Valutare in Sanità
Una voce al servizio della Salute

RIVISTA
SCIENTIFICA

NUMERO SPECIALE



PMA
Special Edition

Numero speciale

VALUTARE IN SANITÀ
Rivista Scientifica
Registrazione presso il Tribunale di Bari
n.36 del 24/09/09

www.valutareinsanita.it

Direttore Responsabile
Maddalena Milone

Direzione, redazione e pubblicità
MP s.r.l.
Via Vincenzo Sassanelli, 56
70125 Bari - Italia
Tel.080 9905360
Fax 080 9905359
e-mail: info@meeting-planner.it
website: www.meeting-planner.it

Editore
MP s.r.l.
Via Vincenzo Sassanelli, 56 - 70125 Bari

Stampa
Mediatipo s.r.l.
Via delle Margherite, 26
70026 - Modugno (BA)

La Direzione non si assume la responsabilità dei contenuti delle inserzioni pubblicitarie.
È vietata la riproduzione, anche parziale, dei testi.
Regolamento UE 679/2016 Tutela della Privacy – Informativa: Responsabile del trattamento MP S.r.l.
Tutte le richieste di consultazione, modifica, integrazione e cancellazione dei dati personali dovranno essere inviate a MP S.r.l.

Numero chiuso in redazione a luglio 2026

Numero realizzato con il supporto non condizionante di Merck Serono.

Sommario



04 **IL RUOLO DELL'ORMONE ANTIMÜLLERIANO (AMH) COME MARKER PREDITTIVO DELLA RISERVA OVARICA**

Dott. Marco Galletta – Responsabile Centro PMA, Ospedale Papardo, Messina

07 **GLI EFFETTI DEL LASER VAGINALE CO2 SULL'OUTCOME GRAVIDICO NELL'INFERTILITÀ**

Dott. Giovanni Rosati – Specialista in Ginecologia, Ostetricia e PMA, Centro Embryos, Battipaglia (SA)

09 **IL POTENZIALE, I PERICOLI E LE INSIDIE DELL'INTELLIGENZA ARTIFICIALE (IA) NELLE TECNICHE DI PROCREAZIONE MEDICALMENTE ASSISTITA (PMA)**

Dott.ssa Elisabetta Trabucco – Direttrice del Centro di Fecondazione Assistita Genera, Napoli

IL RUOLO DELL'ORMONE ANTIMÜLLERIANO (AMH) COME MARKER PREDITTIVO DELLA RISERVA OVARICA

Dott. Marco Galletta – Responsabile Centro PMA, Ospedale Papardo, Messina

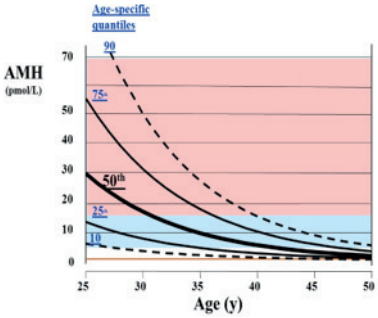
Introduzione

La valutazione della riserva ovarica rappresenta uno degli aspetti fondamentali della medicina della riproduzione moderna. Con tale termine si indica il patrimonio follicolare residuo presente nelle ovaie e la capacità riproduttiva potenziale della donna sia in termini qualitativi che quantitativi. Usata prevalentemente nei percorsi di PMA per prevedere la risposta alla stimolazione, ma anche in altri contesti come la preservazione della fertilità o per la diagnosi di insufficienza ovarica precoce. Ma quali sono i principali markers di riserva ovarica? L'FSH basale, estradiolo basale, conta dei follicoli antrali (AFC) e l'AMH stesso. Negli ultimi anni l'*Anti-Müllerian Hormone* è diventato uno dei biomarcatori più utilizzati per stimare tale riserva, grazie alla sua affidabilità e relativa stabilità durante il ciclo mestruale.

Fisiologia dell'AMH

Si tratta di una glicoproteina dimerica appartenente alla famiglia del *transforming growth factor-beta* (TGF-β) espresso nelle c. del Sertoli dei testicoli fetali (dove induce regressione dei dotti mulleriani). Nella donna è prodotto dalle cellule della granulosa dei follicoli pre-antrali e antrali fino a quando non diventano sensibili all'FSH.

I livelli di AMH sono quasi irrilevabili alla nascita, con un sottile aumento entro i primi 2-4 anni di vita, dopo sono stabili fino all'adolescenza. Da lì diminuiscono longitudinalmente nel tempo come segnale di esaurimento della riserva follicolare e diventano irrilevabili in menopausa.



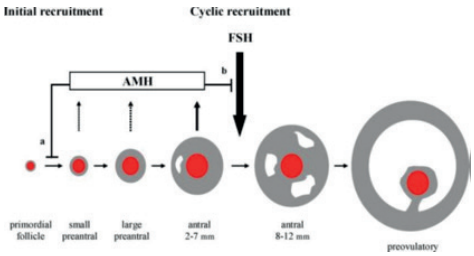
** Guibourdenche J, Lucidarme N, Chevenne D, Rigal O, Nicolas M, Luton D, Le´ger J, Porquet D, Noe¨l M. Anti-Mullerian hormone levels in serum from human foetuses and children: pattern and clinical interest. Mol Cell Endocrinol 2003;211:55–63.*

In ambito clinico, il dosaggio dell'AMH viene oggi impiegato sia nella valutazione della fertilità naturale sia nella programmazione delle tecniche di procreazione medicalmente assistita (PMA) e per la preservazione della fertilità.

L'AMH svolge un ruolo importante nel controllo della follicologenesi, in quanto:

- Limita il reclutamento iniziale dei follicoli primordiali;
- Riduce la sensibilità follicolare all'ormone follicolo-stimolante (FSH);
- Contribuisce alla regolazione della crescita follicolare.

Inibisce il reclutamento follicolare iniziale (a) e inibisce la crescita FSH-dipendente e la selezione dei follicoli antrali (b) e impedendo in questo modo l'esaurimento del pool di tutti i follicoli primordiali contemporaneamente.



Ruolo dell'AMH nello sviluppo del follicolo ovarico (centro rosso rappresenta l'ovocita, area grigia lo strato di cellule della granulosa).

I livelli sierici di AMH riflettono quindi indirettamente il numero dei follicoli ovarici residui e risultano correlati alla quantità di ovociti disponibili.

I valori principali sono qui riportati:

Alto	**> 3.0 ng/ml**
Normale	**> 1.2 ng/ml**
Basso	**0,3- 1.0 ng/ml**
Molto basso	**< 0,3 ng/ml**

AMH e riserva ovarica

L'AMH è considerato uno dei migliori marker della riserva ovarica attualmente disponibili. A differenza di altri parametri endocrini, come FSH ed estradiolo, i livelli di AMH mostrano variazioni relativamente limitate durante il ciclo mestruale, rendendo possibile il dosaggio in qualsiasi momento.

Valori elevati di AMH sono generalmente associati a una maggiore riserva ovarica, mentre valori bassi indicano una riduzione del patrimonio follicolare. Tale caratteristica rende l'AMH particolarmente utile nella valutazione della fertilità femminile, soprattutto nelle donne che intendono posticipare la gravidanza o che devono affrontare trattamenti potenzialmente gonadotossici, come chemioterapia o radioterapia per la eventuale preservazione della fertilità.

L'AMH diminuisce progressivamente con l'età, fino a diventare quasi indosabile in menopausa. Questa correlazione con l'invecchiamento ovarico lo ha reso uno strumento importante anche nella previsione della menopausa precoce (insufficienza ovarica, POF). Nei casi di Obesità *Hi et al*. ha dimostrato che i livelli di AMH nel siero sono più bassi nelle donne obese rispetto

alle donne di pari età e di peso normale, nonostante la stessa conta dei follicoli antrali. Questo suggerisce che i livelli di AMH nelle donne obese possono essere inferiori per motivi fisiologici legati all'obesità stessa piuttosto che dalla riserva ovarica ridotta. Riduzione dei valori anche in conseguenza di chirurgia, come nella Ovariectomia mono o bilaterale dove l'AMH diventa irrilevabile già dopo 3-5 gg post intervento.
L'uso della contraccezione orale rende poco affidabile l'AMH come predittore di riserva ovarica in quanto i sui valori si riducono del 30% nelle utilizzatrici di CO per raggiungere il 50% dopo 9 mesi, ma si normalizzano dopo pochi cicli mestruali naturali.

APPLICAZIONI CLINICHE

Procreazione medicalmente assistita

Attualmente è il migliore test in termini di sensibilità e specificità rispetto a AFC e alle concentrazioni FSH, E2 e inibina B10. Consente, infatti, una più precisa previsione della risposta ovarica alla stimolazione, permette l'individualizzazione del trattamento per migliorare l'*outcome* riproduttivo ed inoltre aiuta nell'identificazione di pazienti che hanno possibilità di gravidanza dopo IVF vicina a zero. Nelle tecniche di PMA, il dosaggio dell'AMH è ampiamente utilizzato per prevedere la risposta ovarica alla stimolazione con gonadotropine. Infatti Donne con livelli elevati di AMH tendono a rispondere meglio alla stimolazione ovarica, producendo un numero maggiore di ovociti e quindi personalizzare i protocolli terapeutici e ridurre il rischio di sindrome da iperstimolazione ovarica (OHSS). Al contrario, livelli molto bassi di AMH possono indicare una scarsa risposta ovarica e un ridotto numero di ovociti recuperabili identificando le pazienti a bassa prognosi riproduttiva Proprio le pazienti a bassa risposta ovarica cosiddette *poor responders* a rischio di cancellazione ciclo 2-30% di pazienti sottoposte a IVF risultano *POOR*.
Permette pertanto l'individualizzazione del trattamento per migliorare *l'outcome* riproduttivo.
Studi presenti in letteratura hanno ritrovato correlazione tra AMH e ovociti recuperati, concludendo definitivamente che AMH è il migliore marker per predire risposta ovarica alla stimolazione, superando AFC e età.

Definizione di *poor*:

Almeno due tra le seguenti tre caratteristiche devono essere presenti.

- età materna avanzata (≥40 anni) o qualsiasi altro fattore di rischio per POR;
- pregressa POR (≤3 ovociti con un protocollo di stimolazione convenzionale);
- test di riserva ovarica (ORT) anormale (i.e. AFC <5-7 follicoli o AMH <0.5-1.1 ng/mL).

In alternativa, due episodi di POR in seguito a stimolazione massima sono sufficienti per definire una paziente come *poor-responder* in assenza di età materna avanzata o anomalie dell'ORT.

