

RUBRICA: LA MEDICINA PRATICA**Camilla Borea**

Laureata Cdl Infermieristica Università Degli Studi Di Cagliari

Pierpaolo Pateri

Presidente Onorario Opi Cagliari

Claudio Pirarba

Infermiere Azienda Ospedaliero Universitaria Di Cagliari

Claudio Mameli

Tutor Meds24/C Cdl Infermieristica Università Degli Studi Di Cagliari

Giuseppe Zingaro

Infermiere Azienda Ospedaliero Universitaria Di Cagliari

Cesar Ivan Aviles Gonzalez

Professore Associato MEDS/24C Università Degli Studi Di Enna Kore

Keywords:

triage, intelligenza artificiale

Corresponding author:

Giuseppe Zingaro

Email: giu.zingaro@gmail.com

Contributions:

the author(s) made a substantial intellectual contribution, read and approved the final version of the manuscript, and agreed to be accountable for all aspects of the work.

Conflict of interest:

the author(s) declare no potential conflict of interest.

Ethics approval and consent to participate:

not applicable.

Patient consent for publication:

not applicable.

Funding:

none.

Availability of data and materials:

data and materials are available from the corresponding author upon request.

Received: 2 February 2025

Accepted: 20 March 2025

APPLICAZIONI DELL'INTELLIGENZA ARTIFICIALE IN CONTESTI DI TRIAGE INFERMIERISTICO: UNA REVISIONE SISTEMATICA

ABSTRACT

Introduzione:

Background: Triage is a nursing competence aimed at assigning a priority code to patients presenting to the emergency department and assessing their potential risk evolution based on observable and reported symptoms. The triage nurse must be highly skilled and trained according to existing national regulations. Triage is a crucial moment to ensure patient management and define the order of access to treatment. Artificial intelligence (AI) is having an increasing global impact, particularly in supporting healthcare activities. This systematic review aimed to evaluate the effectiveness of AI-based systems in the nursing triage process compared to traditional methods.

Methods:

The foreground question was formulated using the PICO method. The research was conducted using three databases (PubMed, Cinahl, Scopus) following PRISMA guidelines. To include as many relevant studies as possible, five different keyword combinations were used (nurs*, triage, artificial intelligence). Studies published between 2019 and 2024, in both Italian and English, were considered for inclusion.

Results:

A total of 12 articles were included. More than half of the selected studies present very recent data. The extracted data primarily focused on study objectives, methodologies, measured outputs, and obtained results. In each included study, the findings were derived from a comparison between the performance of the described, proposed, or developed AI technologies and traditional triage procedures. The main limitation of this review is the limited availability of scientific literature on the subject.

Conclusions:

The use of artificial intelligence in nursing triage is a rapidly evolving field with evident potential, particularly in optimizing nursing workflow. However, as this area is still expanding, most AI-based applications described and studied require further testing and refinement before being widely implemented. Overall, the findings from the reviewed studies are highly promising.

RIASSUNTO

Background:

Il triage è una competenza infermieristica volta all'attribuzione di un codice priorità ai pazienti che si presentano in pronto soccorso e alla valutazione del potenziale rischio evolutivo sulla base dei sintomi osservabili e dei sintomi riferiti. L'infermiere di triage deve essere esperto e deve essere formato secondo le normative nazionali esistenti. Il triage è un momento fondamentale per garantire la presa in carico dei pazienti e per definire l'ordine di accesso al trattamento. L'intelligenza artificiale sta avendo un impatto crescente a livello globale, soprattutto a supporto delle attività sanitarie. Questa revisione sistematica mirava a valutare l'efficacia dell'utilizzo di sistemi basati sull'intelligenza artificiale (IA) durante il processo di triage infermieristico rispetto ai metodi tradizionali.

Metodi:

Il quesito di foreground è stato composto attraverso il metodo PICO. La ricerca è stata condotta con l'utilizzo di tre database (PubMed, Cinahl, Scopus) sulla base delle linee guida PRISMA. Per includere più articoli possibili sono state utilizzate cinque diverse combinazioni delle keywords stabilite (nurs*, triage, artificial intelligence). Sono stati identificati studi pubblicati in un range temporale che va dal 2019 al 2024, in italiano e inglese.

Risultati:

Sono stati inclusi un totale di 12 articoli. Più della metà degli studi inclusi esprimono dati molto recenti. I dati estrapolati dagli studi riguardano principalmente l'obiettivo, i metodi utilizzati, gli outputs misurati e i risultati ottenuti. In ogni articolo incluso, i risultati sono stati ottenuti dal confronto con le prestazioni delle tecnologie descritte, proposte o sviluppate, con le procedure di triage tradizionali. Il limite principale della ricerca risiede nella limitatezza di letteratura scientifica disponibile sull'argomento.

Conclusioni:

L'utilizzo dell'intelligenza artificiale in contesti di triage infermieristico è un campo fortemente in via di sviluppo e ha un potenziale evidente, soprattutto se utilizzato per ottimizzare il lavoro infermieristico. Tuttavia, essendo in forte accrescimento, quasi tutte le applicazioni basate sull'IA descritte e studiate necessitano di ulteriori test e miglioramenti prima di poter essere utilizzate. Nel complesso, i risultati provenienti dagli studi sono molto promettenti.

INTRODUZIONE

Oggigiorno, l'intelligenza artificiale è oggetto di molteplici discussioni, a causa anche del suo sempre maggiore impiego in numerosi ambiti, incluso quello sanitario.

