



UNIVERSITÀ  
DEGLI STUDI  
DI PADOVA



DIPARTIMENTO  
DI INGEGNERIA  
DELL'INFORMAZIONE

**DIPARTIMENTO DI INGEGNERIA DELL'INFORMAZIONE**

**CORSO DI LAUREA MAGISTRALE IN  
BIOINGEGNERIA**

**“INTELLIGENZA ARTIFICIALE IN ENDOSCOPIA DIGESTIVA:  
PROTOCOLLO DI RICERCA  
PER ANALIZZARE LIMITAZIONI E BENEFICI  
IN UN'AZIENDA OSPEDALIERA UNIVERSITARIA”**

**Relatore: Prof. Andrea Facchinetti**

**Laureanda: Simona De Toma**

**Correlatori: Ing. Francesca Menotto  
Ing. Erica Ceresola  
Dott. Erik Rosa Rizzotto**

**ANNO ACCADEMICO 2022 – 2023**

**Data di laurea 09/10/2023**

# Indice

Introduzione.....	1
1 Intelligenza artificiale in medicina.....	3
1.1 Introduzione all'intelligenza artificiale.....	3
1.2 In cosa consiste l'intelligenza artificiale medica? .....	3
1.3 IA nella diagnosi medica.....	4
1.3.1 Intelligenza artificiale in radiologia.....	4
1.3.2 Intelligenza artificiale in patologia.....	5
1.3.3 Intelligenza artificiale in endoscopia.....	6
1.4 IA nel trattamento medico-chirurgico.....	6
2 Intelligenza artificiale come dispositivo medico.....	10
2.1 Quadro normativo in Europa .....	10
2.2 Artificial Intelligence Act .....	12
2.3 Etica e governance dell'IA medica .....	14
2.3.1 Qualità dei dati .....	15
2.3.2 Pregiudizio algoritmico .....	15
2.3.3 Opacità.....	16
2.3.4 Sicurezza e protezione .....	18
2.3.5 Attribuzione di responsabilità.....	18
2.4 Norma ISO/IEC 24028: spiegabilità e controllabilità.....	20
3 Endoscopia Digestiva: Colonscopia.....	22
3.1 Strumentazione ed esecuzione .....	22
3.2 Indicatori di qualità.....	24
3.2.1 Tasso di intubazione cecale .....	24
3.2.2 Qualità della preparazione intestinale.....	24
3.2.3 Tasso di rilevamento dell'adenoma (ADR) e tempo di retrazione .....	25
3.3 Screening del Cancro del Colon-Retto (CCR).....	26
3.3.1 Programma di Screening del tumore del colon retto della Regione Veneto .....	27
4 Valutazione dell'efficacia di un sistema di IA utilizzato in colonscopia: stesura di un protocollo di ricerca.....	30
4.1 Introduzione .....	30
4.2 Materiali e metodi.....	31
4.2.1 Impostazione dello studio e pazienti inclusi.....	31
4.2.2 Progettazione dello studio .....	32
4.2.3 Obiettivi dello studio .....	32
4.2.4 Procedura endoscopica .....	33

4.2.5	Sistema di intelligenza artificiale .....	35
4.2.6	Analisi statistica e calcolo della dimensione del campione.....	36
4.2.7	Acquisizione delle variabili .....	36
4.3	Risultati.....	37
4.3.1	Studio della popolazione .....	37
4.3.2	Analisi del tasso di rilevamento di adenomi per ciascuna popolazione .....	39
4.3.3	Analisi delle istologie .....	40
4.3.4	Analisi dei tempi.....	41
4.3.5	Analisi descrittiva dei dati dello Screening di II livello .....	42
4.3.6	Analisi descrittiva dei dati del follow up.....	44
4.3.7	Confronto delle proporzioni campionarie (WL / WL+AI) .....	47
4.3.8	Regressioni logistiche.....	48
5	Discussione.....	51
6	Conclusioni.....	55
7	Bibliografia.....	56

