Sindrome del dolore vescicale - cistite interstiziale: una recente scoperta dellâ€™intelligenza artificiale

Prof.ssa Alessandra Graziottin
Direttore del Centro di Ginecologia e Sessuologia Medica
H. San Raffaele Resnati, Milano

Commento a:
Okui N.
Unsupervised Machine Learning reveals a vulvodynia-predominant subtype in bladder pain syndrome/interstitial cystitis
Cureus. 2024 Jun 18;16(6):e62585. doi: 10.7759/cureus.62585. PMID: 39027744; PMCID: PMC11257649

La sindrome del dolore vescicale - cistite interstiziale (BPS/IC) è una patologia cronica caratterizzata da dolore pelvico e sintomi urinari che compromettono gravemente la qualità della vita. Molto spesso si presenta in concomitanza con la vulvodinia, con una prevalenza stimata fra il 30% e il 50% a seconda degli studi. Tuttavia, nella pratica clinica corrente si utilizzano questionari standardizzati che tendono a focalizzarsi quasi esclusivamente sulla vescica, senza identificare il dolore vulvare associato.
Questo studio si è posto l'obiettivo di applicare l'intelligenza artificiale per classificare le pazienti in sottogruppi omogenei (cluster) e verificare se fosse possibile intercettare la presenza di vulvodinia basandosi sui profili sintomatologici. In particolare, è stato utilizzato un algoritmo di Machine Learning non supervisionato, ossia che analizza dati grezzi con l'obiettivo di identificare autonomamente schemi, modelli o raggruppamenti nascosti, senza aiuto umano.

Come ha lavorato l'intelligenza artificiale

I ricercatori hanno arruolato 123 pazienti con BPS/IC e 64 controlli sani. Per analizzare i dati clinici, l'algoritmo ha operato secondo precisi passaggi matematici e computazionali:
input dei dati multidimensionali: l'IA ha preso in esame le risposte fornite dalle pazienti a sette questionari clinici validati, che misurano l'intensità del dolore (NRS-11), i sintomi e l'impatto della cistite interstiziale (ICSI, ICPI), l'urgenza e la frequenza urinaria (PUF), i sintomi della vescica iperattiva (OABq SF, OABSS) e i disturbi del pavimento pelvico (PFDI-20);
calcolo della similarità (distanza euclidea): l'algoritmo ha calcolato la distanza matematica fra i dati delle pazienti per quantificare il grado di somiglianza o differenza tra i diversi quadri sintomatologici (box 1);
raggruppamento gerarchico (metodo di Ward): sulla base delle distanze calcolate, l'algoritmo ha iniziato a unire le pazienti con profili simili all'interno di una struttura ad albero chiamata "dendrogramma", procedendo dal basso verso l'alto (box 2);
ottimizzazione statistica (metodo del gomito - Elbow Method): per individuare il numero ottimale di cluster in cui dividere la coorte, l'IA ha analizzato la variazione della somma degli errori al quadrato (SSE). La curva statistica ha mostrato un'inflessione netta nel passaggio da due a tre gruppi (punto di gomito), stabilendo matematicamente che 3 era il numero ottimale di cluster (box 3);
assegnazione e visualizzazione grafica: l'algoritmo ha assegnato formalmente le pazienti ai tre cluster

individuati, mentre l'analisi delle componenti principali (PCA) ha ridotto la complessità dei dati per permettere la visualizzazione dei tre gruppi su un grafico bidimensionale.
Infine, tutte le pazienti sono state sottoposte indipendentemente al **Vulvodynia Swab Test** (il test del cotton fioc per valutare la sensibilità vestibolare). Questo dato clinico oggettivo è stato volutamente escluso dall'analisi iniziale dell'IA per verificare se la macchina fosse in grado di far emergere la vulvodinia basandosi unicamente sui questionari urinari e pelvici generici.

