

ChatGPT-mania in Medicina di Laboratorio? Non è oro tutto quello che luccica

Davide Negrini, Giuseppe Lippi

Sezione di Biochimica Clinica, Scuola di Medicina, Università di Verona

Caro Editore,

ChatGPT (OpenAI Inc., San Francisco, CA, USA) può superare brillantemente il test USMLE (United States Medical Licensing Examination), ma non è in grado di interpretare un referto completo di laboratorio, ovvero una operazione estremamente comune ed eseguita quotidianamente dai medici di tutto il mondo.

Questa è in estrema sintesi il paradosso di come funzionano i recentissimi sistemi di Intelligenza Artificiale (AI) generativa, che sono in grado di analizzare il linguaggio naturale (NLP, Natural Language Processing), analizzati in dettaglio nell'articolo di Cadamuro et al. (1), che rappresenta il punto di vista del gruppo di lavoro sull'Intelligenza Artificiale della Federazione Europea di Chimica Clinica e Medicina di Laboratorio (WG-AI, EFLM).

Lo studio condotto dal gruppo di lavoro europeo ha chiesto a ChatGPT v.4 di interpretare 10 referti verosimili, ottenendo esiti discreti quando i singoli esami erano analizzati indipendentemente, ma evidenziando l'incapacità del sistema di "unire i puntini" (ovvero, vedere il quadro d'insieme) tipica di un essere umano addestrato in tal senso. L'idea originale è quella di vedere come avrebbe risposto ChatGPT a ciò che un paziente generico avrebbe potuto chiedere, ovvero di interpretare i risultati delle proprie analisi del sangue, rilevando se vi fosse motivo per preoccuparsi, oppure per poter rimandare il consulto con il medico.

Anche se la maggior parte delle AI generative e di NLP è tutt'ora in versione dimostrativa ad accesso controllato, poco manca (ma in alcuni casi è già realtà) alla distribuzione su larga scala, identificando quindi un potenziale rischio per la salute dei pazienti che potrebbero affidarsi ai successori del, già in precedenza problematico, "Dottor Google". Questi strumenti sarebbero inoltre da considerare nella macrocategoria dei dispositivi medici, con tutte le conseguenze del caso. Viene infine riportato nell'articolo che fortunatamente, per ora, ChatGPT a tutte le richieste mediche è solitamente cauto, raccomandando una visita con un professionista medico/sanitario, senza lanciarsi in indicazioni di ulteriori esami o terapie da seguire.

A differenza della Medicina di Laboratorio, che malgrado sia molto legata all'avanzamento tecnologico si sta ora trovando un passo indietro sull'argomento, in altre discipline mediche l'assistenza dell'AI è attualmente molto sfruttata. Un esempio importante è quello della diagnostica per immagini, dove la presenza di assistenza all'interpretazione delle immagini, a partire dal periodo COVID-19 (2) in poi è un dato di fatto (3). Questi sistemi consentono di evidenziare ed enfatizzare eventuali anomalie nelle immagini radiografiche pre-valutate dall'AI, velocizzando i processi di analisi da parte dei medici, migliorando infine le prestazioni di identificazione di eventuali lesioni. Interessante anche assistere all'emergere o allo sviluppo di tecniche di AI in ambiti come la medicina dello sport e riabilitativa (4).

Nella lista dei dispositivi medici approvati dall'americana FDA (Food and Drug Administration) (5), solo pochi sono di interesse per la Medicina di Laboratorio (il 75% sono di radiologia, ovvero 392 su 521 al momento in cui è stato scritto questo articolo). I pochi esistenti riguardano algoritmi per analisi digitalizzata dello striscio di sangue, per l'identificazione delle cellule e l'interpretazione ematologica. Nella categoria ematologia, ci sono ad oggi 15 dispositivi medici elencati che, sommati agli altri 11 categorizzati tra chimica clinica e microbiologia, rappresentano meno del 5% del totale dei dispositivi di AI approvati dalla FDA.

Come conclude l'articolo del gruppo di lavoro EFLM (1), il problema non è "se" verrà il momento in cui le AI saranno in grado di interpretare i referti di laboratorio, ma semmai "quando", ovvero dopo che vi siano dei "dataset di training" sufficientemente completi da poterlo fare, raccomandando che tutti i professionisti di laboratorio si tengano informati e stiano al passo con l'avanzamento tecnologico relativo a queste dirompenti novità nella diagnostica medica.

