulteriore valutazione e, se necessario, con quali scadenze e in quale quadro procedurale. Ciò si basa su un'analisi iniziale del potenziale impatto del segnale sulla salute dei pazienti o sulla salute pubblica e sul rapporto rischio/beneficio del medicinale in questione.
- **Signal assessment and recommendation for action:** step finale, guidato dal PRAC, di valutazione di tutti i dati disponibili relativi a un segnale per determinare la necessità di un intervento normativo [1].

Signal Detection

È evidente che il primo step nel processo di gestione del segnale, la *signal detection*, gioca un ruolo chiave [2]. L'obiettivo della *signal detection* può essere espresso in vari modi, ma sono due gli elementi significativi:

- identificare, il prima possibile, qualsiasi indicazione riguardo un potenziale rischio associato a un farmaco, sia nuovo che vecchio.
- destare sospetti meritevoli di indagini più approfondite, come gli studi farmacologici epidemiologici.

I segnali di farmacovigilanza derivano generalmente da osservazioni in singoli pazienti o in popolazioni o da studi sperimentali e possono avere una dimensione sia qualitativa che quantitativa (**Tabella 1**) [3].

Figura 1 Gestione del segnale ed enti regolatori che sovrintendono ai diversi processi.

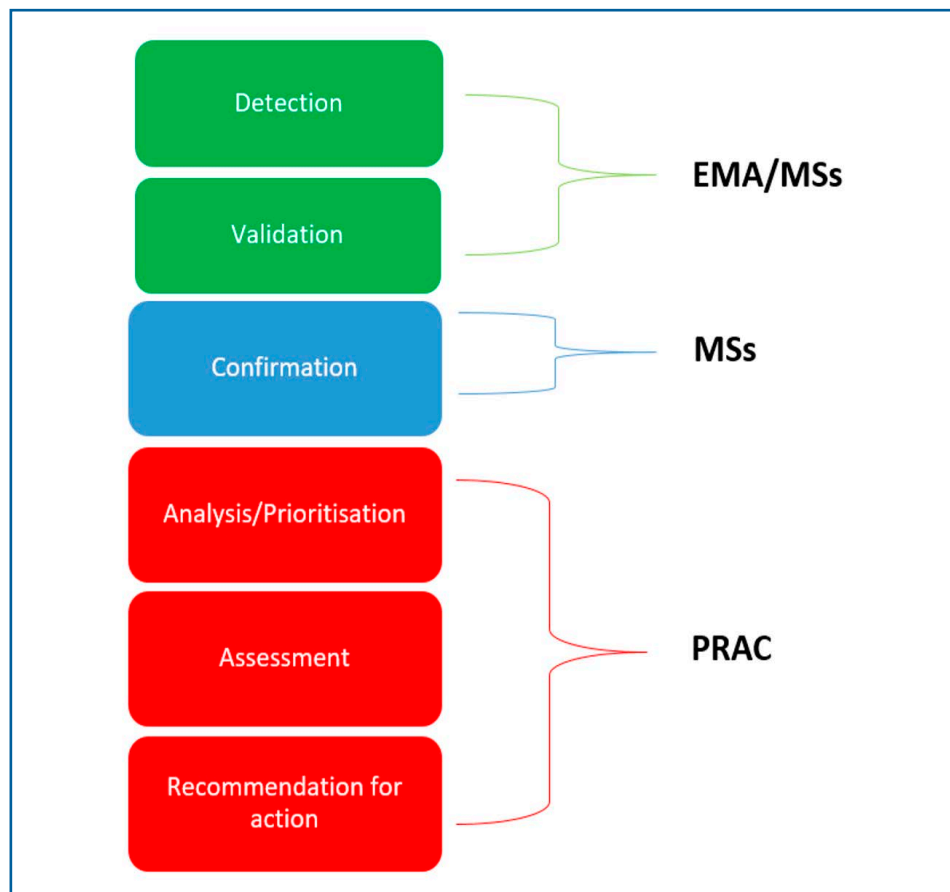


Tabella 1 Fonti di segnali, modificata da Meyboom et al. [3].