I risultati prodotti dall'IA

L'intelligenza artificiale ha suddiviso la popolazione con BPS/IC in tre cluster fenotipici ben differenziati:

cluster 0 (sottotipo a predominanza di dolore vescicale grave): pazienti con una sintomatologia urinaria e dolorosa vescicale estremamente severa (punteggi ICSI, ICPI e PUF molto alti), ma con una sensibilità vulvare al swab test del tutto normale e sovrapponibile a quella del gruppo di controllo sano;
cluster 1 (sottotipo a predominanza di dolore vescicale moderato): pazienti che presentano il classico quadro di dolore vescicale e frequenza urinaria, ma con un livello di gravità intermedio e più sfumato rispetto al Cluster 0;
cluster 2 (sottotipo a predominanza di vulvodinia): rappresenta la scoperta clinica più rilevante dello studio. Le pazienti inserite in questo gruppo mostrano punteggi minimi o moderati di dolore prettamente vescicale e di vescica iperattiva, talvolta vicini a quelli del gruppo sano. Al contrario, presentano punteggi al Vulvodynia Swab Test straordinariamente elevati e statisticamente superiori rispetto a tutti gli altri soggetti ($P < 0.001$).
L'algoritmo ha quindi dimostrato l'esistenza di **un fenotipo specifico di BPS/IC** in cui la componente dolorosa non è intravescicale, bensì extravescicale e localizzata a livello del vestibolo vulvare.

Perché questi risultati sono importanti?

I risultati generati dal Machine Learning offrono risvolti clinici di fondamentale importanza per l'ottimizzazione dei percorsi diagnostici e terapeutici del dolore cronico:

evidenziano i limiti dei soli questionari: lo studio prova che i questionari clinici standard incentrati sulla vescica non sono in grado di catturare la presenza di vulvodinia. Basarsi solo su di essi comporta il rischio elevato di non riconoscere la comorbilità vulvare, portando a trattamenti inappropriati o perlomeno incompleti;
rendono imprescindibile il Vulvodynia Swab Test: viene scientificamente dimostrato che l'esame obiettivo tramite swab test è insostituibile per una corretta mappatura fenotipica. Deve quindi essere integrato sistematicamente nella routine clinica delle pazienti con dolore pelvico cronico;
guidano la transizione verso terapie personalizzate (Tailored Approach): identificare le pazienti del Cluster 2 permette di evitare terapie vescicali invasive e inefficaci, indirizzandole immediatamente verso protocolli terapeutici specifici per la vulvodinia;
validano l'efficacia dei trattamenti laser: lo studio si collega e dà forza scientifica alle evidenze sull'efficacia clinica del laser Erbium YAG non ablativo (VEL) e dei protocolli combinati (Erbium/Neodymium) applicati a livello vaginale e vulvare. Questi trattamenti non solo riducono il dolore vulvare locale, ma migliorano significativamente, per via riflessa, gli stessi sintomi urinari e la salute pelvica complessiva delle pazienti affette da questo specifico sottotipo.

Box 1. La distanza matematica (o euclidea)

La distanza matematica in un contesto di Machine Learning è un concetto che permette di tradurre la **differenza tra due persone** (o due casi clinici) in **un valore numerico**.

Quando i dati derivano da test diversi, il problema principale è che le unità di misura sono eterogenee (ad esempio, il punteggio di una scala del dolore da 0 a 10 non è direttamente confrontabile con una scala di frequenza urinaria, che ha un'unità di misura diversa).

Ecco come viene risolto questo problema nel processo descritto nell'articolo

Fase 1. La normalizzazione dei dati

Prima di calcolare qualsiasi distanza, i dati vengono standardizzati. Questo è il passaggio fondamentale: l'algoritmo converte i punteggi grezzi di ogni test in una scala comune.

Fase 2. Lo spazio multidimensionale

Una volta standardizzato, ogni paziente viene trasformato in **un punto in uno spazio multidimensionale**. Se fossero stati usati solo 2 test, l'algoritmo avrebbe prodotto un grafico bidimensionale; poiché invece i test utilizzati sono 7, è stato prodotto un grafico a 7 dimensioni (assai complesso da visualizzare, ma perfettamente maneggiabile in termini matematici).

Fase 3. La distanza euclidea

La distanza euclidea è, in sostanza, la "linea d'aria" che separa due punti in questo spazio multidimensionale. Se due pazienti hanno dato risposte molto simili in tutti i test, i loro punti saranno vicini nello spazio (distanza euclidea piccola); se hanno sintomi completamente diversi, i loro punti saranno lontani (distanza euclidea elevata).

Grazie a questa operazione, l'intelligenza artificiale non legge più i questionari come testi, ma come **coordinate spaziali**. In tal modo, quando l'algoritmo raggruppa le pazienti in cluster, non sta facendo altro che cercare le "nuvole" create dai punti che sono più vicini fra loro.