A corollario di quanto espresso, è poi utile ribadire come appaia forse eccessiva l'enfasi con cui attualmente si vuole raffigurare il pericolo che l'AI generativa rappresenta in tema di editoria scientifica (6) e appropriatezza delle risposte a quesiti clinici, almeno per ora. Si riporta un esempio paradigmatico nella Tabella 1, in cui un utente generico chiede al sistema cosa siano le troponine cardiache.

Corrispondenza a: Davide Negrini, Sezione di Biochimica Clinica, Scuola di Medicina, Università di Verona, p.le Scuro 10, Verona
E-mail: davide.negrini@aovr.veneto.it

Ricevuto: 09.05.2023

Revisionato: 09.05.2023

Accettato: 09.05.2023

Pubblicato on-line: 23.05.2023

DOI: 10.19186/BC_2023.039

Tabella 1

Esempio paradigmatico di come l'adattatività di ChatGPT non sia ancora del tutto attendibile.

User: cosa sono le troponine

ChatGPT: Esistono tre diverse subunità di troponina: TnI, TnC e TnT. La subunità TnI è quella più comunemente usata come marcatore di danno miocardico, in quanto viene rilasciata nel sangue quando si verifica una lesione del muscolo cardiaco.

User: La TnI non è una subunità ma una proteina autonoma. Come marcatore di danno miocardico sono usate comunemente sia la TnT che la TnI.

ChatGPT: Hai assolutamente ragione, mi scuso per la mia imprecisione. La TnI è una proteina autonoma e, insieme alla TnT, è una delle subunità della troponina. Entrambe le subunità sono comunemente utilizzate come marcatori di danno miocardico, in quanto vengono rilasciate nel sangue in seguito a lesioni del muscolo cardiaco.

User: cosa sono le troponine

ChatGPT: Esistono tre diverse subunità di troponina: TnI, TnC e TnT. Le troponine sono diventate importanti marcatori di danno cardiaco, in particolare la subunità TnI, che viene rilasciata nel sangue quando il muscolo cardiaco subisce una lesione.

La risposta iniziale presenta due sostanziali imprecisioni, la prima relativa al fatto che le troponine non sono "subunità della troponina", ma ognuna è una proteina autonoma, così come il fatto che tra esse la Troponina I sia quella che è misurata per la diagnosi del danno miocardico, perché viene rilasciata in circolo dopo il danno (come se la Troponina T e paradossalmente anche la Troponina C non lo fossero). Quando l'utente cerca di correggere il sistema, riceve una risposta positiva. Nondimeno, riproponendo la medesima domanda, il sistema persevera in entrambi gli errori originari, ribadendo che le troponine sono "subunità" e che la Troponina I è quella (più) usata per la diagnosi del danno miocardico perché è rilasciata in circolo dopo il danno. Come usavano dire gli antichi romani "*errare umanum est, perseverare autem diabolicum*". Oppure... "*non è oro tutto ciò che luccica*" (7).

CONFLITTO DI INTERESSE

Nessuno

BIBLIOGRAFIA

1. Cadamuro J, Cabitza F, Debeljak Z, De Bruyne S, Frans G, Perez SM, et al. Potentials and pitfalls of ChatGPT and natural-language artificial intelligence models for the understanding of laboratory medicine test results. An assessment by the European Federation of Clinical Chemistry and Laboratory Medicine (EFLM) Working Group on Artificial Intelligence (WG-AI). *Clin Chem Lab Med* 2023;61:1158-66.
2. Negrini D, Danese E, Henry BM, Lippi G, Montagnana M. Artificial intelligence at the time of COVID-19: who does the lion's share? *Clin Chem Lab Med* 2022;60:1881-6.
3. Wang S, Cao G, Wang Y, Liao S, Wang Q, Shi J, et al. Review and prospect: artificial intelligence in advanced medical imaging. *Front Radio* 2021;1:781868.
4. Artificial Intelligence and Machine Learning (AI/ML)-Enabled Medical Devices. FDA 2022. <https://www.fda.gov/medical-devices/software-medical-device-samd/artificial-intelligence-and-machine-learning-aiml-enabled-medical-devices> (ultimo accesso maggio 2023).
5. Artificial intelligence innovation: Leading companies in AI-assisted radiology for the medical devices industry. Medical Device Network 2023. <https://www.medicaldevice-network.com/data-insights/innovators-ai-ai-assisted-radiology-medical-devices/> (ultimo accesso maggio 2023).
6. Negrini D, Lippi G. Generative Artificial Intelligence in (Laboratory) Medicine: friend or foe?. *Biochim Clin* 2023;47:259-65.
7. Shakespeare W. Il mercante di Venezia.