OSSERVAZIONI SU PAZIENTI (SEGNALI QUALITATIVI)
Segnalazione spontanea
Segnalazione aneddotica da letterature (<i>case report</i> e <i>case series</i>)
Regimi di monitoraggio intensivo e Siti sentinella
Monitoraggio degli eventi prescrittivi
Segnalazione stimolata (monitoraggio intensivo e addizionale)
Studi di follow-up
OSSERVAZIONI SU PAZIENTI (SEGNALI QUANTITATIVI)
Grandi fonti di dati su uso di farmaci e morbidità (inclusi gli approcci di record linkage)
Studi caso-controllo
Studi di follow-up
Monitoraggio degli eventi prescrittivi
Regimi di monitoraggio intensivo e Siti sentinella
Ampi sistemi di segnalazione spontanea
EVIDENZE SPERIMENTALI
Trial clinici
Esperimenti <i>in vivo</i>
Tossicologia su modelli animali

METODO CLASSICO QUALITATIVO – ANALISI “CASO PER CASO”

Prima che fossero disponibili gli attuali mezzi informatici, la rilevazione del segnale si basava esclusivamente su uno studio caso per caso, che implicava che ogni singolo caso di segnalazione spontanea di una sospetta ADR (reazione avversa da farmaco) fosse esaminata da un valutatore esperto. In particolare, quest’ultimo valutava la pro-

babilità che il quadro clinico fosse stato causato dal farmaco e cercava di comprendere se il caso fosse in qualche modo insolito o strano. Questa rilevazione qualitativa del segnale, basata esclusivamente sull'intelletto umano, mediante l'analisi caso per caso, ha dimostrato per lungo tempo, e continua tutt'ora a mostrare, il suo valore; tuttavia, tale approccio sta diventando sempre più dispendioso a causa dell'incremento del volume dei dati e quindi sta perdendo efficacia, specialmente quando si ha a che fare con associazioni complesse, come le sindromi da interazioni tra farmaci, o quando sono coinvolte numerose covariate.

Nel corso degli ultimi 10-15 anni, la disponibilità di potenti sistemi informatici ha stimolato lo sviluppo del rilevamento quantitativo del segnale basato sul concetto di disproporzionalità, ovvero un metodo che quantifica in che misura l'incidenza di una determinata condizione (come una reazione avversa) eccede un valore atteso sulla base delle informazioni che si hanno fino ad ora. Tuttavia, negli anni recenti è stato dimostrato che un approccio ibrido che utilizza misure quantitative in aggiunta all'analisi qualitativa rappresenta un ulteriore passo avanti nella rilevazione dei segnali in farmacovigilanza [4]. Di fatto, se da un lato una valutazione quantitativa è più efficiente e coerente, dall'altro l'analisi qualitativa dei casi basata su osservazioni nei pazienti (ad esempio attraverso la segnalazione spontanea) spesso è estremamente efficace nel rilevare dei segnali. Infatti, in casi eccezionali, una singola buona segnalazione (ad esempio un caso di rechallenge positivo) può già essere considerato di per sé un segnale valido [5].

METODO QUANTITATIVO – IL CONCETTO DI “DISPROPORZIONALITÀ”

Questa analisi, prevalentemente automatizzata, è condotta su un numero più elevato di casi e utilizza database e tecniche basati principalmente su modelli di epidemiologia analitica. Si tratta di metodi che possono essere applicati indifferentemente per la ricerca di un'associazione tra il farmaco e la reazione avversa, come l'evidenziazione di interazioni farmacologiche o sindromi correlate al farmaco o di fattori di rischio in categorie di pazienti ben definite. Diversi indici di disproporzionalità sono stati proposti e studiati, tra cui il *Proportional Reporting Ratio* (PRR), il *Reporting Odds Ratio* (ROR), l'analisi *Bayesian Propagation Neural Network* (BCPNN) del Centro OMS di Uppsala e altri. Tutti si basano sul concetto di disproporzionalità delle segnalazioni e rispondono alla domanda: “stiamo osservando un numero di casi che supera quelli che ci saremmo aspettati in base alle conoscenze precedenti?” Il PRR e il ROR sono due modelli che non richiedono strumenti particolari e possono anche essere calcolati facilmente. Dopo una preferenza iniziale per il PRR, è stato dimostrato che il ROR fornisce informazioni ugualmente utili per identificare i segnali e quindi l'EMA ha scelto di includere quest'ultimo strumento di calcolo nell'*EudraVigilance Data Analysis System* (EVDAS) [6, 7].