Il triage è una competenza infermieristica in cui viene attribuita una priorità di trattamento al paziente attraverso un processo decisionale complesso che ha origine a partire da una valutazione minuziosa che i professionisti sanitari svolgono sul paziente e sulla sintomatologia da esso riferita.

Nel corso della mia esperienza di tirocinio svolta presso il Pronto Soccorso del Policlinico Universitario "D. Casula", ho avuto modo di entrare in contatto con quelli che sono i processi operativi che si trovano "dietro le quinte" dei dipartimenti d'emergenza.

Successivamente a questa esperienza, ho riflettuto sul potenziale che l'intelligenza artificiale potrebbe avere come supporto all'attività infermieristica che viene svolta in triage.

Sebbene l'interesse riguardo il tema sia sempre più crescente, la letteratura risulta poco fornita di studi che valutino il confronto tra le prestazioni dei sistemi basati sull'IA e i metodi tradizionali basati su scale validate a livello globale e sul giudizio clinico svolto dagli infermieri.

L'obiettivo di questo elaborato è di approfondire, attraverso una revisione sistematica della letteratura, l'utilità dell'applicazione dell'IA come supporto al processo di triage infermieristico.

Nella presente revisione sistematica, sono stati analizzati e raccolti dati presenti nella letteratura da studi che esplorano le proposte di applicazione dell'intelligenza artificiale nel contesto di triage infermieristico.

In particolare, sono stati considerati gli studi che descrivono tecnologie basate sull'intelligenza artificiale che possono essere d'esempio per applicazioni utilizzabili dagli infermieri in un futuro prossimo e che potranno essere implementate nella pratica per la classificazione della gravità dei pazienti e per il supporto al giudizio clinico.

L'obiettivo principale è quello di sintetizzare le informazioni presenti e di valutare l'efficacia di tali applicazioni nel miglioramento dell'accuratezza e dell'efficienza del processo di triage, in modo tale da mettere in luce eventuali lacune presenti nella letteratura che possano essere un punto di partenza per potenziali ricerche future.

METODI

Il quesito che ha stimolato la ricerca è:

In letteratura, sono presenti delle evidenze circa la validità e l'efficacia dell'utilizzo dell'intelligenza artificiale durante il processo di triage infermieristico, rispetto ai metodi tradizionali e l'adeguato utilizzo di algoritmi validati, considerando accuratezza nella valutazione, tempi e casi di sovrastima e sottostima?

Per formulare il quesito di ricerca è stato utilizzato il metodo PICO così strutturato:

P = pazienti che necessitano di triage in contesti d'emergenza;

I = utilizzo dell'intelligenza artificiale nel processo di triage;

C = metodo tradizionale senza l'utilizzo dell'intelligenza artificiale;

O = accuratezza della valutazione del paziente, velocità del processo di triage, casi di sovrastima e sottostima.

È stata eseguita una revisione sistematica tenendo conto delle linee guida PRISMA (*Preferred Reporting Items for Systematic Reviews and Meta-Analyses*).

La ricerca è stata condotta tra giugno e agosto 2024 attraverso l'uso delle banche dati di ricerca come PubMed e l'uso del MeSH (Medical Subject Headings), CINAHL tramite EBSCOhost e infine Scopus.

Le keywords utilizzate per la ricerca sono: "nurs*" (è stata scelta la desinenza per poter includere nursing e nurse), "triage", "artificial intelligence", in combinazione con l'uso degli operatori booleani AND e OR.

Per poter includere più articoli possibili sono state utilizzate diverse combinazioni nella banca dati CINAHL e PubMed, schematizzate nella tabella seguente.

TAB. 1
SCHEMA DELLE COMBINAZIONI DELLE KEYWORDS

CINAHL	PUBMED	SCOPUS
nurs* AND triage AND artificial intelligence OR a.i. OR ai	triage AND (nurse OR nursing) AND artificial intelligence	nurs* AND triage AND artificial intelligence
triage AND artificial intelligence OR a.i. OR a.i.	triage AND artificial intelligence	-----

CRITERI DI INCLUSIONE ED ESCLUSIONE

Sono stati selezionati gli studi pubblicati in un range temporale che va dal 2019 al 2024 (ultimi cinque anni), in lingua italiana e inglese, provvisti di accesso all'abstract e al full text gratuito.

Sono stati esclusi studi pubblicati antecedenti al 2019, in lingue diverse dall'inglese e dall'italiano, che non presentavano disponibilità di abstract e di full text gratuito.

Durante la ricerca nella banca dati Scopus sono stati esclusi gli articoli che non rientravano nella subject area "medicine" e "nursing".

Nella tabella seguente sono riportati in maniera riassuntiva i criteri di eleggibilità della ricerca:

TAB. 2
TABELLA RIASSUNTIVA DEI CRITERI DI ELEGGIBILITÀ

	INCLUSIONE
POPOLAZIONE	Pazienti in contesti ospedalieri (o di triage a distanza)
INTERVENTO	Utilizzo dell'IA applicata al processo di triage
CONFRONTO	Metodi tradizionali di triage, utilizzo di scale (KTAS, CTAS, TTAS, ESI, ATS).
OUTCOME	Accuratezza e capacità della classificazione sulla base della priorità, il tasso di sovrastima e sottostima.
TIPOLOGIA DI STUDI	Studi primari riguardanti proposte basate sull'IA potenzialmente utilizzabili dagli infermieri durante il processo di triage intraospedaliero
DATA DI PUBBLICAZIONE	2019-2024
LINGUA	Inglese, italiano
ACCESSO	Accesso al free full text, presenza di abstract

SELEZIONE DEGLI ARTICOLI

Il processo di selezione degli articoli è avvenuto in più fasi, descritte di seguito per garantire la trasparenza del processo.