Previsione di hyper-response e OHSS

Esagerata risposta ovarica a gonadotropine

- Più di 15 ovociti recuperati (Broer et al Human Reprod Update 2010);
- 15-20% of IVF treatments can occur: mild, moderate OHSS;
- 1-3% of IVF treatments: severe (Practice Committee of ASRM 2008);
- Sensitività 90,5%;
- Specificità 81,3%.

AMH *cut of value* per hyper response 3,5 ng/ml.

Sindrome dell'ovaio policistico (PCOS)

Nelle pazienti affette da sindrome dell'ovaio policistico, i livelli di AMH risultano spesso aumentati a causa dell'elevato numero di piccoli follicoli antrali presenti nelle ovaie. Per questo motivo, l'AMH è stato proposto anche come possibile *marker* diagnostico della PCOS, sebbene non sia ancora universalmente incluso nei criteri diagnostici ufficiali.

Oncofertilità

In oncologia riproduttiva, l'AMH viene utilizzato per valutare il danno ovarico indotto da trattamenti antitumorali. Il monitoraggio dei livelli di AMH prima e dopo chemioterapia permette di stimare la compromissione della funzione ovarica e di pianificare eventuali strategie di preservazione della fertilità.

Limiti dell'AMH

Nonostante i numerosi vantaggi, l'AMH presenta alcuni limiti. Innanzitutto, i valori possono variare in base al metodo di laboratorio utilizzato, rendendo talvolta difficile il confronto tra diversi studi clinici.
Inoltre, l'AMH rappresenta un indice quantitativo della riserva ovarica, ma non fornisce informazioni dirette sulla qualità ovocitaria, che dipende invece dall'età della donna. Pertanto, una paziente giovane con AMH basso può comunque avere una buona probabilità riproduttiva rispetto a una donna più anziana con valori apparentemente normali.
Un ulteriore limite riguarda l'assenza di valori soglia universalmente accettati, poiché i livelli considerati "normali" possono differire in funzione dell'età e del contesto clinico.

Conclusioni

L'ormone antimülleriano rappresenta oggi uno dei più importanti *marker* della riserva ovarica nella pratica clinica ginecologica e riproduttiva. Grazie alla sua correlazione con il numero di follicoli residui e alla relativa stabilità durante il ciclo mestruale, l'AMH costituisce uno strumento utile nella valutazione della fertilità femminile, nella personalizzazione dei trattamenti di PMA e nella preservazione della fertilità.
Tuttavia, il dosaggio dell'AMH deve essere interpretato nel contesto clinico complessivo della paziente, considerando età, anamnesi e altri parametri endocrini e strumentali. Futuri studi saranno necessari per standardizzare i metodi di dosaggio e migliorare ulteriormente il valore predittivo di questo biomarcatore.

Bibliografia

1. *LG ESHRE 2025*
2. *Cate RL, Mattaliano RJ, Hession C, Tizard R, Farber NM, Cheung A, Ninfa EG, Frey AZ, Gash DJ, Chow EP. Isolation of the bovine and human genes for Mullerian inhibiting substance and expression of the human gene in animal cells. Cell 1986;45:685–698*
3. *Weenen C, Laven JS, Von Bergh AR, Cranfield M, Groome NP, Visser JA, Kramer P, Fauser BC, Themmen AP. Anti-Mullerian hormone expression pattern in the human ovary: potential implications for initial and cyclic follicle recruitment. Mol Hum Reprod 2004;10:77–83*
4. *Broekmans et al (2008) Anti-Müllerian hormone and ovarian dysfunction. Trends Endocrinol and Metab 19:340-347*
5. *HI, Sammel MD, Freeman EW, Lin H, DeBlasis T, Gracia CR. Body size affects measures of ovarian reserve in late reproductive age women. Menopause 2008;15:857–861*
6. *La Marca A, De Leo V, Giulini S, Orvieto R, Malmusi S, Giannella L, Volpe A. Anti-Mullerian hormone in premenopausal women and after AMH and ART 127*
7. *Bentzen JG, Forman JL, Pinborg A, Lidegaard O, Larsen EC, Friis-Hansen L, Johannsen TH, Nyboe Andersen A. Ovarian reserve parameters: a comparison between users and non-users of hormonal contraception. Reprod Biomed Online 2012; 25:612–619.*
8. *Van den Berg MH, van Dulmen-den Broeder E, Overbeek A, Twisk JW, Schats R, van Leeuwen FE, Kaspers GJ, Lambalk CB. Comparison of ovarian function markers in users of hormonal contraceptives during the hormone-free interval and subsequent natural early follicular phases. Hum Reprod 2010;25:1520–1527*
9. *Kallio S, PuurunenJ, RuokonenA, Vaskivuo T, Piltonen T, TapanainenJS. Antimullerian hormone levels decrease in women using combined contraception independently of administration route. Fertil Steril2013;99:1305–1310*
10. *Practic Committee ASRM. Testing and interpreting measures of ovarian reserve: a committee opinion. Fertil Steril 2012;98:1407–1415*
11. *Lekamge DN, Barry M, Kolo M, Lane M, Gilchrist RB, Tremellen KP. Anti-Mullerian hormone as a predictor of IVF outcome. Reprod Biomed Online 2007;14:602–610.*
12. *Nelson SM, Yates RW, Fleming R. Serum anti-Mullerian hormone and FSH: prediction of live birth and extremes of response in stimulated Nelson SM, Yates RW, Fleming R. Serum anti-Mullerian hormone and FSH: prediction of live birth and extremes of response in stimulated cycles–implications for individualization of therapy. Hum Reprod 2007;22:2414–2421.*
13. *Hendriks DJ, Mol BW, Bancsi LF, Te Velde ER, Broekmans FJ. Antral follicle count in the prediction of poor ovarian response and pregnancy after in vitro fertilization: a meta-analysis and comparison with basal follicle-stimulating hormone level. Fertil Steril 2005;83:291–301).*
14. *European Society of Human Reproduction and Embryology (ESHRE 2011)*
15. *Lee TH, Liu CH, Huang CC, Wu YL, Shih YT, Ho HN, Yang YS, Lee MS.Serum anti- Mu¨llerian hormone and estradiol levels as predictors of ovarian hyperstimulation syndrome in assisted reproduction technology cycles. Hum Reprod 2008;23:160–167*
16. *Clinical Gynecologic Endocrinology and Infertility, Speroff L., Fritz M.A. Clinical Gynecologic Endocrinology and Infertility. 8th ed. Lippincott Williams & Wilkins, 2010.*
17. *Broer S. L. et al. "Anti-Müllerian hormone: ovarian reserve testing and its potential clinical implications". Human Reproduction Update, 2014; 20(5):688–701.*
18. *La Marca A., Volpe A. "Anti-Müllerian hormone (AMH) in female reproduction: is measurement of circulating AMH a useful tool?". Clinical Endocrinology, 2006; 64(6):603–610.*
19. *Dewailly D. et al. "The physiology and clinical utility of anti-Müllerian hormone in women". Human Reproduction Update, 2014; 20(3):370–385.*
20. *ESHRE. Guideline on ovarian stimulation 2025*
21. *ASRM. "Testing and interpreting measures of ovarian reserve". Fertility and Sterility, 2020.*
22. *Nelson S. M. et al. "Anti-Müllerian hormone-based approach to controlled ovarian stimulation for assisted conception". Human Reproduction, 2009; 24(4):867–875.*

GLI EFFETTI DEL LASER VAGINALE CO2 SULL'OUTCOME GRAVIDICO NELL'INFERTILITÀ

Dott. Giovanni Rosati - Specialista in Ginecologia, Ostetricia e PMA, Centro Embryos, Battipaglia (SA)

Con il termine "microbiota vaginale" si intende l'insieme di quei microrganismi (in particolar modo batteri) che vivono fisiologicamente nell'ambiente vaginale contribuendo all'omeostasi dell'apparato genitale femminile.
Nella stragrande maggioranza delle donne in età fertile la composizione del microbiota vaginale è dominata dalla presenza dei Lattobacilli, batteri a struttura bastoncellare, Gram-positivi, anaerobi facoltativi. Tra i lattobacilli più noti ricordiamo: *Lactobacillus Crispatus, Lactobacillus Gasseri, Lactobacillus Jensenii, Lactobacillus Iners.* Questi batteri, spesso denominati con l'appellativo di "batteri buoni", svolgono numerose funzioni, volte al mantenimento dello stato di salute dell'ambiente vaginale.
Tra le funzioni più importanti ricordiamo la produzione di acido lattico, che contribuisce al mantenimento del caratteristico pH acido della vagina (3,8-4,5) e che impedisce la proliferazione di microrganismi patogeni. Inoltre, i lattobacilli hanno anche la capacità di produrre sostanze antimicrobiche (come il perossido di idrogeno) che presentano una notevole funzione battericida.
Il microbiota vaginale svolge innumerevoli funzioni che ricoprono un ruolo critico nella vita delle donne:

- Protegge dalle infezioni vaginali;
- Riduce il rischio di contagio e trasmissione delle MST;
- Dona un notevole contributo alla fertilità;
- Supporta l'evoluzione della gestazione.

Diversi studi, negli anni, hanno evidenziato un forte collegamento tra il microbiota vaginale e la fertilità femminile. Difatti, l'equilibrio dei batteri vaginali può influenzare il naturale concepimento, l'impianto embrionario così come il successo delle tecniche di procreazione medicalmente assistita.
È ormai noto che un ambiente vaginale equilibrato, dominato in particolar modo dalla presenza di Lattobacilli, faciliti la sopravvivenza spermatica e il passaggio degli spermatozoi dapprima verso l'utero e successivamente verso le tube. Spesso, però, questo equilibrio è minato dalla disbiosi vaginale che comporta un aumento della proliferazione dei microrganismi opportunisti e/o patogeni.
Nella disbiosi vaginale, infatti, si assiste ad un aumento del fisiologico pH (>4,5), ad un incremento dell'infiammazione locale e ad un tasso maggiore di infezioni batteriche (causate ad esempio da batteri come la *Gardnerella Vaginalis*). Numerosi sono i fattori che causano disbiosi vaginale: abuso di antibiotici e/o antimicotici e di lavande vaginali, rapporti non protetti, numerosi partner sessuali, stress, squilibri ormonali, utilizzo di IUD.