EVDAS

Il 22 novembre 2017, l'EMA ha lanciato una nuova versione di Eudravigilance, con funzionalità potenziate, necessaria per sostenere i cambiamenti significativi introdotti dalla legislazione europea in materia di farmacovigilanza, soprattutto in termini di requisiti per segnalare le reazioni avverse sospette. All'interno di questo nuovo sistema, EMA ha permesso ai MAH di accedere all'*EudraVigilance Data Analysis System* (EVDAS). Questo sistema è stato progettato per consentire agli utenti di analizzare i dati di sicurezza raccolti in EudraVigilance, permettendo di avere un quadro più consapevole sul profilo di sicurezza dei medicinali e, soprattutto, per fornire strumenti di rilevamento del segnale.

EVDAS analizza la banca dati delle segnalazioni con tecniche di disproporzionalità, principalmente ROR. Questi ultimi vengono riportati negli *Electronic Reaction Monitoring Reports* (eRMR). Inoltre, il sistema prevede anche la possibilità di accedere ai singoli casi (*report form*). Il 22 febbraio 2018 è iniziato un periodo pilota di un anno. Durante questo periodo, solo i MAH dei principi attivi contenuti nell'elenco dei medicinali sottoposti a monitoraggio aggiuntivo (definito il 25 ottobre 2017) sono stati obbligati a monitorare Eudravigilance e utilizzare EVDAS per eseguire la *signal detection*. Dopo un anno, l'EMA baserà la fase successiva di implementazione

sull'esperienza acquisita durante la fase pilota, rendendo obbligatorio l'uso di questo sistema per ciascun MAH. In particolare, EMA prevede che l'uso di questo strumento avanzato di rilevamento del segnale e analisi dei dati riguardo il profilo di sicurezza dei farmaci consentirà una migliore individuazione di problematiche di *safety* nuove o in evoluzione, consentendo quindi eventuali azioni rapide per salvaguardare la salute pubblica [8].

Scopo di questo studio è stato di confrontare i risultati, in termini di dati di *safety*, ottenuti con il metodo classico di rilevamento del segnale (caso per caso) rispetto al nuovo approccio di rilevamento del segnale attraverso l'uso di EVDAS, sottolineando e argomentando le principali differenze.

Metodi

Sono stati analizzati i dati di *safety* relativi a un farmaco specifico (X) di proprietà dell'azienda (Y). Come fonte delle segnalazioni è stata adottata la "Rete Nazionale di Farmacovigilanza (RNF)".

Nel periodo da 01.01.2018 a 30.06.2018 è stata eseguita una classica valutazione caso per caso, considerando solo le segnalazioni contenute nel database aziendale, seguendo quindi l'approccio classico di *signal detection*.

Il farmaco X è stato incluso da EMA nell'elenco dei farmaci sotto ulteriore monitoraggio; quindi anche per questo farmaco è iniziato il progetto pilota con EVDAS.

Nel secondo semestre del 2018 (da 01.07.2018 a 31.12.2018), dopo questo passaggio, sono stati applicati due approcci: quello tradizionale e quello sperimentale con EVDAS. Lo scopo è che le nuove ADR possibili emerse dalla prima analisi fossero controllate o confermate in un sistema mondiale come EVDAS al fine di ottenere una visione complessiva del profilo di sicurezza per il farmaco X.

In primo luogo, abbiamo sviluppato un protocollo di valutazione del rischio in cui è stata inclusa la frequenza della rilevazione del segnale relativa al farmaco X. In particolare, abbiamo deciso di continuare a monitorare il farmaco e di eseguire il rilevamento del segnale con una periodicità di 6 mesi. Successivamente, basandoci sui risultati ottenuti con EVDAS, saremmo poi stati in grado di determinare se confermare, prolungare o ridurre la frequenza di monitoraggio del farmaco X per l'anno successivo. I parametri selezionati per determinare la valutazione del rischio sono evidenziati nella **Figura 2**.