Una prima selezione è avvenuta per mezzo di applicazione dei criteri di inclusione ed esclusione e procedendo con l'eliminazione dei duplicati.

Durante la fase successiva si è proseguito leggendo il titolo e l'abstract degli articoli per accertare l'appropriatezza dei contenuti.

Gli articoli non pertinenti e che non hanno soddisfatto i criteri di eleggibilità sono stati esclusi.

Un'ulteriore revisione è stata svolta leggendo il testo integrale degli articoli che hanno superato le fasi precedenti.

Durante l'ultima fase sono stati riesaminati gli articoli inclusi per appurarne la pertinenza.

Il processo è stato documentato attraverso l'utilizzo del diagramma di flusso PRISMA e gli articoli selezionati sono stati inclusi in una tabella riassuntiva dei dati che comprende la tipologia di studio analizzata, l'obiettivo di ogni studio, gli interventi attuati e gli esiti misurati.

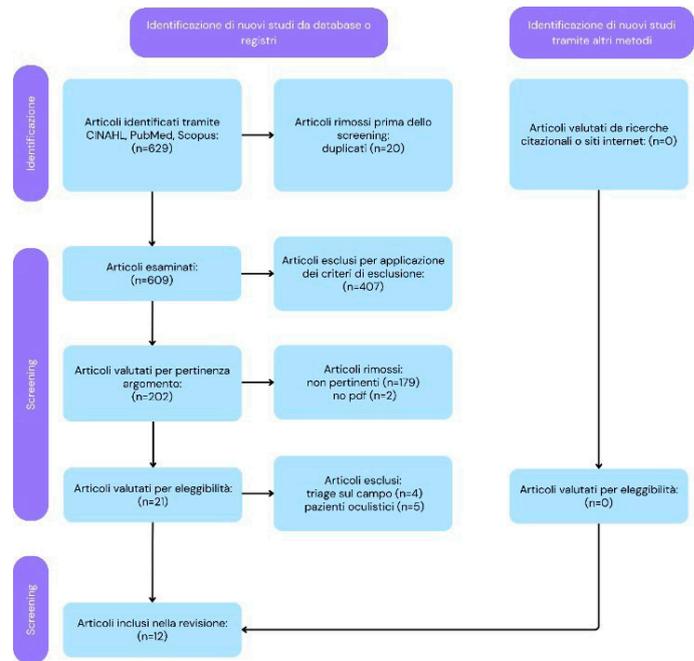


Figura 1 Diagramma di flusso PRISMA.

Di seguito, la tabella riassuntiva degli studi inclusi nella revisione:

TABELLA RIASSUNTIVA DEGLI STUDI					
N°	Titolo	Tipologia	Obiettivo	Interventi	Outcomes
1	“ChatGPT: Friend or Foe? - Utility in Trauma Triage” Jerry Jacob, 2023,India.	Studio prospettico	L'obiettivo di questo studio è quello di valutare se la piattaforma ChatGPT è in grado di ottimizzare il triage per i pazienti politraumatizzati.	Gli scenari sono stati inseriti testualmente e a ChatGPT è stato chiesto di svolgere la valutazione di base attraverso le scale Australasian Triage System (ATS) ed Emergency Severity Index (ESI).	L'abilità di predizione della piattaforma dei punteggi delle scale ATS ed ESI per ogni scenario proposto.
2	“Evaluating the Reliability of a Remote Acuity Prediction Tool in a Canadian Academic Emergency Department” Laila Nasser et al., 2024, Canada.	Studio retrospettivo	L'obiettivo di questo studio è quello di esaminare l'affidabilità di un algoritmo di machine learning virtuale (VTriage) nel prevedere a distanza i punteggi di gravità per i pazienti che sono stati visti in un pronto soccorso di un centro accademico di cure terziarie da gennaio 2021 ad agosto 2022, applicando l'algoritmo a dati retrospettivi.	Sono stati raccolti dati quali la data, l'ora della visita in pronto soccorso ed il motivo, l'età del paziente, il sesso, il sintomo presentato e il punteggio di dolore riportato. L'algoritmo è stato progettato per sovrastimare il livello di gravità dei sintomi per assicurare un alto livello di sicurezza nell'uso a distanza. I punteggi sono stati poi confrontati con i punteggi di triage assegnati dagli infermieri in tempo reale che utilizzavano il CTAS elettronico.	L'affidabilità del giudizio, i casi di sovrastima e di sottostima dei pazienti arrivati in pronto soccorso con ambulanza e con mezzi propri.