La disbiosi vaginale può pertanto influenzare il concepimento, sia naturale che artificiale (ottenuto tramite tecniche di PMA).
Le interferenze con il concepimento sono molteplici:

- Un ambiente dominato dai processi infiammatori e da un pH meno acido comporta una ridotta sopravvivenza spermatica, causata in particolar modo dalla riduzione della motilità degli spermatozoi e dall'aumento dello stress ossidativo;
- La presenza di un numero inferiore rispetto alla norma di lattobacilli comporta una modificazione della composizione del muco cervicale in termini di quantità, viscosità e qualità, rendendo in questo modo difficile il passaggio delle cellule germinali maschili;
- Le infezioni vaginali possono risalire anche a livello della cavità uterina, andando a compromettere la recettività endometriale.

In questo modo si vanno a ridurre le possibilità di impianto embrionario e ad aumentare le possibilità di aborto spontaneo e/o precoce.
L'approccio alla risoluzione della disbiosi vaginale deve essere ovviamente personalizzato in base alla storia clinica, all'eziologia e ai segni e/o sintomi.
L'obiettivo principale, oltre alla riduzione dei fastidi avvertiti dalla paziente, consiste nel ristabilire l'omeostasi dell'ambiente vaginale andando a ripristinare la dominanza dei lattobacilli e il fisiologico pH acido.
La valutazione iniziale consiste in una visita ginecologica approfondita, completata dall'effettuazione di un tampone vaginale con esame microbiologico.
Se la disbiosi risulta essere causata da una vaginosi batterica o da una candidosi recidivante, allora la risoluzione è soprattutto di natura farmacologica, ovvero con l'assunzione da parte della paziente di antibiotici e/o antimicotici specifici combinati con l'utilizzo di probiotici. Se la disbiosi, invece, risulta essere aspecifica, quindi senza una causa scatenante di natura infettiva, allora l'approccio consiste o nell'assunzione di probiotici per via orale, o nell'applicazione di gel vaginali a base di acido lattico volti al ripristino del pH vaginale, oppure nell'effettuazione di diverse sedute di laser vaginale CO2.

Il laser vaginale CO2 è un trattamento ginecologico, non chirurgico, volto al miglioramento del trofismo, dell'elasticità e dell'idratazione della mucosa vaginale.
Il laser agisce creando micro-stimolazioni termiche controllate sulla mucosa vaginale. In questo modo si va a stimolare la produzione di collagene ed elastina che migliorano l'idratazione e la vascolarizzazione della zona interessata e trattata.
Sebbene l'applicazione di questa tipologia di tecnica-laser sia ascrivibile soprattutto all'atrofia vaginale, alla dispareunia e alla lieve incontinenza urinaria, il suo utilizzo si è rivelato essere di fondamentale importanza anche come supporto nella disbiosi vaginale recidivante e/o aspecifica. Il protocollo prevede di solito circa 5/6 sedute, con una cadenza temporale di circa 1 seduta al mese, con richiami annuali se necessari.
I benefici che le pazienti traggono da questo trattamento sono numerosi:

- Aumento dell'idratazione della mucosa vaginale grazie

all'incremento della produzione di collagene;
- Aumento della tonicità dei tessuti vaginali grazie ad una maggiore produzione di elastina;
- Riduzione di bruciore e fastidi intimi in quanto rinforza la mucosa vaginale;
- Risoluzione dei problemi di lieve incontinenza urinaria grazie al supporto fornito anche a livello dei tessuti uretrali e della cupola vaginale.

Stimolando l'aumento della produzione di collagene, il laser vaginale CO2 conferisce alle pareti vaginali maggiore tono ed elasticità e favorisce un incremento della lubrificazione. La lubrificazione vaginale e la disbiosi vaginale sono strettamente collegate, in quanto l'equilibrio della flora batterica influisce direttamente sulla salute della mucosa e quindi sulla sua capacità di produrre secrezioni naturali.
Le secrezioni vaginali, a loro volta, sono coinvolte nel mantenimento dell'omeostasi dell'ambiente vaginale, in quanto mantengono stabile il valore del pH, fungono da barriera nei confronti degli agenti patogeni e creano l'ambiente ideale per il transito degli spermatozoi.

Il laser vaginale CO2 gioca un ruolo fondamentale anche nel supporto e nella cura della lieve incontinenza urinaria. Questa condizione, molto comune, può essere caratterizzata dalla perdita di poche gocce d'urina così come da minzioni abbastanza frequenti e ripetute. Il laser agisce rinforzando la cupola vaginale, prevenendo l'insorgenza di condizioni patologiche come il cistocele e limitando le perdite di urina che possono contaminare l'ambiente vaginale provocando infezioni ed infiammazioni recidivanti che, a loro volta, possono influire sull'*outcome* gravidico.
Il laser vaginale CO2 stimola la rigenerazione tissutale incrementando anche la produzione di elastina. L'elastina è una proteina che consente ai tessuti di allungarsi e poi di ritornare alla loro morfologia iniziale. Provocando, quindi, un miglioramento strutturale delle pareti vaginali, questa tecnica laser comporta una diminuzione significativa della dispareunia, che rappresenta un vero e proprio ostacolo al concepimento naturale.
In conclusione, sottoporre le pazienti affette da disbiosi vaginale aspecifica, lieve incontinenza urinaria e dispareunia ad almeno 5/6 sedute di laser-terapia vaginale CO2, comporta un miglioramento dell'*outcome* gravidico, sia naturale che artificiale (ossia tramite l'ausilio di tecniche di fecondazione assistita).

Bibliografia

1. *Microbiota in vaginal health and pathogenesis of recurrent vulvovaginal infections: a critical review Namarta Kalia 1, Jatinder Singh 2, Manpreet Kaur 3*
2. *Microbial diversity in the vaginal microbiota and its link to pregnancy outcomes Agnes Baud 1, Kenzo-Hugo Hillion 1, Céline Plainvert 2, Véronique Tessier 3, Asmaa Tazi 2, Laurent Mandelbrot 4, Claire Poyart 2 5, Sean P Kennedy 6*
3. *Laser therapy for the restoration of vaginal function Marco Gambacciani 1, Santiago Palacios 2*

IL POTENZIALE, I PERICOLI E LE INSIDIE DELL'INTELLIGENZA ARTIFICIALE (IA) NELLE TECNICHE DI PROCREAZIONE MEDICALMENTE ASSISTITA (PMA)

Dott.ssa Elisabetta Trabucco – Direttrice del Centro di Fecondazione Assistita Genera, Napoli

Abstract

Le applicazioni dell'intelligenza artificiale (IA) sono aumentate in modo esponenziale in tutti gli aspetti della vita, compreso l'ambito delle Tecniche di Procreazione Medicalmente Assistita (PMA). Sebbene gli strumenti assistiti dall'IA mostrino enormi promesse, vi sono limitazioni preoccupanti e cautele da tenere in considerazione prima che venga adottato un loro uso diffuso nella pratica clinica quotidiana. Lo scopo di questa revisione è offrire agli operatori sanitari e ai ricercatori una panoramica delle applicazioni dell'IA e assisterli nel valutare e analizzare in modo indipendente la ricerca sull'IA nella sua applicazione alla PMA.

Esaminiamo diversi tipi di tecnologie di IA esplorando al contempo la ricerca attuale sulle sue applicazioni all'interno della PMA. Le applicazioni specifiche includono la selezione di embrioni e gameti e l'IA come sistema di supporto alle decisioni cliniche. Riassumiamo i vantaggi dell'integrazione dell'IA in un contesto di PMA e descriviamo i limiti e le sfide noti di questa tecnologia nella sua applicazione alle nostre missioni di clinici, educatori e studiosi. Consideriamo le implicazioni etiche dell'IA all'interno della PMA e forniamo a clinici e ricercatori un quadro di riferimento per valutare, esaminare e rivedere criticamente la ricerca che utilizza l'IA specificamente nel contesto della PMA. Presentiamo le attuali linee guida e la governance internazionale che possono essere adattate alla pubblicazione e alla rendicontazione della ricerca sull'IA nella PMA.

Introduzione

L'intelligenza artificiale (IA) ha fatto irruzione in tutti gli aspetti della vita quotidiana, compresi la medicina, la ricerca e non solo[1]. L'accessibilità, la funzionalità e la facilità d'uso rendono l'IA uno strumento affascinante e potenzialmente trasformativo, che si prevede avrà un impatto positivo sulla cura del paziente e sui risultati clinici. Tuttavia, questi primi giorni dell'IA sono caratterizzati da una regolamentazione amministrativa minima e da una convergenza di influenze che stanno portando a una crescita a ritmo esponenziale, dovuta a miliardi di spese in capitale annuali da parte di imprese private globali. Tutto ciò contribuisce a un senso di meraviglia ed entusiasmo per il potenziale di trasformazione del pensiero sociale e della produttività.

Un'aspirazione decantata dell'IA è quella di ridurre il carico finanziario della sanità, abbassando il costo dell'intelligenza e del lavoro cognitivo e potenziando al contempo il processo decisionale clinico. Tuttavia, vi sono serie e concomitanti preoccupazioni circa l'impatto sulla sicurezza del lavoro, l'affidabilità del prodotto, la riservatezza dei dati e il modo in cui l'IA possa ridurre il nostro pensiero profondo come specie [2]. In quanto utilizzatori di questa tecnologia innovativa e potente, abbiamo la responsabilità di approcciare questo rapido cambiamento della società sia con curiosità che con prudente cautela.

Nel campo dell'Endocrinologia della Riproduzione e dell'Infertilità (REI), le tecniche di procreazione medicalmente assistita (PMA) sono state all'avanguardia nell'adozione di nuove tecnologie volte a migliorare i risultati dei trattamenti riproduttivi dei pazienti [3]. Si stima che l'infertilità colpisca 186 milioni di persone in tutto il mondo [4]. Enormi progressi tecnici, sociali e clinici hanno portato all'uso diffuso della fecondazione in vitro (FIVET) nel corso di quasi 50 anni dalla nascita del primo bambino da FIVET nel 1978; circa il 2,6% di tutti i bambini nati nel 2023 negli Stati Uniti e più di 10 milioni di bambini nel mondo sono stati concepiti tramite FIVET [5]. In quanto professionisti medici nel campo della PMA e utenti *leader* di tali tecnologie mediche in evoluzione, abbiamo la responsabilità di essere i primi a vagliarle.

Dobbiamo esaminare attentamente i limitati dati a disposizione e, se una adeguata verifica mostra che tali progressi sono promettenti e che i benefici superano i rischi, diventiamo spesso i primi ad adottare tecnologie all'avanguardia. Questo articolo affronta i benefici e le applicazioni dell'IA nel campo della PMA, evidenziando al contempo i potenziali pericoli, le insidie e le sfide di questa tecnologia nella sua applicazione alle nostre missioni di clinici, educatori e studiosi. Soprattutto, mira a fornire agli specialisti in infertilità gli strumenti per valutare, esaminare e rivedere criticamente la ricerca sulla PMA che utilizza metodologie di IA, fornendo al contempo un'ampia panoramica delle relative piattaforme di IA.