Risultati

Per quanto riguarda la prima analisi del segnale, abbiamo ricevuto 6 segnalazioni riguardanti il prodotto X nel periodo considerato. Nella tabella riepilogativa dei casi (**Tabella 2**), sono evidenziati gli eventi non presenti nell'SmPC (riassunto delle caratteristiche del prodotto) dei medicinali della società Y.

Quindi, abbiamo anche analizzato la letteratura internazionale (**Tabella 3**) al fine di trovare nuovi segnali dai Paesi di non commercializzazione della società Y o di confermare i segnali ottenuti in precedenza e discussi sopra.

Nel primo semestre del 2018 (01.01.2018 - 30.06.2018), i dati di vendita relativi alle singole confezioni di farmaci basati sul prodotto X dell'azienda Y sono i seguenti:

- Unità vendute con X (40 mg) 30 compresse: 191;
- Unità vendute con X (10 mg) 30 compresse: 18.828;
- Unità vendute di X (20 mg) 30 compresse: 25.760.

Al fine di identificare il numero di pazienti esposti, dobbiamo considerare che la durata media della terapia per tutte le dosi (40 mg, 20 mg e 10 mg) è di 1 cpr/giorno per 20 settimane (140 giorni).

Figura 2 Protocollo di valutazione del rischio per il farmaco X.

Tabella 2 ADR trovate attraverso lo screening della RNF nel primo semestre del 2018.

Tipologia di paziente/i	Codifica e fonte	ADR	SOC
Maschio (21 anni)	IT-002-18	Epistassi	Patologie respiratorie, toraciche e mediastiniche
Femmina (29 anni)	IT-023-18	Astenia	Patologie generali e condizioni relative alla sede di somministrazione
		Affaticamento	
		Malessero gastrointestinale	Patologie gastrointestinali
		Sonnolenza	Patologie del sistema nervoso
Maschio (16 anni)	IT-030-18	Insonnia	Disturbi psichiatrici
		Depressione	
		Pensieri negativi	
		Pianto	Patologie generali e condizioni relative alla sede di somministrazione
Maschio (19 anni)	IT-036-18	Cheratocono	Patologie dell'occhio
		Disturbi visivi	
Femmina (Adulta)	IT-058-18	Anedonia	Disturbi psichiatrici
		Sensibilità genitale ridotta	Patologie dell'apparato riproduttivo e della mammella
Femmina (17 anni)	IT-073-18	Mancanza di effetto farmacologico	Patologie generali e condizioni relative alla sede di somministrazione
		Perdita di capelli	Patologie della cute e del tessuto sottocutaneo

Tabella 3 ADR trovate attraverso lo screening della letteratura internazionale nel primo semestre del 2018.

SOC	ADR	L	U	S	NS	Nr. caso
Patologie dell'apparato riproduttivo e della mammella	Problemi nella raccolta follicolare	/	1	/	1	LIT18-008
	Problemi nella maturazione degli oociti	/	1	/	1	
Traumatismi, intossicazioni e complicazioni da procedura	Esposizione a farmaco durante la gravidanza	/	2	2	/	LIT18-025, LIT18-039
Disturbi psichiatrici	Depressione	1	/	/	1	LIT18-029

L: Listed, U: Unlisted, S: Serious, NS: Non serious

Complessivamente, il numero di pazienti che potrebbero essere stati esposti al farmaco X sono:

- 40 per X (40 mg) 30 compresse;
- 4.034 per X (10 mg) 30 compresse;
- 5.520 per X (20 mg) 30 compresse.

Come accennato in precedenza, dal secondo semestre del 2018 il farmaco X è stato incluso nella lista dei farmaci sottoposti a monitoraggio aggiuntivo. Pertanto, abbiamo eseguito sia la classica *signal detection* caso per caso che tramite EVDAS.

Partendo quindi dall'analisi caso per caso, abbiamo individuato le seguenti nuove ADR, elencate nelle **Tabella 4** e **5**.