TABELLA RIASSUNTIVA DEGLI STUDI

3	<p>“The diagnostic and triage accuracy of the GPT-3 artificial intelligence model: an observational study” David M Levine et al., 2024, USA.</p>	Studio osservazionale	<p>L’obiettivo di questo studio è quello di testare l’accuratezza predittiva della capacità di diagnosi e di triage di GPT-3, un modello di machine learning auto-supervisionato che viene addestrato solamente a prevedere la parola successiva, basato sul completamento automatico che usufruisce di una larga collezione di testi reperiti su internet.</p>	<p>Le prestazioni di GPT-3 sono state comparate con le prestazioni di un campione di 5000 utenti internet non sanitari negli Stati Uniti e di un campione di 21 medici praticanti della Harvard Medical School. Sono state create 48 vignette costituite da un linguaggio semplice, che contenevano un’informazione seguita da dettagli riguardanti specifiche condizioni cliniche. Ogni vignetta era associata ad un livello di triage e ad una diagnosi precedentemente assegnate da due medici internisti della Harvard Medical School.</p>	<p>L’accuratezza della diagnosi e la corretta classificazione di triage da parte degli utenti internet, dei medici e di GPT-3. L’intervallo di confidenza per la diagnosi e per il triage. L’accuratezza di GPT-3 in relazione con l’aumento della gravità dei casi proposti. I casi di errore, riguardanti il triage.</p>
4	<p>“Patient stratification based on the risk of severe illness in emergency departments through collaborative machine learning models” Jui-Ying Chen et al., 2024, Taiwan.</p>	Studio retrospettivo	<p>L’obiettivo di questo studio è di definire un modello di machine learning collaborativo basato su quattordici variabili che possa originare punteggi di rischio per patologie gravi, a partire da dati raccolti in un ospedale universitario tra il 2015 e il 2022 che potrebbe aiutare il personale di triage a fornire una migliore classificazione dei pazienti per cure critiche tempestive.</p>	<p>Sono stati sviluppati modelli di rete neurale artificiale (ANN): uno che utilizzava la TTAS (ANN-MH) come valore predittore mentre un altro che non ne usufruiva (ANN-MO). La capacità predittiva di entrambi i modelli è stata valutata utilizzando due aree sotto la curva (AUROC e AUPRC) e sono stati confrontati sensibilità, specificità, valore predittivo positivo e negativo (PPV e NPV).</p>	<p>La capacità predittiva di entrambi i modelli, l’AUROC e l’AUPCR di entrambi i modelli, sensibilità, precisione, specificità, PPV e NPV.</p>
5	<p>“Assessing the utility of artificial intelligence throughout the triage outpatients: a prospective randomized controlled clinical study” Xiaoni Liu et al., 2024, Cina.</p>	Studio prospettico randomizzato e controllato	<p>L’obiettivo di questo studio è quello di valutare l’efficacia di ChatGPT 3.5 nel triage.</p>	<p>Sono stati utilizzati i dati di uno studio retrospettivo di coorte svolto a marzo 2023 in cui sono state selezionate in maniera casuale 30 cartelle cliniche estratte da un vasto pool e a ChatGPT è stato richiesto di rispondere a 30 domande attinenti ad esse, le cui risposte sono state valutate da cinque esperti e i dati di un altro studio di coorte prospettico in cui è stato somministrato un modulo da compilare a 45 pazienti. Il triage è stato eseguito in maniera tradizionale da infermieri addestrati e con l’utilizzo di ChatGPT. La coerenza tra i due tipi di triage è stata valutata da 5 esperti.</p>	<p>Comprensione, competenza sui concetti di primo soccorso e sulle gerarchie di trattamento, doti umane e coerenza tra le due tipologie di triage.</p>

TABELLA RIASSUNTIVA DEGLI STUDI

6	<p>“Automatic Classification of the Korean Triage Acuity Scale in Simulated Emergency Room Using Speech Recognition and Natural Language Processing: A Proof-of-Concept Study” Dongkyun Kim et al., 2021, Korea.</p>	Studio prospettico	<p>L’obiettivo di questo studio è quello di sviluppare un sistema di triage automatizzato, basato sul riconoscimento vocale e l’IA, che sia in grado di classificare i pazienti secondo lo score KTAS attraverso l’estrazione delle informazioni da conversazioni tra i pazienti e gli operatori del pronto soccorso dell’Ospedale Universitario di Hanyang.</p>	<p>Sono stati scelti 762 accessi al pronto soccorso avvenuti tra settembre e dicembre 2019. Per ogni caso è stato simulato un triage ipotetico in cui è stata eseguita una conversazione tra il paziente e l’operatore sanitario in cui venivano esternati dei dati utili per la classificazione KTAS (sintomi, parametri vitali, punteggio di dolore, ecc.). In ogni simulazione sono stati espressi 6 sintomi principali (dolore toracico, dispnea, febbre, ictus, dolore addominale, mal di testa) e le conversazioni sono state registrate e successivamente trascritte.</p>	<p>La percentuale di accuratezza della trascrizione della conversazione, l’accuratezza sulla classificazione KTAS e l’importanza dei sintomi espressi per la determinazione e della classe KTAS da parte dell’algoritmo</p>
7	<p>“Improving ED emergency severity index acuity assignment using machine learning and clinical natural language processing” Oleksandr Ivanov, et al., 2021</p>	Studio retrospettivo	<p>L’obiettivo di questo studio è quello di determinare se i dati delle cartelle cliniche elettroniche possano essere estrapolati e utilizzati con KATE, algoritmo basato sul machine learning, per produrre dei modelli predittivi altamente accurati che utilizzano l’ESI.</p>	<p>I dati utilizzati per sviluppare il modello KATE sono stati estratti da 166.175 accessi al pronto soccorso di due ospedali diversi. Il modello è stato testato su un campione casuale del dataset ai quali era già stata associata una priorità dai clinici utilizzando la scala ESI.</p>	<p>L’accuratezza delle attribuzioni di priorità secondo la scala ESI rispetto all’assegnazione data dagli infermieri e dai clinici ai casi estrapolati.</p>
8	<p>“Leveraging graph neural networks for supporting automatic triage of patients” Annamaria Defilippo et al., 2024, Italy.</p>	Studio osservazionale	<p>L’obiettivo di questo studio è di presentare una proposta di algoritmo basato sull’IA applicata alla scienza delle reti che sia in grado di utilizzare dati storici e informazioni attuali dei pazienti, in modo da assegnare correttamente la priorità durante il triage.</p>	<p>Il modello è stato testato utilizzando un dataset contenente 6962 casi clinici e 16 caratteristiche per ciascuno, disponibile sulla piattaforma Kaggle. Tra le caratteristiche presenti vi erano l’età, i parametri vitali, l’abitudine al fumo, predisposizione genetica al diabete, ecc. In ogni esempio, era incluso anche il codice priorità assegnato in base ad un ordine decrescente di rischio, da rosso a verde.</p>	<p>La capacità dell’algoritmo di classificare correttamente i pazienti secondo la priorità e l’efficienza confrontata con i metodi tradizionali.</p>
9	<p>“The Future of Triage: The Analysis of Traditional Methods Compared to ChatGPT” Helena Mayerhoffer, 2024, Croatia.</p>	Studio osservazionale	<p>L’obiettivo di questo studio è quello di comparare l’accuratezza nella classificazione di triage e nell’assegnazione di diagnosi tra ChatGPT e i metodi di triage tradizionali.</p>	<p>Sono state selezionate le diagnosi più comuni per cui i pazienti si recano in pronto soccorso e le categorie di triage dall’“Emergency Nursing: 5-Tier Triage Protocols”. Sono stati successivamente creati dei casi di studio e sono stati somministrati a ChatGPT e confrontati poi con la letteratura.</p>	<p>La differenza tra le risposte di ChatGPT e quelle rilevate nel libro, la diagnosi individuata, la categoria di triage assegnata e l’accuratezza delle risposte.</p>