Definizione delle tecnologie di IA utilizzate nella PMA

Cos'è l'Intelligenza Artificiale (IA)?

Per decenni, i progressi nella programmazione informatica hanno permesso ai compiti umani quotidiani di diventare efficienti, automatizzati e precisi. L'IA descrive un insieme di sistemi informatici in grado di eseguire compiti più complessi, che generalmente richiedono un pensiero umano di ordine superiore (es. processo decisionale, ragionamento, pensiero creativo) [6] e in grado di applicare conoscenze/esperienze preesistenti per risolvere nuovi problemi. I sistemi di IA possono eseguire compiti in circostanze imprevedibili utilizzando l'esperienza per migliorare le prestazioni, dimostrare capacità di pensiero cognitivo di livello superiore simili a quelle umane e agire razionalmente per raggiungere tali obiettivi [6].

Caratterizzare le diverse tecnologie di IA è difficile a causa dei rapidi progressi e delle sovrapposizioni tra i vari modelli. I primi modelli di IA negli anni '50 erano basati su regole, per cui la logica, gli alberi di ricerca e i codici manuali venivano usati per affrontare problemi limitati, come giocare una partita a scacchi. A ciò è seguito un breve periodo di mancanza di fondi e di interesse per l'IA, noto come "inverno dell'IA". Alla fine degli anni '90 e all'inizio dei primi anni 2000 sono emersi modelli basati sull'apprendimento automatico o *machine learning* (ML). Il

cambiamento fondamentale con questa tecnologia di IA è stato il passaggio dall'uso di regole all'uso di grandi volumi di dati, insieme alla capacità di apprendere nuovi modelli piuttosto che seguire regole preprogrammate.
All'interno del *machine learning* (ovvero l'IA che impara dai dati), i sistemi possono utilizzare l'apprendimento "supervisionato", "non supervisionato" o "per rinforzo". Ad esempio, l'apprendimento supervisionato comporta l'apprendimento da esempi predefiniti ed è spesso utilizzato per la previsione o la classificazione. L'apprendimento non supervisionato comporta la ricerca di pattern senza etichette predefinite. Il framework dell'apprendimento per rinforzo descrive un altro modello in cui l'apprendimento avviene per tentativi ed errori.
La maggior parte delle tecnologie di IA applicate oggi alla FIVET utilizza generalmente un sottoinsieme del ML denominato "apprendimento profondo" o *deep learning* (DL), in cui enormi set di dati vengono utilizzati per addestrare e sviluppare modelli che presentano molte reti neurali (NN) a più strati, in modo molto simile al cervello umano. Esistono numerosi tipi di NN, come le NN convoluzionali (CNN), che utilizzano filtri convoluzionali per rilevare pattern locali, e le NN artificiali (ANN), che presentano nodi artificiali (che agiscono come neuroni), tra le altre. I filtri convoluzionali sono piccole matrici di apprendimento che scansionano i dati in input e possono rilevare pattern, mentre i nodi artificiali sono unità computazionali all'interno delle NN che possono ricevere input e trasmettere dati ad altri strati (proprio come i neuroni biologici)[7].
Altre tecnologie di IA comunemente riscontrate includono l'elaborazione del linguaggio naturale (NLP) per comprendere e generare il linguaggio, l'IA generativa che crea nuovi contenuti, o l'IA multimodale in grado di elaborare simultaneamente più tipi di dati. L'IA generativa comprende i grandi modelli linguistici (LLM) e i grandi modelli multimodali (LMM), che utilizzano rispettivamente input linguistici e multimediali [7, 8].

La Tabella 1 confronta i due tipi di strumenti di IA generativa e i loro possibili utilizzi.

È importante notare che nessuno dei modelli di IA generativa è validato come strumento medico diagnostico, pertanto dovrebbero essere usati con cautela nel contesto della medicina, proprio come i medici/ricercatori usano qualsiasi altro riferimento, come motori di ricerca o libri di testo. OpenEvidence, un'IA generativa specializzata, può essere utilizzata per informare il processo decisionale medico in modo simile ad altre piattaforme come *UpToDate*.

Per confermare la qualità e la legittimità della risposta dell'IA a qualsiasi quesito, raccomandiamo una verifica controllando e convalidando i riferimenti bibliografici primari che vengono citati come fonti primarie di dati e come base per le conclusioni e le raccomandazioni offerte dalla piattaforma di IA generativa. Una volta che questo passaggio ha verificato la presenza di riferimenti legittimi e di qualità accettabile, si consiglia di aggiungere un ulteriore livello di revisione confermando le informazioni con un'altra fonte indipendente e fidata (es. linee guida di società professionali note e/o *Cochrane Reviews*) prima di intraprendere qualsiasi azione clinica.
Le applicazioni dell'IA nella sanità possono essere generalmente categorizzate in 5 ampi domini [7]. L'IA può essere utilizzata come strumento predittivo (usare i dati per prevedere esiti futuri), strumento di classificazione (smistare i dati in normali/anormali), strumento di associazione (trovare relazioni tra diverse variabili per trarre conclusioni cliniche), strumento di regressione (valutare la forza della relazione di una variabile rispetto ad altre variabili) o strumento di ottimizzazione (automatizzare compiti pratici, es. compiti amministrativi) [7].
Nella sezione successiva viene fornita una breve panoramica dell'IA nella PMA, che prevede principalmente lo sfruttamento delle capacità predittive, di classificazione e di ottimizzazione dell'IA.

Garg and Seifer • Reproductive Biology and Endocrinology (2026) 24:36

Tabella 1: Un confronto tra gli strumenti di Intelligenza Artificiale Generativa comunemente utilizzati

Tipo di IA generativa	Esempi	Pro	Contro	Possibili utilizzi
Fondazionale: Modelli ampi e generalizzati, addestrati su vasti set di dati e utilizzati dal grande pubblico.	ChatGPT Gemini Claude	• Insieme di conoscenze più ampio. • Più potente in quanto addestrato su dataset più grandi. • Migliore nel ragionamento logico generale.	• Manca di focus e profondità specifica sui singoli argomenti. • Manca di tutele specifiche per determinati settori o applicazioni. • Non è conforme alla normativa HIPAA.	• Generare riassunti su un argomento. • Generare testi complessi. • Generare immagini.
Specializzata: Modelli ottimizzati (fine-tuned) per casi d'uso specifici o settori industriali verticali, utilizzati da professionisti medici certificati in uno specifico campo di competenza.	Open Evidence	• Integra conoscenze e tutele (guardrails) specifiche. • Possiede una conoscenza più approfondita nel campo della medicina. • È conforme alla normativa HIPAA.	• Non può generare immagini. • Funzionalità limitate.	• Esplorare quesiti medici specifici utilizzando una ricerca medica ad alta precisione. • Sintesi della letteratura con fonti primarie, basandosi su riviste scientifiche peer-reviewed ad alto impatto.

Applicazioni dell'IA nella PMA

Selezione degli embrioni

La selezione di embrioni di alta qualità è una fase cruciale della FIVET per massimizzare le possibilità di una gravidanza di successo [13]. Tradizionalmente, ciò avviene tramite la selezione embrionale basata sulla morfologia, durante la quale un embriologo esperto ispeziona visivamente e seleziona gli embrioni dall'aspetto più sano; tuttavia, questo metodo è soggetto a variabilità intra-osservatore e soggettività [14]. La selezione embrionale tradizionale si basa sull'annotazione delle immagini dell'embrione e sul posizionamento in graduatoria della qualità all'interno di una determinata coorte di embrioni [13]. Le immagini possono essere statiche (es. embrioni allo stadio di clivaggio valutati in base a criteri specifici) o possono utilizzare la videografia *time-lapse* (TLV), che si ritiene riduca la soggettività tra gli operatori e consenta l'identificazione di nuovi biomarcatori di vitalità [13, 15]. La TLV prevede l'acquisizione di immagini di microscopia ottica a intervalli di tempo prestabiliti per consentire la valutazione dello sviluppo embrionale [16]. Tuttavia, la sola TLV non ha dimostrato di migliorare i tassi di gravidanza clinica o di nati vivi rispetto alla selezione morfologica standard da immagini statiche [17]. Diverse tecnologie di IA che incorporano immagini statiche, TLV e/o parametri clinici aggiuntivi sono state sviluppate con l'obiettivo di migliorare la selezione degli embrioni.

Diversi studi individuali riportano esiti differenti quando confrontano i modelli di IA con la pratica corrente o come modello di previsione per la selezione degli embrioni. Uno dei più grandi studi controllati randomizzati in doppio cieco di non inferiorità ha confrontato la valutazione morfologica standard rispetto a un modello di IA basato su DL in 14 centri in Australia ed Europa. Su 1066 pazienti, il tasso di gravidanza clinica è stato del 46,5% nel gruppo IA rispetto al 48,2% nel gruppo con classificazione morfologica, mostrando una mancanza di non inferiorità per il deep learning in termini di tasso di gravidanza clinica *[Deep learning versus manual morphology-based embryo selection in IVF. A randomized, double-blind noninferiority trial. Nat Med. 2024;30(11):3114-20]*.

Wang e colleghi hanno confrontato prospetticamente la classificazione dell'embriologo (sistema di classificazione di *Gardner*) rispetto all'IA (utilizzando TLV e modelli di *deep learning*) e hanno dimostrato tassi di impianto migliori nel gruppo IA (81% contro 68%, p =0.02); tuttavia, non vi è stata alcuna differenza in termini di nati vivi o tassi di aborto spontaneo [19]. Una meta-analisi di 20 studi ha dimostrato come i modelli di IA abbiano previsto la probabilità di una gravidanza clinica di successo con maggiore accuratezza (accuratezza mediana dell'81,5%) rispetto agli embriologi clinici (accuratezza mediana del 51%) [20], suggerendo che la selezione embrionale basata sull'IA potrebbe superare quella guidata dall'uomo; tuttavia, i dati provenienti da studi randomizzati su larga scala e di alta qualità che valutino il tasso di nati vivi come esito sono attualmente carenti.