La scoperta del **cluster 2** (quello a predominanza di vulvodinia) è avvenuta proprio perché l'IA ha rilevato che, nonostante le differenze nei singoli test, un gruppo di pazienti occupava costantemente una "regione" dello spazio in cui le distanze reciproche erano minime, rivelando una coerenza clinica invisibile all'occhio umano.

In sintesi, la distanza matematica è lo strumento che trasforma la complessità dei sintomi in una mappa geografica dove i pazienti con lo stesso disturbo finiscono per trovarsi nello stesso "quartiere".

Box 2. Il dendrogramma e metodo del gomito

Nel dendrogramma, l'asse verticale (la "distanza") rappresenta **il grado di diversità o variabilità** fra i gruppi che si stanno unendo.

Nella parte bassa del dendrogramma, la distanza è prossima allo zero. Questo significa che i casi (e le pazienti) sono estremamente simili tra loro. Nella parte alta, invece, la distanza aumenta sino a diventare massima. Qui l'algoritmo è costretto a unire gruppi che sono sempre meno simili tra loro. La variabilità (o distanza) cresce perché sta aggregando pazienti che hanno profili clinici più distanti.

Perché questo è cruciale per la ricerca? Nel grafico contenuto nello studio, le ramificazioni

basse indicano che esistono piccoli gruppi di pazienti con una coerenza interna fortissima. Quando il ricercatore usa il "metodo del gomito", cerca il punto di equilibrio in corrispondenza del quale non ha più senso salire longo il grafico e aggregare ulteriormente, perché si perderebbero le differenze cliniche fondamentali.

Box 3. Il metodo del gomito

Il metodo del gomito, in altre parole, è lo strumento matematico che permette di rispondere a una domanda cruciale: quanti gruppi esistono realmente in questo insieme di pazienti? Senza un criterio oggettivo, suddividere le persone in cluster sarebbe arbitrario. Questo metodo fornisce un rigore statistico al processo.

Ecco come funziona in termini semplici:

1. Il concetto di Sum of Squared Errors (SSE)

Per capire il metodo, dobbiamo prima definire l'errore (SSE), che misura quanto le pazienti all'interno di un cluster sono diverse, ossia distanti, dal centro del cluster stesso (detto "centroide"). Se la SSE è bassa, significa che le pazienti in quel gruppo sono molto simili tra loro (cluster compatto); se invece la SSE è alta, le pazienti sono molto distanti dal centro (cluster disperso).

2. Il meccanismo: aggiungere gruppi

L'algoritmo viene eseguito più volte, aumentando progressivamente il numero di cluster da 1 fino a 10. Quando si passa da 1 a 2 cluster, la SSE cala drasticamente, perché si è articolato un gruppo molto eterogeneo in due più piccoli e coerenti. Continuando ad aumentare i cluster, la SSE continuerà a scendere, ma sempre più lentamente.

3. Perché si chiama "metodo del gomito"

Se si tracciano su un grafico bidimensionale i valori della SSE (in asse Y) rispetto al numero di cluster (in asse X), si ottiene una curva che scende rapidamente per poi appiattirsi. La forma che si ottiene assomiglia a un braccio piegato: la parte iniziale è l'avambraccio, la parte piatta è la braccio vero e proprio, e il punto di piega è il gomito, che rappresenta il numero ottimale di cluster: esso è infatti il punto oltre il quale aggiungere ulteriori gruppi non porterebbe più un vantaggio significativo in termini di riduzione dell'errore (ovvero, non si guadagnerebbe più ulteriore e significativa coerenza interna).

4. Applicazione nel caso specifico

Nel caso dello studio che abbiamo esaminato, l'analisi ha mostrato che:

-passando da 2 a 3 cluster, si verifica ancora un calo significativo della SSE;

-superato il cluster 3, la curva inizia a diventare molto più piatta: creare un quarto o un quinto cluster non avrebbe offerto una distinzione clinica rilevante tra le pazienti.

Pertanto, il numero 3 è stato selezionato perché rappresenta il miglior equilibrio matematico fra **semplicità** (non avere troppi gruppi) e **precisione** (avere tutti i gruppi che servono, ossia che riflettono le reali differenze cliniche tra le pazienti).