TABELLA RIASSUNTIVA DEGLI STUDI

10	“A comparative vignette study: Evaluating the potential role of a generative AI model in enhancing clinical decision-making in nursing” Mor Saban et al., 2024.	Studio trasversale	L'obiettivo di questo studio è quello di valutare se ChatGPT possa dimostrare un giudizio decisionale clinico equivalente a quello di infermieri esperti e studenti infermieri principianti.	Sono stati reclutati 30 infermieri esperti del pronto soccorso e 38 studenti di infermieristica. Sia gli infermieri esperti che i principianti hanno compilato dei questionari online, mentre le risposte di ChatGPT sono state ottenute attraverso la piattaforma stessa.	L'accuratezza delle valutazioni, l'idoneità degli esami raccomandati, l'uso corretto delle risorse e la capacità di rivalutazione del paziente.
----	---	--------------------	--	--	---

TABELLA RIASSUNTIVA DEGLI STUDI CON SETTING “A DISTANZA”

a	“An Artificial Intelligence Based Application for Triage Nurses in Emergency Department, Using the Emergency Severity Index Protocol” George Kipourgos et al., 2022, Greece	Studio retrospettivo prospettico	L'obiettivo di questo studio è quello di progettare un'applicazione chiamata iTRIAGE per poter supportare gli infermieri nel processo di triage, aiutandoli a prendere decisioni accurate e veloci. L'applicazione dovrebbe anche essere in grado di indicare il medico più adatto per affrontare i problemi riscontrati dal paziente. L'applicazione si propone anche come strumento di training per studenti di medicina e infermieristica.	Le informazioni utilizzate per sviluppare e testare i modelli sono state raccolte durante il triage di 616 pazienti presso l'Ospedale Universitario di Patraso. Sono stati creati due modelli di IA: il primo utilizza il machine learning per categorizzare i pazienti e identificare il medico più adatto per quel trattamento, attraverso la definizione di sottosistemi (patologico, cardiologico, muscolo scheletrico, ecc.). Il secondo modello, basato sull'algoritmo fuzzy clips, mira a classificare i pazienti attraverso l'utilizzo della scala ESI	Percentuale di successo nella categorizzazione dei pazienti nei vari sottosistemi attraverso il modello machine learning ed il tasso di corretta classificazione, con il modello Fuzzy Clips, del punteggio ESI.
b	“Use Characteristic and Triage Acuity of a Digital Symptom Checker in a Large Integrated Health System: Population-Based Descriptive Study” Keith E Morse et al., 2020, USA.	Studio descrittivo	L'obiettivo di questo studio è quello di valutare i dati demografici provenienti dagli utenti e i livelli di sensibilità di triage forniti da un chatbot di controllo dei sintomi distribuito in cooperazione con un sistema sanitario integrato negli Stati Uniti.	Sono stati inclusi nello studio tutte le valutazioni dei sintomi effettuate nel sistema Sutter Health tra aprile 2019 e febbraio 2020. Per quanto riguarda il triage, il chatbot era in grado di riportare otto tipi di consigli diversi sulla base del sintomo riferito, divisi in tre macrocategorie di rischio: alto, chiama l'ambulanza, medio, consulta un medico entro la giornata, basso rivolgiti ad un medico nelle prossime settimane o gestisci la situazione a casa.	I dati demografici sugli utenti, l'orario relativo alla richiesta di valutazione dei sintomi, il tipo di sintomo/i riferito/i, grado di urgenza e livello di accuratezza del triage.

RISULTATI

Jerry Jacob, ha valutato la capacità predittiva di ChatGPT. Inizialmente ChatGPT sembrava aver commesso un errore declassando un paziente, utilizzando i criteri ATS ed ESI.