Le tecnologie di IA che utilizzano la morfologia e/o i parametri clinici possono essere impiegate anche come modelli predittivi per rilevare la qualità dell'embrione, lo stato di ploidia, l'aborto spontaneo, la gravidanza e i tassi di nati vivi come esiti. Uno studio retrospettivo ha dimostrato che un modello di selezione embrionale basato su IA a immagini statiche aveva un'accuratezza del 67% nel prevedere gli aborti nel primo trimestre [21]. Diversi studi hanno valutato algoritmi basati sull'IA che utilizzano tecnologie ML e ANN, dimostrando AUC (Area sotto la curva) comprese tra 0,70 e 0,84 in termini di previsione del potenziale di impianto [22-26]. Un altro modello di IA basato su DL, addestrato su 10.378 embrioni, ha utilizzato immagini statiche e informazioni cliniche per prevedere lo stato di ploidia negli embrioni, mostrando un'accuratezza del 70% [27]. Le tecnologie di IA sono state applicate anche alla metilazione del DNA dell'intero genomico, che si ritiene sia un biomarcatore per valutare la qualità dell'embrione, dove i dati preliminari suggeriscono che i modelli basati sull'IA possono prevedere con successo i tassi di nati vivi e aiutare nella selezione degli embrioni (con un'AUC di 0,9) [28].

Selezione dei gameti

Selezionare spermatozoi e ovociti della massima qualità durante un ciclo di FIVET è essenziale per ottenere una fecondazione ottimale. Valutare la qualità dei gameti e prevedere i tassi di fecondazione è utile anche per fornire consulenza ai pazienti, ad esempio per coloro che si sottopongono a crioconservazione degli ovociti. La capacità di prevedere gli esiti in base alla qualità/numero di ovociti prelevati può guidare il processo decisionale in termini di pianificazione del numero di ovociti necessari per adattarsi agli obiettivi di pianificazione familiare della paziente [29].

Le tecniche tradizionali di selezione degli spermatozoi includono la valutazione della morfologia, della motilità e dell'integrità del DNA, mentre le tecniche avanzate per gli spermatozoi mancano di dati di alta qualità che dimostrino un miglioramento degli esiti clinici [30]. Nella maggior parte dei contesti clinici, un embriologo seleziona manualmente gli spermatozoi in base ai criteri dell'Organizzazione Mondiale della Sanità (OMS) [31]. Questo è un processo che richiede tempo, con una significativa variabilità inter- e intra-osservatore.

Sono stati sviluppati diversi sistemi di IA per valutare la morfologia degli spermatozoi, con i modelli più comuni che utilizzano reti neurali profonde (DNN). Gli analizzatori di spermatozoi assistiti da computer (CASA) esistono dagli anni '80 e possono ridurre la soggettività dell'operatore; l'ultimo manuale dell'OMS sull'analisi dello sperma ha infatti riconosciuto come i sistemi CASA possano determinare accuratamente la concentrazione spermatica e i parametri di motilità progressiva attraverso colorazioni fluorescenti del DNA e algoritmi di rilevamento della coda [32]. Gli algoritmi di previsione ML che utilizzano reti neurali possono raggiungere un'accuratezza fino al 92,9% nel prevedere una concentrazione spermatica anormale e dell'85,7% nel prevedere anomalie spermatiche [33].

Oltre alla previsione, le tecnologie di IA basate su DL possono valutare anche la motilità, la morfologia e la frammentazione del DNA. Le metodologie DL che utilizzano video sono state in grado di consentire la correlazione dei pattern cinetici di motilità per supportare la selezione per l'iniezione intracitoplasmatica di spermatozoi (ICSI) [34]. In termini di morfologia, l'OMS descrive 11 diverse anomalie della testa dello spermatozoo, tuttavia può essere difficile classificarle accuratamente a occhio nudo [35]. I progressi del DL che utilizzano reti neurali possono rilevare e classificare oggettivamente queste anomalie morfologiche in tempo reale [36]. I test attuali che valutano la frammentazione del DNA sono invasivi, il che significa che i singoli spermatozoi selezionati non possono essere successivamente utilizzati per la ICSI. Non esistono linee guida sull'uso dell'analisi della

frammentazione del DNA nella selezione degli spermatozoi, tuttavia è stato dimostrato che le tecnologie di IA che utilizzano reti neurali prevedono con successo l'integrità del DNA utilizzando singole immagini, eliminando la necessità di test invasivi [37].

Insieme a uno spermatozoo sano, anche la selezione di un ovocita maturo è vitale per consentire la fecondazione. Per valutare veramente la maturità nucleare, deve essere confermata l'estrusione del globulo polare, il che richiede la rimozione del cumulo ooforo [38]. Attualmente, la maturità nucleare viene valutata in modo non invasivo e visivo dagli embriologi, in modo molto simile ad altri aspetti della fecondazione descritti sopra. Gli ovociti vengono valutati in base a vari parametri. Numerosi sistemi di IA basati su DL sono stati sviluppati per assistere questo processo. Gli esempi includono VIOLET™ (*Future Fertility*), in grado di analizzare le caratteristiche morfologiche di immagini 2D e prevedere la fecondazione con un'accuratezza del 91,2%, risultando così più veloce e più accurato degli embriologi esperti [39]. Oltre all'analisi basata sulle immagini, i modelli di IA possono potenziare i test non invasivi di espressione genica. L' "OsteraTest" utilizza il ML e la bioinformatica per prevedere in modo non invasivo lo sviluppo dell'ovocita con un'accuratezza dell'86% [40]. Sebbene siano necessari RCT su larga scala e successive meta-analisi, tali applicazioni dell'IA per migliorare la selezione di spermatozoi e ovociti rappresentano strade promettenti.

Processo decisionale clinico

Una delle maggiori sfide nella FIVET è la previsione e l'ottimizzazione degli esiti, la stratificazione del rischio di effetti avversi e, di conseguenza, la personalizzazione dei protocolli di trattamento per bilanciare e valutare questi fattori. Ad esempio, si potrebbe massimizzare il numero di ovociti prelevati utilizzando un approccio più aggressivo alla stimolazione ovarica; tuttavia, ciò potrebbe aumentare il rischio di sindrome da iperstimolazione ovarica (OHSS). Pertanto, gran parte della relazione medico-paziente consiste nel mettere a punto i protocolli in base alle circostanze mediche e personali individuali (es. la tolleranza al rischio del paziente, l'accesso ai finanziamenti, lo stato assicurativo, le tempistiche personali e la motivazione) e nel fornire una cura olistica che racchiuda gli obiettivi a breve e lungo termine del paziente. Decisioni così complesse e stratificate vengono spesso prese sulla base del giudizio clinico, dell'esperienza e del processo decisionale condiviso. Le tecnologie di IA hanno mostrato promesse nell'assistere in vari bivi del processo decisionale clinico.

Durante la consulenza pre-trattamento, vengono utilizzati numerosi parametri per informare i pazienti sulle loro possibilità di successo, tra cui l'età, la riserva ovarica, l'etnia e la storia dei trattamenti precedenti. I modelli di IA vengono sviluppati e utilizzati da pazienti, compagnie assicurative e cliniche per fornire l'accesso alle cure (es. offrire programmi finanziari a rischio condiviso) [41]. Tali modelli possono anche ridurre i tempi durante l'accettazione iniziale e la valutazione dei servizi. Ad esempio, un modello basato su LLM intitolato *"Fast Track to Fertility"* ha dimostrato di aumentare l'accesso dei pazienti del 24% e di ridurre il tempo necessario per il trattamento del 50%, con un maggiore coinvolgimento dei pazienti [42].

I protocolli di stimolazione ovarica e di induzione dell'ovulazione (*trigger*) utilizzati sono aspetti chiave di un ciclo di FIVET. Esistono varie permutazioni di protocolli e l'ottimizzazione dei regimi farmacologici può massimizzare la resa follicolare e la possibilità di un nato vivo di successo [43]. È stato dimostrato che l'utilizzo di modelli di IA (per lo più basati sulla regressione) per assistere nel dosaggio delle gonadotropine migliora la sicurezza, riduce i tassi di cancellazione del ciclo e di OHSS e migliora la resa degli ovociti, la gravidanza clinica e i tassi di nati vivi [8]. Un grande studio condotto sul database nazionale statunitense (*SART CORS*) con 365.473 pazienti ha dimostrato come la modellizzazione tramite IA che utilizza i dati di pazienti simili possa essere impiegata per migliorare la resa degli ovociti e i tassi di nati vivi [44].

Le tecniche di ML possono anche supportare il processo decisionale relativo al momento in cui indurre la maturazione ovocitaria durante il ciclo di stimolazione, incorporando sia le immagini ecografiche che i valori di laboratorio. Studi retrospettivi mostrano che i follicoli tra i 16 e i 19 mm nel giorno del trigger sono più frequentemente associati a ovociti maturi, e l'utilizzo di tecnologie di IA può migliorare il numero di ovociti maturi e di blastocisti; tuttavia, mancano attualmente dati sui nati vivi [44-46].

Benefici dell'IA

Esistono numerosi vantaggi pratici e clinici delle tecnologie di IA, in particolare quando si ha accesso ai big data [47], come dimostrato nelle sezioni precedenti. L'IA è in grado di rilevare, riconoscere e classificare grandi quantità di dati per analizzare e prevedere esiti al di là di quanto ci si possa aspettare da qualsiasi sforzo umano, riducendo al contempo il tempo di completamento del compito e riducendo i bias soggettivi [24]. Ciò si è riflesso nel confronto dei risultati clinici sopra descritti, in cui le tecnologie di IA generalmente eguagliano o superano le prestazioni umane nei compiti presi in esame. Il fatto che le tecnologie di IA apprendano in tempo reale da grandi set di dati può anche migliorare la standardizzazione, la riproducibilità e la scalabilità. L'IA è inoltre meno incline a commettere errori rispetto agli esseri umani, migliorando così potenzialmente la validità, la riproducibilità e l'affidabilità dei risultati.

Le tecnologie di IA possono anche migliorare l'accessibilità ai servizi di FIVET che in molti paesi rimangono a finanziamento privato. L'IA è in grado di eseguire funzioni molto più velocemente rispetto ai compiti guidati dall'uomo, il che a sua volta può migliorare la produttività in un contesto clinico [42]. Ad esempio, l'IA che utilizza il DL può automatizzare il punteggio degli embrioni senza richiedere l'annotazione umana, eliminando potenzialmente la necessità di un punteggio manuale e risparmiando una notevole quantità di tempo [48, 49]. Sebbene vi siano costi associati alla configurazione delle tecnologie di IA, una volta operative esse possono contribuire a ridurre le spese a lungo termine in termini di forza lavoro, migliorando al contempo la produttività. Ad esempio, se l'IA viene integrata nella pratica clinica e riesce davvero a migliorare la selezione dei gameti e degli embrioni, potrebbe teoricamente ridurre il numero di fallimenti dei cicli.