## **Introduzione**

L'intelligenza artificiale (IA) è una delle rivoluzioni tecnologiche più straordinarie del ventunesimo secolo e la sua applicazione in campo medico rappresenta una delle frontiere più promettenti nella promozione della salute e nella lotta contro le malattie. La capacità dell'IA di analizzare, interpretare e apprendere da enormi quantità di dati sta apportando trasformazioni nel settore sanitario, e in particolare nell'endoscopia digestiva, un ambito diagnostico e terapeutico di fondamentale importanza per la salute dell'apparato gastrointestinale. In particolare, l'endoscopia digestiva svolge un ruolo importante nello screening e nella diagnosi precoce del cancro del colon-retto (CCR), una delle principali cause di morte legate ai tumori a livello globale. Gli esami di screening del CCR, come la colonscopia, hanno dimostrato di essere efficaci nella riduzione dell'incidenza e della mortalità da questa malattia [1]. Tuttavia, l'interpretazione delle immagini endoscopiche richiede competenza e precisione e non è possibile eliminare del tutto l'errore umano.

Il presente lavoro di tesi si propone di esaminare l'applicazione dell'IA nel contesto medico, con un focus sull'ospedale di Padova, dove è stato implementato un sistema basato su IA per migliorare e supportare il programma di screening del CCR. La tesi è strutturata in quattro capitoli, ognuno dei quali affronta diverse sfaccettature del tema.

Il primo capitolo introduce l'IA in generale, con un approfondimento del suo utilizzo nel campo medico, concentrandosi in particolare sull'immagine diagnostica (radiologia, patologia ed endoscopia digestiva) e sui trattamenti medico-chirurgici.

Il secondo capitolo analizza l'IA come dispositivo medico, esaminando le normative vigenti che la regolamentano. Vengono affrontate anche questioni etiche e di governance legate all'uso dell'IA in medicina, con un'attenzione particolare alla spiegabilità e alla controllabilità, secondo la norma ISO/IEC 24028.

Il terzo capitolo tratta in modo dettagliato la colonscopia, illustrando la strumentazione necessaria per svolgere l'esame e il sistema basato sull'IA, gli indicatori di qualità della colonscopia e l'importanza dei programmi di screening.

Il quarto capitolo è dedicato allo studio sperimentale condotto presso l'Ospedale Sant'Antonio di Padova. In questo contesto, dove dal 2009 è in corso un programma di screening per il CCR, è stato introdotto nel 2021 un sistema basato sull'IA per la rilevazione e la caratterizzazione dei polipi e degli adenomi. Questo studio mira a valutare in che misura lo strumento basato su IA può supportare gli endoscopisti nelle diagnosi.

I due capitoli successivi discutono l'elaborato, sottolineando le limitazioni dello studio e possibili sviluppi futuri.

Il tema centrale della mia tesi è emerso da un connubio tra il mio interesse per l'ambiente ospedaliero e la mia curiosità riguardo a quanto la sinergia tra l'ingegneria, l'informatica e la medicina possano contribuire a ottenere risultati di eccellenza.

# 1 Intelligenza artificiale in medicina

La ricerca sull'applicazione ed uso dell'intelligenza artificiale (IA) nel campo medico sta crescendo molto rapidamente. Come riportato nell'articolo "*Healthcare Remains The Hottest AI Category For Deals*" [2], nel 2016, i progetti di IA nel settore sanitario hanno ricevuto più investimenti rispetto a qualsiasi altro settore dell'economia globale. Tuttavia, insieme all'entusiasmo, sorgono anche dubbi e richieste di cautela riguardo alle aspettative eccessive.

## 1.1 Introduzione all'intelligenza artificiale

Il concetto di IA si riferisce a un campo dell'informatica dedicato alla creazione di *software* in grado di effettuare prestazioni che, agli occhi di un osservatore comune, sembrerebbero essere frutto esclusivo dell'intelligenza umana. Queste abilità includono, ad esempio, la capacità di risolvere problemi attraverso la logica deduttiva e la capacità di trarre conclusioni basate su input sensoriali visivi o uditivi.