Quando gli è stato richiesto di prendere in esame la situazione del paziente, ChatGPT ha riconosciuto l'errore e ha generato una nuova risposta in cui era contenuta una nuova classificazione.

L'algoritmo è stato in grado di fornire indicazioni tempestive sulla classificazione per entrambe le scale di valutazione. Nonostante possa sembrare uno strumento valido per migliorare il triage, non è esente da complicazioni. D'altro canto, potrebbe risultare un ottimo strumento nei contesti rurali grazie alla capacità di prevedere di quali interventi futuri avrebbero bisogno i pazienti, in modo da facilitare la pianificazione dell'assistenza per poter offrire la qualità migliore di quest'ultima (Jacob, 2023).

Nel secondo studio, viene esaminata l'affidabilità dell'algoritmo VTriage. Sono stati ottenuti risultati esatti per il 48,4% dei casi. L'affidabilità del giudizio è cambiata da scarsa a discreta utilizzando le statistiche kappa, sia nei casi di pazienti giunti in pronto soccorso in ambulanza e sia nei casi di coloro arrivati con l'utilizzo di mezzi propri.

Utilizzando il punteggio CTAS elettronico, VTriage ha eseguito: una valutazione sovrastimata nel 44,8% dei casi e sottostimata nel 5,6% dei casi nei pazienti arrivati con ambulanza; una valutazione sovrastimata nel 46,5% dei casi e sottostimata nel 5,4% dei casi nei pazienti giunti in pronto soccorso con mezzi propri. VTriage, necessita di ulteriori miglioramenti prima di poter essere implementato in sicurezza per l'uso con i pazienti in tempo reale (Nasser et al., 2024).

In questo studio, GPT-3 ha fornito la diagnosi corretta, riportata tra le prime opzioni, nell'88% dei casi, rispetto al 54% delle persone comuni e al 96% dei medici. GPT-3 ha effettuato un triage corretto nel 70% dei casi, rispetto al 74% delle persone comuni e al 91% dei medici.

Attraverso le valutazioni statistiche è risultato che l'intervallo di confidenza era ben calibrato sia per quanto riguarda la predizione della diagnosi e sia per l'assegnazione della categoria di triage.

È stato osservato che l'accuratezza di GPT-3 tendeva a diminuire dell'8,33% per ogni livello di aumento della gravità del caso. Nell'analisi degli errori di triage è spiccato che GPT-3 ha declassato i casi urgenti per un totale di 7 volte.

In generale, è emerso che un modello basato sull'IA e programmato senza alcun addestramento specifico, potrebbe svolgere diagnosi a livelli vicini, ma inferiori, a quelli di medici e superiori a quelli di persone comuni e non sanitarie.

Per quanto riguarda il triage, le prestazioni di GPT-3 sono risultate inferiori rispetto a quelle dei medici ma vicine a quelle delle persone comuni non sanitarie che hanno accesso alle informazioni presenti su internet (Levine et al., 2024).

Nel caso dell'articolo di Chen et al., analizzando i risultati dei due modelli, il modello ANN-MH, ovvero quello che utilizzava la scala TTAS come valore predittivo, ha superato il modello ANN-MO ed entrambi hanno superato la classificazione del triage tradizionale.

Lo studio di coorte retrospettivo condotto da Liu et al., ha rivelato che tra le 30 risposte valutate dagli esperti, 17 hanno ricevuto un punteggio molto alto (tra 9.5 e 10 punti) e hanno dimostrato comprensione, analisi professionale, concetti di pronto soccorso adeguati, gerarchie di diagnosi e di trattamento coerenti e buone doti umane, 7 hanno ricevuto un punteggio alto (9 punti) ed erano, perlopiù, professionali e comprensive ma da migliorare nel complesso, e 6 hanno ricevuto un punteggio più basso (8 e 7 punti) risultando non professionali e incomplete.

Lo studio prospettico di coorte ha rivelato che il triage tradizionale svolto sui 45 pazienti studiati risultava particolarmente coerente con il triage svolto da ChatGPT.

Tra gli esperti che hanno valutato la coerenza, 3 di loro hanno considerato un tasso del 100%, 1 del 95,6% e 1 del 93,3%.

Nel complesso, i risultati hanno indicato che le risposte restituite da ChatGPT hanno fornito delle informazioni di triage accurate e professionali ai 45 pazienti analizzati.

Nel sesto studio, la percentuale di errore relativa

al sistema di triage automatizzato è diminuita del 25,21% utilizzando una particolare tecnica di machine learning (il transfer learning).

Per quanto riguarda la classificazione si sono raggiunte delle prestazioni eccellenti.

Sono stati individuati dei vocaboli più importanti di altri che hanno influito sulla determinazione del livello KTAS, tra questi troviamo: “stress”, “punteggio di dolore”, “febbre”, “respiro”, “testa” e “torace” (D. Kim et al., 2021).

Nello studio di Ivanov et al., il modello di triage KATE ha previsto correttamente le assegnazioni di priorità attraverso l'utilizzo dell'ESI nel 75,7% dei casi rispetto agli infermieri coinvolti (59,8%) e alla media dei clinici che hanno partecipato allo studio (75,3%). L'accuratezza di KATE è stata del 26,9% superiore rispetto alla media degli infermieri.

Riguardo le assegnazioni di urgenza ESI di tipo 2 e 3, l'accuratezza di KATE è stata del 93,2% superiore mostrando un'accuratezza dell'80% rispetto al 41,4% degli infermieri.