Tabella 4 ADR individuate attraverso lo screening della RNF nel secondo semestre del 2018.

Tipologia di paziente/i	Codifica e fonte	ADR	SOC
Femmina (29 anni)	IT-096-18	Perdita di peso non intenzionale	Disturbi del metabolismo e della nutrizione
		Bilirubina indiretta aumentata	Esami diagnostici
Maschio (non riportato)	IT-118-18 (Letteratura)	Dolore	Patologie generali e condizioni relative alla sede di somministrazione
Non noto (45 anni)	IT-171-18 (Spontaneo)	Nausea	Patologie gastrointestinali
		Tachicardia	Patologie cardiache
		Ipoacusia	Patologie dell'orecchio e del labirinto

Tabella 5 ADR individuate attraverso lo screening della letteratura internazionale nel secondo semestre del 2018.

SOC	ADR	L	U	S	NS	Nr. caso
Disturbi psichiatrici	Disturbo ossessivo-compulsivo	1	/	1	/	LIT18-107
Patologie della cute e del tessuto sottocutaneo	Dermatosi neutrofila	/	1	1	/	LIT18-108

L: Listed, U: Unlisted, S: Serious, NS: Non serious.

Nel secondo semestre del 2018 (01.07.2018 - 31.12.2018) i dati di vendita delle singole confezioni dei farmaci a base di X sono i seguenti:

- Unità vendute di X (40 mg) 30 compresse: 96;
- Unità vendute con X (10 mg) 30 compresse: 18.527;
- Unità vendute con X (20 mg) 30 compresse: 24.047;
- Unità vendute di X (30 mg) 30 compresse: 1758.

In questo caso, il farmaco è anche stato commercializzato per la prima volta in formulazione (40 mg) 30 compresse. Utilizzando gli stessi criteri sopra menzionati per il primo semestre, il numero di pazienti che possono essere stati esposti al farmaco X è di circa:

- 20 per X (40 mg) 30 compresse;
- 3.970 per X (10 mg) 30 compresse;
- 5.152 per X (20 mg) 30 compresse;
- 12 per X (30 mg) 30 compresse.

Per valutare il possibile verificarsi di un nuovo segnale utilizzando i dati forniti da EVDAS, è stato utilizzato il seguente approccio:

1. Creare un elenco di eventi imprevedibili e rari e di tutti i casi che riguardano situazioni speciali (uso durante la gravidanza/l'allattamento, in pazienti pediatrici, uso improprio, abuso, ecc.) riscontrati durante il periodo di riferimento dell'analisi (secondo semestre del 2018) e applicabili al medicinale oggetto di revisione.
2. Valutare il numero di volte in cui ciascuno di questi eventi è stato segnalato durante il periodo di riferimento dell'analisi del segnale (che, per essere significativo, deve essere ≥ 3).
3. Valutare per ciascun evento selezionato, che si è verificato almeno 3 volte durante il periodo di riferimento, il valore medio mondiale del ROR nei eRMR. Per fare ciò, è stato effettuato prima l'accesso alla piattaforma con le credenziali della *Qualified Person Responsible For Pharmacovigilance* (QPPV). Quindi, l'intervallo di tempo per l'analisi dei dati è stato selezionato scegliendo l'opzione ad-hoc e la sostanza principio attivo (X) è stata selezionata dall'elenco fornito. Successivamente, è stato scaricato il file e-RMR (come spiegato dal manuale per gli utenti di EudraVigilance [9]) e sono state raccolte le ADR. A questo punto, è stato possibile calcolare gli ROR, ed è stato considerato l'elenco dei possibili segnali. Il valore di ROR per essere statisticamente significativo deve essere ≥ 1 .
4. Verificare che l'evento che soddisfa i requisiti 2 e 3 rientri nella IME list (*Important Medical Event list*, utilizzata per facilitare la classificazione delle sospette reazioni avverse) e/o nella DME list (*Designated Medical Event list*, contenente condizioni mediche intrinsecamente gravi e spesso farmaco-correlate).

Nel caso in cui si verificano le condizioni di cui sopra, vi sarà un segnale di disproporzionalità (SDR, *Signal of Disproportion Reporting*) che dovrà poi essere ulteriormente esaminato.