Il termine "IA", sebbene oggi ampiamente diffuso, ha avuto origine negli stessi anni della scienza informatica, con la sua prima menzione nel 1956 da parte del professore John McCarthy dell'Università di Stanford [3]. Le prime applicazioni dell'IA nel settore sanitario risalgono agli anni '70, quando furono sviluppati i primi algoritmi computazionali per supportare decisioni cliniche e analizzare dati sanitari informatizzati [4].

Negli ultimi anni, l'IA ha vissuto un rinnovato interesse grazie alla crescente disponibilità di dati digitali e all'evoluzione delle capacità di calcolo delle tecnologie informatiche. Il *Data Mining* (estrazione dati) e il *Machine Learning* (apprendimento automatico), termini sempre più frequenti nella quotidianità, consentono ai computer di estrarre informazioni significative dai *Big Data* e apprendono autonomamente da essi, senza la necessità di una programmazione specifica [5]. Il *Deep Learning* (DL) rappresenta un'evoluzione di queste tecniche, sfruttando reti neurali artificiali ispirate al funzionamento del cervello umano. Il vantaggio del DL è la possibilità di elaborare dati in formati diversi, come video e audio, senza etichette predefinite, apprendendo progressivamente in modo sempre più autonomo [5].

## 1.2 In cosa consiste l'intelligenza artificiale medica?

Con il concetto "medicina basata sulle evidenze" si intende l'approccio alla pratica medica che si basa sull'utilizzo delle migliori evidenze scientifiche disponibili per prendere decisioni cliniche informate e personalizzate. L'IA medica segue un approccio simile, ovvero viene allenata, attraverso il *Machine Learning*, con un'ampia gamma di dati medici in modo da migliorare il processo decisionale clinico e l'assistenza sanitaria in generale [6].

In passato, i metodi statistici affrontavano questa sfida cercando di descrivere i modelli presenti nei dati attraverso equazioni matematiche. Ad esempio, l'analisi statistica tradizionale, come la regressione lineare, cercava di identificare una "linea di migliore adattamento" ai dati. Tuttavia, con l'avvento del *Machine Learning*, l'IA offre approcci diversi. Utilizzando il *Machine Learning*, infatti, vengono utilizzate tecniche che permettono di scoprire relazioni complesse tra i dati, ovvero relazioni che spesso non possono essere facilmente ridotte ad una semplice equazione matematica. Un esempio concreto sono le reti neurali, le quali sono un modello che, come il cervello umano, rappresentano i dati attraverso un vasto numero di "neuroni" interconnessi. Questa configurazione permette ai sistemi di *Machine Learning* di affrontare la risoluzione di problemi complessi in modo simile a una mente umana, valutando attentamente le prove disponibili per giungere a conclusioni ragionate. Tuttavia, a differenza del singolo medico, questi sistemi sono in grado di elaborare contemporaneamente e rapidamente un numero quasi illimitato di dati in ingresso [7].

Tali sistemi hanno la capacità di apprendere da ogni nuovo caso che affrontano e possono prendere in considerazione una quantità di dati decisamente superiore a quella che un medico potrebbe mai vedere nell'arco di molte vite, il tutto in pochi minuti. Questo è il motivo per cui applicazioni basate sull'IA possono fornire, per esempio, classificazioni più accurate di lesioni cutanee sospette rispetto a quelle che avrebbero individuato i dermatologi senza questo ausilio [8]. Altri studi [9] dimostrano che all'IA vengono persino affidate attività in cui gli esperti spesso non raggiungono un accordo, come, ad esempio, l'identificazione della tubercolosi polmonare dalle radiografie del torace.

### **1.3 IA nella diagnosi medica**

L'utilizzo dell'IA per la diagnosi della malattia o per la condizione di un paziente permetterà sempre più una significativa riduzione dei tempi rispetto a quelli necessari per un medico. L'IA analizza dati clinici provenienti da vari tipi di esami, tra cui quelli radiologici come radiografie, TAC e risonanze magnetiche, test patologici, endoscopici, ultrasonografici e biochimici ed è in grado di fornire risultati rapidi, superando il modello medico tradizionale che spesso fa fatica a fornire diagnosi tempestive e accurate, soprattutto per le diagnosi complesse. Data la celerità dell'IA nello svolgere diagnosi e problemi, i medici sono in grado di elaborare piani di trattamento più ponderati e ben fondati in base alle condizioni specifiche del paziente.