Gli strumenti alimentati dall'IA, come la trascrizione vocale e la cartella clinica automatizzata, possono far risparmiare dal 37% al 46% del tempo del medico dedicato a compiti amministrativi come la registrazione di ammissioni, trasferimenti e dimissioni [50]. È stato inoltre riscontrato che l'integrazione di strumenti prognostici guidati dall'IA aumenta l'utilizzo della FIVET tra i nuovi pazienti in una clinica per la fertilità [51], suggerendo che l'IA possa aiutare ad ampliare l'accesso e la consapevolezza dei

trattamenti di FIVET.
Inoltre, le tecnologie di IA generativa (es. ChatGPT) sono in grado di adattare la propria interfaccia in base alla lingua, al livello di alfabetizzazione sanitaria (*health literacy*) e alla personalità dell'utente, il che le rende uno strumento potente per ampliare l'accesso alle cure e interagire con pazienti di contesti diversi [52]. Ad esempio, è stato dimostrato che il consenso chirurgico assistito da ChatGPT prima di un'artroplastica totale di ginocchio riduce i punteggi di ansia e migliora l'esperienza complessiva di ospedalizzazione del paziente rispetto al consenso tradizionale guidato dal chirurgo [53]. Inoltre, l'IA generativa si è dimostrata efficace nel generare accuratamente note e istruzioni di dimissioni in altre lingue, il che può migliorare le barriere linguistiche nell'assistenza sanitaria [54].
Numerosi modelli predittivi mirano ad anticipare la risposta della singola paziente alla stimolazione ovarica e a suggerire modi per migliorarla; tuttavia, non esiste un unico modo stabilito per "personalizzare" i piani di trattamento per i pazienti al fine di massimizzare le loro possibilità di successo [55, 56]. Ad esempio, un grande studio multicentrico che ha utilizzato l'IA è stato in grado di prevedere quali dimensioni dei follicoli migliorassero i tassi individuali di nati vivi [8]. Poiché l'IA può analizzare molti più punti dati rispetto al processo umano, porta con sé l'enorme speranza di potersi muovere verso la personalizzazione dei trattamenti individuali per i pazienti e di massimizzare i risultati dei loro trattamenti.
Pertanto, sebbene l'IA appaia promettente in ciascuno degli esempi di lavoro sopra citati, c'è ancora molta strada da fare prima che questi strumenti possano essere incorporati in modo affidabile nell'uso clinico. La sezione successiva discute le sfide attuali con le piattaforme di IA e la necessità di dati più robusti e di alta qualità per l'addestramento e la convalida, al fine di produrre modelli migliori che siano accurati, precisi e testati.

I pericoli dell'IA

Esperienza limitata e non comprovata dell'IA

Per quanto le tecnologie di IA sopra menzionate siano entusiasmanti e promettenti, dobbiamo usare cautela quando consideriamo il loro ruolo estensivo nella pratica clinica. Sebbene vi siano numerosi studi che confrontano gli strumenti guidati dall'IA con la pratica corrente, vi è una mancanza di studi controllati randomizzati prospettici, ampi e di alta qualità. Negli studi esistenti si riscontra una significativa eterogeneità tra i disegni di studio, la loro qualità e una mancanza di standardizzazione degli esiti, il che limita le conclusioni e la generalizzabilità che possiamo trarre dalle meta-analisi. A seconda dell'intervento testato, gli esiti possono variare dalla "non inferiorità" alla valutazione dell'impatto sui tassi di gravidanza clinica e di nati vivi. In definitiva, lo scopo della FIVET non è solo una nascita di successo, ma anche ottenere una gravidanza sana e un bambino sano, tutti elementi che costituiscono importanti considerazioni cliniche quando si valuta lo scopo della PMA.

Mancanza di trasparenza su come vengono prese le decisioni dal modello di IA

I sistemi di IA basati su ML e DL contengono una moltitudine di strati che sono spesso non trasparenti e nascondono l'esatta logica e i percorsi che guidano il processo decisionale. Ciò ha creato l'effetto di una "scatola nera" (*black box*), in cui la mancanza di trasparenza, interpretabilità e tracciabilità rende difficile per i professionisti avere una fiducia sufficiente in questi sistemi [13]. Una domanda importante da porsi è se questi strati nascosti aggiuntivi migliorino effettivamente i risultati clinici [57]. Ad esempio, un confronto tra 12 algoritmi che prevedono la vitalità delle blastocisti ha mostrato che la regressione logistica superava i sistemi di *machine learning* più complicati e di tipo *"black-box"* [58].
Se il medico che utilizza uno specifico strumento di IA per prendere una decisione clinica non riesce a comprendere appieno come tale strumento giunga a una conclusione, si crea un senso di sfiducia [59]. Questa naturale esitazione nell'utilizzare uno strumento "non spiegabile" è valida, poiché le tecnologie di IA sono a rischio di *bias* e di *overfitting* (sovra-adattamento) basati sulla presenza o assenza di bias nei dati utilizzati per costruire il modello stesso. La fiducia e l'accettazione da parte del medico dipendono dalla trasparenza della validazione della piattaforma di IA e dalle popolazioni di pazienti su cui si basa il suo *output*, poiché questi due elementi influenzeranno la generalizzabilità delle raccomandazioni dell'IA. In altre parole, la trasparenza potenzia il giudizio clinico mentre l'opacità enfatizza la vigilanza clinica.

Limitazioni dovute alla qualità e ai bias nei dati utilizzati per addestrare e convalidare i modelli di IA

Dato che i modelli di IA vengono addestrati utilizzando un particolare set di dati, qualsiasi limitazione nei dati stessi si riflette sulla funzionalità del modello. *L'overfitting* descrive un fenomeno in cui un modello apprende i dati di addestramento "fin troppo bene" e include il "rumore" di fondo che è irrilevante per la domanda posta. È importante sottolineare che questa non è una critica alla metodologia in sé, ma ai dati da cui l'IA impara e su cui viene poi convalidata. In definitiva, la capacità dell'IA di prevedere accuratamente un esito o di fungere da strumento diagnostico è limitata dalla qualità e dalla quantità di dati su cui è stata addestrata e convalidata [13, 60].
Esiste infatti un rischio elevato di *bias* con il ML e il DL a causa delle distorsioni presenti nel set di dati, e queste possono variare, includendo discriminazioni basate sul genere, sull'etnia e sullo status socioeconomico [61]. Ad esempio, uno studio ha dimostrato come un modello basato su ML utilizzato per prevedere le esacerbazioni dell'asma nei bambini producesse sovrastime errate per i bambini di status socioeconomico inferiore, suggerendo prestazioni peggiori del modello predittivo guidate da bias [62]. Le popolazioni emarginate e sottorappresentate sono spesso escluse a causa di barriere economiche e culturali all'accesso, limitando così la loro rappresentazione in qualsiasi database di addestramento o convalida per lo sviluppo di un modello di IA [63].

Responsabilità e rischio di disinformazione dovuti ad allucinazioni, dati fabbricati, data leakage e/o fragilità

All'interno del DL, i modelli LLM e LMM comportano un alto rischio di "allucinazione", per cui possono fabbricare dati falsi e attingere a informazioni inesistenti o suggerire esiti errati che in superficie possono sembrare plausibili ma che costituiscono informazioni false. Ciò solleva anche la questione della responsabilità. Chi è ritenuto responsabile di un particolare sistema di IA? Il team che lo ha sviluppato o l'individuo che lo sta utilizzando? Questo solleva questioni etiche più profonde

(discusse più avanti); tuttavia, l'IA deve essere vista come qualsiasi altro strumento utile in un contesto clinico (es. un emocromo completo) e la responsabilità clinica finale ricade sul medico.
Anche la gestione dei dati da parte dell'IA per trarre conclusioni cliniche comporta dei rischi. Ad esempio, le applicazioni di ML sono a rischio di *data leakage* (perdita di dati), *undercutting* e *data poisoning* (avvelenamento dei dati). Il *data leakage* si verifica quando i dati utilizzati per l'addestramento "trapelano" nel set di dati di test. Il *data poisoning* si riferisce all'introduzione intenzionale e dannosa di dati nocivi nel sistema. Dato che l'IA basata su ML e DL si affida all'essere "addestrata" con un set di dati, un'altra preoccupazione riguarda la validità esterna, la scalabilità e la trasferibilità di un algoritmo non polarizzato. Se il set di dati non è rappresentativo della popolazione più ampia o è intrinsecamente distorto, ciò può ridurre l'applicazione nel mondo reale di un algoritmo (es. uno strumento prognostico di IA addestrato su dati derivati esclusivamente da una popolazione caucasica ma testato su una popolazione più ampia) con conseguenti predittori imprecisi per i gruppi emarginati.
La fragilità dell'IA, definita anche come *"brittleiteness"* (fragilità) dell'IA, è la tendenza del modello a produrre risposte aberranti quando incontra dati con lievi deviazioni rispetto ai dati di addestramento in originali. Pertanto, l'IA si comporta bene con dati familiari ma si degrada e produce informazioni imprecise quando viene stimolata dal rumore. Ad esempio, gli LLM possono produrre informazioni contraddittorie o disinformazione quando interagiscono con formulazioni idiosincratiche che differiscono dal loro set di addestramento.

Considerazioni etiche

Informazioni disaccoppiate dalla moralità

E.O. Wilson ha affermato che "il problema dell'umanità è che abbiamo cervelli paleolitici, istituzioni medievali e una tecnologia divina del XXI secolo". Data questa perspicace osservazione, dobbiamo essere intenzionalmente vigili nel nostro approccio all'adozione dell'IA e consapevoli del suo impatto sull'umanità. Per quanto l'IA sia sofisticata, una macchina rimane pur sempre una macchina e manca di una comprensione semantica degli aspetti che definiscono l'umanità rispetto alla morale, alla compassione, al conforto e alla cura. Pertanto, le applicazioni dell'IA in medicina sono e devono rimanere strumenti che devono essere utilizzati da un essere umano con formazione professionale e umanità, anziché agire in modo autonomo rispetto alla supervisione umana.

Propagazione di bias incorporati e intrinseci

In una società in continua evoluzione con crescenti disparità e disuguaglianze sociali, i medici hanno enormi obblighi etici nel fornire cure mediche sicure, ingegnose ed eque. Come evidenziato sopra, i sistemi di IA comportano un rischio significativo di introdurre e amplificare i bias preesistenti intrinseci nei dati utilizzati per addestrare i modelli. Ad esempio, se per addestrare un modello si utilizzano dati di scarsa qualità e non rappresentativi delle popolazioni del mondo reale, le distorsioni prestazionali possono danneggiare ulteriormente gruppi già emarginati e sottorappresentati [64]. Questa è una sfida in quanto tali bias sono spesso strutturali e incorporati a molteplici livelli sociali e politici. Pertanto, i medici devono essere consapevoli di tali possibili distorsioni quando interpretano o utilizzano strumenti di IA.