I valori dell'ROR per ogni evento inatteso rilevato sono stati i seguenti:

- Perdita di peso involontaria: ROR = 0,53
- Aumento della bilirubina indiretta: ROR = 0,91
- Dolore: ROR = 0,36
- Tachicardia: ROR = 0,32.

Poiché gli eventi, non riportati nel SmCP della Società Y, non soddisfano i requisiti 2 e 4 dell'elenco sopra (non compaiono almeno 3 volte nella serie del caso e non sono elencati nella lista IME o nella lista DME), possiamo concludere che nessun possibile nuovo segnale è stato evidenziato da questa analisi.

Alla luce di quanto è stato descritto fino ad ora, possiamo concludere che al momento della stesura di questo studio, il rapporto rischio/beneficio del farmaco X è rimasto invariato.

Conclusioni

Considerando il nostro studio, il confronto tra la *signal detection classica* (caso per caso) e EVDAS mostra che non ci sono particolari differenze in termini di dati di safety e risultati ottenuti.

Sebbene non abbiamo raccolto nuovi potenziali segnali di sicurezza, EVDAS rappresenta un passo avanti nel presente e nel futuro per l'esecuzione della *signal detection* da parte delle aziende farmaceutiche. Infatti, l'aspetto più rilevante è che è possibile ottenere un dato per una ADR che include tutti i casi mondiali relativi a un prodotto specifico. Questo aspetto è considerato il più rilevante e consente a questo nuovo sistema di essere più completo di una classica *signal detection* caso per caso. Di fatto, con quest'ultimo metodo, non è possibile avere una visione globale e completa dei segnali ed è possibile solamente confrontare i dati del database di sicurezza di un'azienda con la letteratura internazionale, con il rischio di generare falsi positivi.

Tuttavia, il sistema EVDAS è caratterizzato anche da significativi inconvenienti. Innanzi tutto, è ancora un progetto pilota e, di conseguenza, non ci sono linee guida chiare e definite su come utilizzare correttamente questo strumento. Pertanto, ogni azienda può decidere la frequenza e il modo di valutare i dati ottenuti. Inoltre, i risultati riportati nell'eRMR sono calcolati considerando solo il principio attivo e non la forma farmaceutica o il dosaggio. Pertanto, anche se un evento si distingue come un nuovo possibile segnale, potrebbe non essere correlato al farmaco specifico o alla formulazione prodotta da un'azienda. Questi due aspetti rendono EVDAS un sistema poco consistente al momento. Tuttavia, una volta terminata la fase pilota, verranno apportate ulteriori modifiche, rendendo EVDAS lo strumento futuro per la rilevazione dei segnali.

Bibliografia

- [1] Guideline on good pharmacovigilance practices (GVP) - Module IX (Rev 1) EMA/827661/2011 Rev 1
- [2] Disponibile dal sito: <https://www.primevigilance.com/resources/complexities-drug-safety/signal-detection-in-pharmacovigilance>
- [3] Meyboom RHB, Egberts ACG, Edwards IR, et al. Drug-Safety, 1997; 16: 355. <https://doi.org/10.2165/00002018-199716060-00002>
- [4] Egberts TCG. Drug-Safety, 2007; 30: 607.
- [5] Meyboom RHB, Gribnau FWJ, Hekster YA, et al. Characteristics of topics in pharmacovigilance in the Netherlands. Clin Drug Invest, 1996; 4: 207-19.
- [6] Wisniewski FZ, Good Signal Detection Practices: Evidence from IMI PROTECT, Drug Saf. 2016; 39: 469-490.
- [7] Disponibile dal sito: http://www.sefap.it/web/upload/GIFF2017-3_13_28.pdf.
- [8] Signal Management. European Medicines Agency. Disponibile dal sito: http://www.ema.europa.eu/ema/index.jsp?curl=pages/regulation/general/general_content_000587.jsp&mid=WC0b01ac0580727d1b
- [9] Disponibile dal sito: https://www.ema.europa.eu/documents/regulatory-procedural-guideline/eudravigilance-user-manual-marketing-authorisation-holders_en.pdf