#### **1.3.1 Intelligenza artificiale in radiologia**

La radiologia svolge attualmente un ruolo cruciale nel processo diagnostico per una vasta gamma di malattie. La domanda di diagnosi radiologica sta crescendo ad un ritmo elevato ogni

anno ma, allo stesso tempo, la formazione di medici esperti in radiologia è un processo che richiede tempo e non può essere accelerato. Una soluzione per cercare di colmare la discrepanza tra l'offerta e la domanda di medici radiologi è stata l'introduzione dell'IA nell'ambito radiologico [10]. Negli ultimi anni sono emerse numerose applicazioni dell'IA nella radiologia diagnostica, offrendo nuove prospettive per migliorare l'efficienza e la precisione del processo diagnostico. Un esempio è quello discusso nell'articolo "*Application of Artificial Intelligence in Targeting Retinal Diseases*" [11] ovvero un algoritmo sviluppato con un'alta sensibilità per lo screening precoce e che permette una diagnosi rapida e la classificazione delle malattie della retina. Un'altra applicazione interessante è il sistema diagnostico assistito dall'IA (CADx) [12], il quale, una volta allenato con dati clinici della TC, è in grado di classificare i noduli polmonari maligni e benigni.

Uno studio condotto da Rodriguez-Ruiz [13] ha esaminato l'efficacia clinica tra medici radiologi e l'IA; in particolare ha confrontato le prestazioni autonome di un sistema di IA con quelle dei radiologi nella diagnosi del tumore al seno mediante raggi X. Attraverso l'analisi di un numero cospicuo di esami è emerso che le prestazioni del sistema di IA possono raggiungere un livello di accuratezza paragonabile a quello dei radiologi clinici esperti nella diagnosi del tumore al seno.

### **1.3.2 Intelligenza artificiale in patologia**

La patologia svolge un ruolo fondamentale nella diagnosi di tumori e altre lesioni. Grazie ai progressi nelle tecniche di scansione patologica e all'evoluzione del *software* correlato, l'uso della tecnologia di *imaging* a schermo intero è diventato una pratica comune nel campo della patologia. Nonostante ciò, sussistono ancora sfide riguardanti l'analisi automatica e veloce delle immagini patologiche al fine di ottenere diagnosi precise in tempi rapidi.

Dall'articolo "*Pathology Image Analysis Using Segmentation Deep Learning Algorithms*" [14] si evince che l'evoluzione degli algoritmi di IA ha permesso di migliorare i processi di segmentazione delle immagini patologiche, l'identificazione dei tumori e la rilevazione delle metastasi. Tutto ciò ha portato a risultati di lavoro di qualità superiore e tempi di esecuzione più brevi durante le analisi patologiche. Gli studi hanno evidenziato che, in determinate situazioni, l'algoritmo di IA può portare a ottenere risultati migliori rispetto ai medici patologi esperti nella diagnosi basata su immagini patologiche [15]. I successi dell'IA in questo campo hanno fatto sì che venisse utilizzata a supporto in diverse diagnosi come, ad esempio, quella dei tumori epiteliali e polmonari [16] [17]. Gli sviluppi menzionati, seppure rappresentino una minima parte di quelli attuali e futuri, mettono in evidenza la fattibilità dell'applicazione dell'IA campo patologico.



### 1.3.3 Intelligenza artificiale in endoscopia

Anche nel campo dell'endoscopia digestiva l'IA ha permesso di ottenere significativi progressi tecnologici che hanno rivoluzionato il tradizionale approccio e migliorato l'efficienza. Secondo esperti del settore [18], l'attivazione di IA durante le sedute endoscopiche ha dimostrato di essere altamente efficace nella rilevazione di lesioni, polipi coloretali e tumori gastrici ed esofagei.