I risultati della ricerca dell'ottavo studio evidenziano i vantaggi dell'integrazione dell'IA nel settore sanitario, in particolare nel contesto di triage dei pazienti. Attraverso i test eseguiti, il modello di IA è stato in grado di classificare con precisione i pazienti nei livelli di triage, mostrando prestazioni superiori rispetto ai metodi tradizionali. Inoltre, il fatto che il modello analizzi la storia clinica dettagliata del paziente in unione con i parametri vitali del momento della visita in triage, consente una valutazione molto più dettagliata e consente un'ottimizzazione nel processo di prioritizzazione dei pazienti (Defilippo et al., 2024).

Dai risultati provenienti dalla ricerca di Mayerhoffer, ChatGPT ha assegnato la categoria opportuna di triage, secondo le raccomandazioni della letteratura, nel 43,33% dei casi.

Sono emersi degli errori di categoria in alcuni casi ma la tendenza era quella di sovrastimare il problema, riscontro utile poiché è indicativo di buoni livelli di sicurezza per il paziente.

Per quanto riguarda la capacità diagnostica, ha mostrato buone prestazioni pur non avendo una formazione medica specifica e non possedendo un'ottimizzazione per tali scopi.

La sua precisione non è ancora sufficientemente alta,

però, per poter fare affidamento esclusivamente sulle risposte del chatbot ma, considerando la prospettiva del miglioramento e dell'ottimizzazione specifica, i risultati evidenziano il potenziale per l'utilizzo in triage.

Alla luce dello studio di Saban et al., ChatGPT ha mostrato dell'indecisione nelle valutazioni iniziali, ha avuto la tendenza a suggerire dei test diagnostici non necessari e durante il processo di rivalutazione ha dimostrato una comprensione inaccurata e delle modifiche non appropriate. In termini di prestazioni, le risposte mantenevano una struttura logica e dimostravano chiarezza.

Complessivamente, ha mostrato una tendenza alla sovrastima dei casi di triage.

Per quanto riguarda lo studio di Kipourgos et al., su iTRIAGE la percentuale di successo nella categorizzazione dei pazienti attraverso il machine learning è stata tra il 72 e il 95% con un solo caso di categorizzazione errata su 616.

Per quanto concerne il secondo metodo basato sul modello Fuzzy Clips, i parametri hanno ottenuto dei livelli soddisfacenti: accuratezza/sensibilità/specificità = 99% e precisione = 93%.

Lo studio ha dimostrato che, in futuro, iTRIAGE potrebbe essere uno strumento d'aiuto per il processo di triage del pronto soccorso e potrebbe essere utilizzato come strumento d'apprendimento per gli studenti di medicina e di infermieristica.

I risultati dello studio di Morse et al., hanno dimostrato che le raccomandazioni di triage distribuite dal chatbot di controllo sono risultate accettabili e generalmente paragonabili a quelle delle linee di triage telefonico gestite da infermieri con sede negli Stati Uniti.

DISCUSSIONE

Gli articoli che hanno soddisfatto i criteri di eleggibilità sono stati 10 ma ne sono stati inclusi altri 2 che presentano delle proposte basate sull'intelligenza artificiale con un setting "a distanza", potendo includere nella presente revisione anche l'efficacia di nuove tecnologie che in futuro potrebbero essere adottate come supporto all'attività infermieristica di triage e come strumento di training per la formazione di studenti di infermieristica e di medicina.

Tale scelta è data dalla volontà di racchiudere quanti più dati possibili per poter validare la ricerca, data la scarsità di articoli scientifici presenti sull'argomento.

Tutti i dodici articoli presenti nella revisione esprimono dati molto recenti: in particolare, sette sono del 2024 e gli altri sono stati pubblicati nel periodo compreso tra il 2023 e il 2020.

In ogni articolo incluso in questa revisione, i risultati estrapolati sono stati ottenuti dal confronto delle prestazioni delle tecnologie descritte, proposte o sviluppate, con le procedure di triage tradizionali rappresentate dal giudizio degli infermieri (o clinici) o dall'utilizzo di scale validate a livello mondiale.

Di seguito, è riportata una tabella in cui sono schematizzati i metodi utilizzati per il confronto in ogni studio.

Circa la metà degli studi analizzati hanno come focus principale l'utilizzo di ChatGPT come strumento impiegabile nel processo di triage, permettendo così un'analisi più completa del suo potenziale relativo alla valutazione dei pazienti in contesti d'emergenza.

Nel complesso, tutti e cinque gli articoli in cui è stato analizzato il potenziale di ChatGPT nelle pratiche di triage hanno messo in evidenza esiti in linea di massima positivi. Tuttavia, Mayerhoffer riporta che, sebbene i risultati siano abbastanza promettenti, il tasso di precisione di ChatGPT non è ancora così elevato da consentirne la totale fiducia verso le valutazioni. Similmente, Saban et al., sottolineano la presenza di criticità riguardo a dei comportamenti di indecisione dimostrati da ChatGPT durante le valutazioni iniziali, manifestando una tendenza alla sovrastima della gravità dei casi presentati e l'attitudine a suggerire esami diagnostici non necessari.