Violazione della privacy

La riservatezza del paziente e la gestione sicura dei dati sono integranti in qualsiasi sistema sanitario. Una preoccupazione importante con qualsiasi tecnologia che si basa sulla memorizzazione e sull'accesso a vaste quantità di dati è il rischio di violazioni della sicurezza dei dati e della privacy. L'IA è particolarmente a rischio di attacchi informatici, *data leakage* e *poisoning*. Una sicurezza tecnica di livello superiore non è sempre accessibile o comprensibile, il che rende i pazienti e gli operatori vulnerabili a enormi quantità di potenziali danni. I dati mostrano una crescente sfiducia nella gestione dei dati e preoccupazioni circa la privacy e la cybersicurezza tra i pazienti, pertanto è fondamentale garantire che i sistemi sanitari che utilizzano tecnologie di IA dispongano di misure di sicurezza a prova di errore [65].

L'impatto sull'occupazione

Il possibile impatto dell'IA sulla struttura e sul funzionamento della società è monumentale. Essendo l'IA più efficiente e accurata degli esseri umani, molte industrie vengono ridisegnate e i posti di lavoro vengono ridefiniti. L'IA sta portando sia alla creazione di nuovi posti di lavoro sia alla perdita di quelli attuali. Con la crescita dell'IA e la trasformazione delle industrie, anche i profili lavorativi cambieranno in modo significativo, richiedendo agli esseri umani di riqualificarsi e incorporare l'IA nel proprio bagaglio di competenze.
Dal punto di vista etico, ciò amplia la portata delle disparità, poiché non tutti hanno i mezzi e l'accesso per apprendere le tecnologie di IA per rimanere competitivi. Solleva anche la questione di chi sia la responsabilità di fornire formazione sulle tecnologie di IA emergenti. Nell'ambito della PMA, ad esempio, molti modelli di IA possono eseguire la selezione degli embrioni a un livello simile o superiore rispetto agli embriologi. Ciò potrebbe ridisegnare il ruolo degli embriologi, per cui parte del loro attuale lavoro potrebbe essere automatizzato dall'IA [66].

L'impatto sull'ambiente

L'IA ha un impatto anche sul nostro pianeta, in quanto i calcoli ad alta intensità energetica lasciano una significativa impronta di carbonio (*carbon footprint*). Basti pensare che sono stati utilizzati 5,4 milioni di litri di acqua per addestrare i modelli GPT-3 nei data center Microsoft degli Stati Uniti [67]. Pertanto, è fondamentale considerare le fonti di energia rinnovabile che alimenteranno i sistemi di IA e quanto ciò sia sostenibile per il nostro pianeta. Vi è una crescente preoccupazione circa l'impatto dei data center di IA sui rischi per la salute pubblica e sull'ambiente. I generatori in loco rilasciano inquinanti atmosferici nocivi, come gli ossidi di azoto che sono collegati a malattie respiratorie come l'asma e il cancro al polmone [68].
Data la natura h24 e 7 giorni su 7 dei centri di IA, il diesel viene attualmente utilizzato per alimentare i generatori di riserva, aggiungendo ulteriore carico all'ambiente. Si prevede che i data center di IA utilizzeranno fino al 12% del consumo di elettricità degli Stati Uniti entro il 2028, rispetto al 4,4% del 2023, il che dimostra il rapido aumento dell'impronta di carbonio associata all'IA. L'uso di energia rinnovabile per alimentare tali data center è vitale e dobbiamo rimanere consapevoli dell'impatto ambientale derivante dall'incorporazione dell'IA su scala più

ampia nella vita di tutti i giorni [69]. Alternative ad altre fonti di energia non rinnovabili emergenti per i data center includono piccoli reattori modulari (SMR) che forniscono energia nucleare priva di emissioni di carbonio.

La prospettiva del paziente

Man mano che l'IA diventa più integrata in tutti gli aspetti della vita e della sanità, evolveranno anche le esperienze e le aspettative delle persone nei confronti dell'IA. Un aspetto etico chiave da considerare è il modo in cui i pazienti che ricorrono alla PMA percepiscono il coinvolgimento dell'IA nelle loro cure. Un sondaggio condotto su 200 pazienti sottoposti a FIVET ha riportato una minore fiducia nelle cure riproduttive informate dall'IA e ha dimostrato una preferenza per le raccomandazioni basate sul medico nelle decisioni relative al trattamento; tuttavia, un numero maggiore di pazienti ha favorito le raccomandazioni generate dall'IA per la selezione dei gameti e degli embrioni. I pazienti non erano inoltre disposti a pagare di più per cure di FIVET informate dall'IA [70]. È importante notare che questo studio era soggetto a bias di selezione, in quanto la popolazione di pazienti presentava livelli elevati di alfabetizzazione sanitaria e tecnologica, e questo potrebbe non rappresentare le opinioni della popolazione generale. Dal punto di vista etico, è importante considerare come i pazienti percepiscono il coinvolgimento dell'IA nelle loro cure e come questo influenzi il loro coinvolgimento e la loro fiducia nei trattamenti e nei medici.

Linee guida etiche dell'OMS sull'IA nella sanità

In considerazione delle preoccupazioni di cui sopra, l'Organizzazione Mondiale della Sanità (OMS) ha sviluppato principi etici per guidare l'uso dell'IA nella sanità. I principi etici chiave includono [71]:

- **Tutelare l'autonomia umana:** Con l'IA in grado di competere con il pensiero cognitivo umano e persino di superarlo, è vitale che gli esseri umani mantengano il controllo del processo decisionale medico e che i pazienti prendano decisioni autonome con trasparenza nell'uso dell'IA. Ciò si estende oltre le interazioni medico-paziente nel regno della ricerca medica, in cui il pensiero intellettuale umano e l'analisi rimangono vitali nella valutazione e nell'uso dell'IA. Il design dell'IA che richiede specificamente l'intervento umano nel ciclo (*human-in-the-loop,* HITL) dà priorità allo sforzo di integrare una specifica IA con il giudizio umano e la supervisione per l'accuratezza e il processo decisionale etico.
- **Promuovere il benessere e la sicurezza umana:** Al centro stesso della comunità medica risiede la protezione del benessere della nostra popolazione e il mantra del "non nuocere". L'IA non dovrebbe danneggiare le persone e dovrebbe seguire gli standard di sicurezza, accuratezza ed efficacia applicabili a qualsiasi altro aspetto della società.
- **Trasparenza, spiegabilità e intelligibilità:** L'IA dovrebbe essere comprensibile (in misura diversa) da qualsiasi essere umano che sia uno *stakeholder* (es. sviluppatori, fornitori, pazienti, regolatori). Il livello di comprensione dipenderà dal contesto; tuttavia, l'IA dovrebbe essere comprensibile in base alla capacità di coloro ai quali viene spiegata. Inoltre, uno strumento di IA dovrebbe essere tenuto a identificarsi come tale in qualsiasi interazione con parti sia IA sia non-IA (ovvero umane), in modo che tutte le parti siano consapevoli della fonte dell'interazione in ogni momento.
- **Responsabilità e rendicontazione *(Responsibility and Accountability)*:** Le tecnologie di IA dovrebbero essere valutate dagli *stakeholder*, inclusi pazienti e medici.
- **Inclusività ed equità:** L'IA dovrebbe essere progettata per "incoraggiare l'uso e l'accesso più ampi possibili, appropriati ed equi, indipendentemente da età, sesso, genere, reddito, etnia, gruppo etnico, orientamento sessuale, abilità o altre caratteristiche protette dai codici dei diritti umani". L'IA non deve attivamente perpetuare bias verso gruppi già minoritari.
- **Reattività e sostenibilità:** L'IA dovrebbe essere costantemente rivalutata durante l'uso da parte di utenti e sviluppatori. Le organizzazioni di governo rimangono responsabili di garantire che le conseguenze ambientali dell'incorporazione dell'IA siano minime e che le interruzioni sul posto di lavoro, comprese le perdite di posti di lavoro, la formazione degli operatori sanitari e la ristrutturazione del posto di lavoro, siano navigate in modo appropriato.

Un framework per la valutazione delle tecnologie di IA nella PMA

Come evidenziato sopra, gli strumenti di IA hanno un potenziale enorme ma comportano anche la necessità di cautela e vigilanza man mano che vengono integrati nella pratica clinica. È importante che i clinici e i ricercatori abbiano la capacità di interpretare e valutare criticamente la ricerca sulla PMA che coinvolge l'IA. Questa sezione mira a fornire ai ricercatori e ai clinici di FIVET un quadro di riferimento su come approcciare la lettura e l'interpretazione di manoscritti che utilizzano sistemi di IA nel contesto della PMA.

Esistono numerosi strumenti di valutazione (es. *APPRAISE-AI*) in grado di valutare la metodologia e la qualità di rendicontazione degli studi sull'IA; una recente revisione sistematica ha identificato almeno 26 studi e 9 checklist di rendicontazione [72, 73], tuttavia, a nostra conoscenza, non ce ne sono nel contesto specifico della PMA. Il primo passo quando si valutano gli strumenti di IA nella PMA è classificare cosa fa il sistema di IA. Nella PMA, essi sono generalmente modelli di previsione, strumenti diagnostici o forniscono un supporto alle decisioni cliniche (come descritto sopra). È importante assicurarsi che il disegno dello studio e i metodi di sviluppo dello strumento di IA siano robusti. Ad esempio, le metodologie di IA devono essere convalidate esternamente e il set di dati dovrebbe essere ampio, diversificato, trasparente e tracciabile. Se per l'addestramento viene utilizzato un piccolo set di dati accuratamente selezionato (*cherry-picked*), ciò amplifica il rischio di bias e risultati inaffidabili. È inoltre importante considerare se lo strumento di IA sia utile e venga confrontato con un parametro equo. Ad esempio, la valutazione degli strumenti di IA per la selezione degli embrioni dovrebbe essere confrontata con il gold standard (valutazioni multiple di embriologi in cieco) con metodologie robuste e non utilizzata come parametro autonomo. Ciò rischia di portare a "rivendicazioni eccessive" (overclaiming) o a una generalizzazione eccessiva, in cui il modello di IA può pretendere di "migliorare gli esiti" (es. migliorare i tassi di nati vivi) ma senza effettivi comparatori clinici per valutare il vero impatto del modello di IA. La Tabella 2 evidenzia 8 domini chiave da considerare quando si legge un manoscritto che descrive le tecnologie di IA nella PMA [7].