Negli altri articoli sono state trattate una varietà di approcci innovativi, tra cui algoritmi di machine learning, un modello di machine learning supervisionato addestrato per il completamento automatico, un modello collaborativo di machine learning, un sistema di triage automatizzato che funziona con il riconoscimento vocale e l'IA, un algoritmo di machine learning per la creazione di

TAB. 5 TABELLA RIASSUNTIVA DEL CONFRONTO DELLE PRESTAZIONI	
STUDIO	CONFRONTO
Jacob, 2023	ATS, ESI
Nasser et al., 2024	Triage eseguito da infermieri
Levine et al., 2024	Punteggi assegnati da medici praticanti
Chen et al., 2024	TTAS
Liu et al., 2024	Triage eseguito da infermieri
Kim et al., 2021	KTAS
Ivanov, 2021	ESI
Defilippo et al., 2024	Triage eseguito da infermieri
Mayerhoffer, 2024	"Emergency Nursing: 5-Tier Triage Protocols"
Saban et al., 2024	Triage eseguito da infermieri e studenti di infermieristica
Kipourgos et al., 2022	ESI
Morse et al., 2020	Triage telefonico eseguito da infermieri

modelli predittivi e un algoritmo clinico basato sull'IA applicata alla scienza delle reti.

Per quanto riguarda gli altri due articoli inclusi, sebbene con un setting differente rispetto al target stabilito per la presente revisione sistematica, viene descritto l'uso di algoritmi di machine learning e fuzzy clip per sviluppare un'applicazione di supporto al triage infermieristico e di training per gli studenti, e l'impiego di un chatbot di controllo della sintomatologia descritta dagli utenti.

Nell'insieme, i risultati degli studi che sono stati inclusi risultano promettenti. Nonostante ciò, il limite principale risiede nella limitatezza di letteratura scientifica disponibile sulla materia in esame e nel fatto che si tratta ancora di un campo in via di sviluppo.

Il potenziale che si evince da questi sistemi è, però, evidente, soprattutto se vengono sfruttati come sistemi di supporto per ottimizzare il lavoro infermieristico.

Bibliografia

- Chen, J.-Y., Hsieh, C.-C., Lee, J.-T., Lin, C.-H., & Kao, C.-Y. (2024). Patient stratification based on the risk of severe illness in emergency departments through collaborative machine learning models. *The American Journal of Emergency Medicine*, 82, 142–152. <https://doi.org/10.1016/j.ajem.2024.06.015>
- Defilippo, A., Veltri, P., Liò, P., & Guzzi, P. H. (2024). Leveraging graph neural networks for supporting automatic triage of patients. *Scientific Reports*, 14(1), 12548. <https://doi.org/10.1038/s41598-024-63376-2>
- Ivanov, O., Wolf, L., Brecher, D., Lewis, E., Masek, K., Montgomery, K., Andrieiev, Y., McLaughlin, M., Liu, S., Dunne, R., Klauer, K., & Reilly, C. (2021). Improving ED Emergency Severity Index Acuity Assignment Using Machine Learning and Clinical Natural Language Processing. *Journal of Emergency Nursing*, 47(2), 265–278.e7. <https://doi.org/10.1016/j.jen.2020.11.001>
- Jacob, J. (2023). ChatGPT: Friend or Foe?—Utility in Trauma Triage. *Indian Journal of Critical Care Medicine*, 27(8), 561–564. <https://doi.org/10.5005/jp-journals-10071-24498>
- Kim, D., Oh, J., Im, H., Yoon, M., Park, J., & Lee, J. (2021). Automatic Classification of the Korean Triage Acuity Scale in Simulated Emergency Rooms Using Speech Recognition and Natural Language Processing: A Proof of Concept Study. *Journal of Korean Medical Science*, 36(27), e175. <https://doi.org/10.3346/jkms.2021.36.e175>
- Kipourgos, G., Tzenalis, A., Diamantidou, V., Koutsojannis, C., & Hatzilygeroudis, I. (2022). An Artificial Intelligence Based Application for Triage Nurses in Emergency Department, Using the Emergency Severity Index Protocol. *International Journal of Caring Sciences*, 15(3), 1764–1772
- Levine, D. M., Tuwani, R., Kompa, B., Varma, A., Finlayson, S. G., Mehrotra, A., & Beam, A. (2024). The diagnostic and triage accuracy of the GPT-3 artificial intelligence model: An observational study. *The Lancet Digital Health*, 6(8), e555–e561. [https://doi.org/10.1016/S2589-7500\(24\)00097-9](https://doi.org/10.1016/S2589-7500(24)00097-9)
- Liu, X., Lai, R., Wu, C., Yan, C., Gan, Z., Yang, Y., Zeng, X., Liu, J., Liao, L., Lin, Y., Jing, H., & Zhang, W. (2024). Assessing the utility of artificial intelligence throughout the triage outpatients: A prospective randomized controlled clinical study. *Frontiers in Public Health*, 12. <https://doi.org/10.3389/fpubh.2024.1391906>
- Mayerhoffer, H. (2024). The Future of Triage: The Analysis of Traditional Methods Compared to ChatGPT. *Croatian Nursing Journal*, 8(1), 29–36. <https://doi.org/10.24141/2/8/1/3>
- Morse, K. E., Ostberg, N. P., Jones, V. G., & Chan, A. S. - (2020). Use Characteristics and Triage Acuity of a Digital Symptom Checker in a Large Integrated Health System: Population-Based Descriptive Study. *Journal of Medical Internet Research*, 22(11), <https://doi.org/10.2196/20549>
- Page MJ, et al. *BMJ* 2021;372:n71. doi: 10.1136/bmj.n71
- Saban, M., & Dubovi, I. (2024). A comparative vignette

Publisher's note: all claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article or claim that may be made by its manufacturer is not guaranteed or endorsed by the publisher.
©Copyright: the Author(s), 2025