Tabella 2: Un framework per valutare criticamente la ricerca sull'IA nella PMA (ART)

Dominio / Elemento di valutazione	Considerazioni e domande chiave
1. Quesito clinico e scopo del sistema di IA	Considerare i seguenti aspetti nel formato 'PICO' (Popolazione/Problema, Intervento, Confronto e Outcome): • Qual è la popolazione? È diversa da quella a cui si pianifica di applicare il modello? • Qual è l'intervento/controllo? • Gli outcome (risultati) sono significativi e rilevanti per la pratica clinica quotidiana? • Per cosa viene utilizzato il sistema di IA? (predizione, classificazione, associazione, regressione, ottimizzazione)
2. Qualità dei dati e metodologia dell'IA	**Qualità dei dati:** • Centro singolo o centri multipli? • Quanto è rappresentativa la popolazione rispetto al quesito formulato? • Quali erano i criteri di inclusione/esclusione e come sono stati gestiti i dati mancanti? • Come sono stati gestiti e acquisiti i dati? Il sistema di IA è descritto chiaramente, inclusi il software/hardware utilizzati e l'algoritmo alla base? • Le insidie specifiche della PMA (IVF) includono il bias di selezione (es. considerare solo gli embrioni che sono stati trasferiti) e i fattori di confondimento legati al medico/clinica (es. una tendenza verso i protocolli/scelte adottati da quella specifica clinica) (concetto di "rubbish in-rubbish out", "spazzatura dentro, spazzatura fuori"). **Metodologia dell'IA:** • In quale fase del flusso di lavoro clinico è stato utilizzato il sistema di IA? Questo può aumentare il bias di ancoraggio ("anchoring bias") (es. gli utenti erano all'oscuro dei risultati del sistema di IA prima di prendere la decisione, se questo è ciò che si sta valutando?)
3. Addestramento e sviluppo del modello	• Come sono stati suddivisi i dati tra addestramento ("training") e validazione? A seconda del quesito clinico, ciò potrebbe richiedere una suddivisione per paziente, per tempo, ecc. • Come è stata gestita la suddivisione dei dati in una progettazione multicentrica? Un centro è stato escluso per essere usato come test esterno? • Qual è il rischio che gli stessi dati siano presenti sia nelle coorti di addestramento, sia in quelle di validazione e di test? • Chi ha sviluppato il modello? Un team diversificato dovrebbe essere composto da scienziati, clinici, statistici e ingegneri in grado di gestire adeguatamente il modello.

Linee guida per la rendicontazione della ricerca sull'IA

Trasparenza sull'uso dell'IA nella ricerca

Dato l'aumento esponenziale della ricerca sull'IA, sono state emanate numerose linee guida su cosa e come gli autori dovrebbero riferire quando pubblicano ricerche assistite dall'IA. In primo luogo, i ricercatori devono rivelare e dichiarare se l'IA è stata utilizzata per assistere la generazione di ricerche non-IA. La sottomissione standard, la revisione paritaria (*peer-review*) e la pubblicazione della ricerca richiedono la dichiarazione di conflitti di interesse, e questo deve essere esteso all'uso dell'IA con l'obiettivo di migliorare la trasparenza e la fiducia. Ad esempio, la rivista *Frontiers* ha rilasciato un *framework* per una governance responsabile dell'IA che include la trasparenza su quando l'IA viene utilizzata, la formazione e il rafforzamento delle capacità, i guardrail etici, la governance dell'equità e dell'accesso, il *feedback* della comunità e il monitoraggio e *l'audit* dell'IA nella ricerca [74].

Uso appropriato dell'IA per il quesito di ricerca posto

Per la ricerca che utilizza metodologie di IA, sono disponibili numerose *checklist* di rendicontazione per gli autori, ed è fondamentale che venga utilizzata quella corretta. Ad esempio, la *checklist TRIPOD-AI* è destinata agli studi di previsione, *STARD-AI* agli studi di accuratezza diagnostica e *SPIRIT-AI/CONSORT-AI* agli studi controllati randomizzati [75]. Linee guida aggiuntive che possono essere adattate alla metodologia di IA sono in fase di sviluppo, come *CHART* (studi di valutazione dei chatbot), *TRIPOD-LLM* (studi sui modelli linguistici), *CLAIM* (imaging) e *CHEERS-AI* (studi di valutazione economica) [7, 72]. La Tabella 3 elenca le varie checklist di rendicontazione disponibili per la ricerca sull'IA.

Tabella 3: Checklist di rendicontazione disponibili per la ricerca sull'AI in ambito sanitario

NOME DELLA LINEA GUIDA / CHECKLIST	DISEGNO DELLO STUDIO (STUDY DESIGN)
TRIPOD-AI Transparent Reporting of a Multivariable Prediction Model for Individual Prognosis Or Diagnosis-AI (Rendicontazione Trasparente di un Modello di Predizione Multivariata per la Prognosi o Diagnosi Individuale - AI)	Valutazione dei modelli di predizione
STARD-AI Standards for Reporting of Diagnostic Accuracy Studies-AI (Standard per la Rendicontazione degli Studi di Accuratezza Diagnostica - AI)	Studi di accuratezza diagnostica
DECIDE-AI Developmental and Exploratory Clinical Investigations of Decision Support Systems Driven by Artificial Intelligence (Indagini Cliniche di Sviluppo ed Esplorative dei Sistemi di Supporto Decisionale Guidati dall'Intelligenza Artificiale)	Molteplici (es. studi di coorte prospettici, studi non randomizzati)
SPIRIT-AI Standard Protocol Items: Recommendations for Interventional Trials (Elementi di Protocollo Standard: Raccomandazioni per gli Studi Interventistici)	Studi clinici controllati randomizzati
CONSORT-AI Consolidated Standards of Reporting Trials (Standard Consolidati di Rendicontazione degli Studi Clinici)	Studi clinici controllati randomizzati
CLAIM Checklist for Artificial Intelligence in Medical Imaging (Checklist per l'Intelligenza Artificiale nell'Imaging Medico)	Imaging medico (Diagnostica per immagini)
CHART Chatbot Assessment Reporting Tool (Strumento di Rendicontazione per la Valutazione dei Chatbot)	Studi di valutazione dei chatbot
CHEERS-AI Consolidated Health Economic Evaluation Reporting Standards (Standard Consolidati di Rendicontazione della Valutazione Economica Sanitaria)	Studi economici

Governance dell'IA nella sanità e nella PMA

Al momento non esistono agenzie governative dedicate che sovrintendono all'uso dell'IA nella sanità, a eccezione dell'OMS che ha rilasciato considerazioni etiche come sopra riportato. Ogni paese sta adattando il proprio quadro normativo per incorporare l'IA in modi diversi. Ad esempio, nel Regno Unito, l'IA nella sanità rientra nell'ombrello del *Data Protection Act 2018* e dei *UK Medical Device Regulations 2002* [76]. La Medicines and Healthcare products *Regulatory Agency* (MHRA) ha creato il programma *"Software and AI as a Medical Device Change Programme"* come quadro normativo per supervisionare i dispositivi medici basati su IA. L'Unione Europea (UE) ha creato *l'AI Act*, che stabilisce un quadro basato sul rischio per utilizzare l'IA in sicurezza [77].

Negli Stati Uniti, la *Food and Drug Administration* (FDA) ha autorizzato oltre 1000 dispositivi medici che utilizzano l'IA, tuttavia non esiste un organismo normativo formale che sovrintenda a ciò [78]. Secondo l'*"AI/ML-based SaMD (software as a medical device) Action Plan"* della FDA, gli sviluppatori sono ritenuti responsabili delle prestazioni nel mondo reale dei sistemi di IA. Al momento della stesura di questo articolo, all'inizio del 2026, non ci sono proposte di regolamenti federali negli Stati Uniti, nonostante diverse normative statali. In Australia e Nuova Zelanda, la regolamentazione dei SaMD ricade sotto la *Therapeutic Goods Administration* (TGA) [77].

Dal punto di vista della sicurezza dei dati, i dati gestiti dall'IA rientrano nell'ombrello di varie agenzie di governance dei dati come l'HIPAA (*Health Insurance Portability and Accountability Act*) e il GDPR (*General Data Protection Regulation*) rispettivamente negli Stati Uniti e in Europa. Nell'ambito della PMA, due importanti organizzazioni internazionali includono la *European Society of Human Reproduction and Embryology* (ESHRE) e la *American Society for Reproductive Medicine* (ASRM). Al momento della stesura di questo articolo, nessuna delle due organizzazioni ha rilasciato linee guida formali sulla regolamentazione dell'IA nella PMA.

Conclusioni e raccomandazioni

Se usata responsabilmente e con cautela, l'IA ha un potenziale tremendo per rivoluzionare la PMA. Tuttavia, tutti i modelli di IA sono robusti solo quanto la qualità dei dati utilizzati per addestrarli. Se i dati di addestramento non sono rappresentativi, l'output sarà probabilmente distorto e meno accurato a seconda della natura della popolazione a cui viene applicato. I ricercatori hanno l'obbligo etico e sociale di ridurre i *bias* e di essere il più trasparenti possibile sull'uso e sui limiti dei dati utilizzati per creare i modelli di IA di cui si servono. Allo stesso modo, i clinici dovrebbero essere alfabetizzati sul tema dell'IA, essendo dotati delle conoscenze e delle competenze necessarie per valutare criticamente la ricerca sull'IA nella PMA.

Per confermare la qualità e la legittimità della risposta dell'IA a qualsiasi quesito, raccomandiamo di non avere una fiducia cieca nelle informazioni fornite dall'IA e di verificare utilizzando risorse indipendenti. La verifica può avvenire controllando e convalidando i riferimenti forniti e quelli non forniti, e aggiungendo un ulteriore livello di revisione controllando con un'altra fonte indipendente e fidata (es. linee guida di società professionali note e/o *Cochrane Reviews,* esperti in materia con comprovata esperienza). Quando si valuta la ricerca sull'IA nella PMA, si raccomanda di esaminare attentamente la checklist riassunta nella Tabella 2.

Inoltre, le tecnologie di IA emergenti dovrebbero essere sottoposte a rigorose analisi metodologiche e soddisfare standard etici e di governance appropriati. Con un certo grado di regolamentazione e uno sviluppo attento, le tecnologie di IA presentano numerose applicazioni promettenti nella PMA, nell'ottica di migliorare la cura del paziente e gli esiti clinici.

Bibliografia

1. *Akanksha Garg e David B. Seifer*[1,2]

CON LA SPONSORIZZAZIONE NON CONDIZIONANTE DI



www.valutareinsanita.it