In uno studio prospettico randomizzato e controllato [19] è stato valutato l'effetto di un modello CADe (*Computer-Aided Detection*), basato sul *deep learning*, sui tassi di rilevamento di polipi e adenomi. Dopo una randomizzazione di 1058 pazienti, dei quali 536 hanno svolto una colonscopia standard e i restanti 522 hanno eseguito una colonscopia dotata di un sistema di rilevamento assistito da computer (CADe), è emerso che nel gruppo CADe i risultati ottenuti hanno evidenziato un aumento sia dell'ADR (*Adenoma Detection Rate*) sia del numero medio di adenomi identificati per paziente rispetto ai risultati ottenuti dal gruppo delle colonscopie standard. Altri studi [20] hanno utilizzato l'IA per analizzare un grande numero di immagini cliniche acquisite tramite endoscopia, sviluppando algoritmi di reti neurali capaci di rilevare automaticamente le lesioni intestinali. I risultati hanno dimostrato che l'associazione tra endoscopia e il nuovo algoritmo basato sull'IA ha migliorato notevolmente la sensibilità diagnostica e la classificazione (Fig.1) delle lesioni intestinali rispetto ai metodi tradizionali. I due casi studio discussi mostrano, ancora una volta, che l'introduzione dell'IA migliora le prestazioni mediche e che, anche in endoscopia, ci sarà un futuro per questa nuova tecnologia.

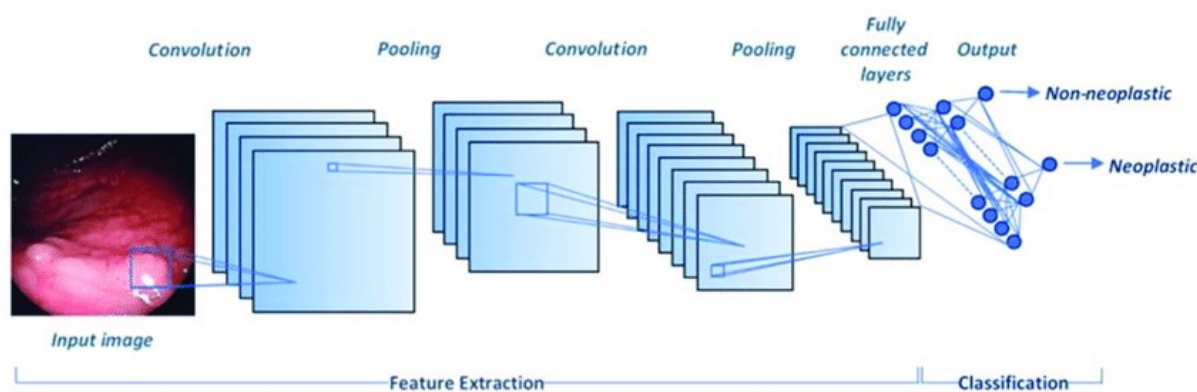


Figura 1: Progetto di rete neurale convoluzionale (CNN) per la classificazione dei polipi del colon-retto. [21]

### 1.4 IA nel trattamento medico-chirurgico

L'avvento del sistema chirurgico Da Vinci [10], dotato di IA, rappresenta un'innovazione importante che ha trasformato l'approccio chirurgico, rendendolo notevolmente meno invasivo e offrendo una serie di vantaggi, tra cui procedure più precise, minimamente invasive ed economiche e persino la possibilità di operare a distanza.

Un'operazione chirurgica è composta da tre fasi distintive: la fase preoperatoria, il momento cruciale dell'intervento e il periodo di convalescenza post-operatoria. L'integrazione delle tecnologie di IA ha portato a notevoli progressi in ciascuna di queste tappe.

Oltre a migliorare le diagnosi, le scelte terapeutiche e l'operatività chirurgica, l'IA potrebbe aiutare i chirurghi nel prendere decisioni riguardo la necessità di determinati interventi chirurgici. Ad esempio, come descritto in un articolo [22], le reti *bayesiane* possono prevedere l'esito della rivascolarizzazione dell'arto durante la valutazione iniziale della ferita. Queste informazioni possono essere utilizzate per sostenere il giudizio clinico, suggerire le decisioni terapeutiche, favorire la condivisione delle decisioni tra il medico e il paziente, stabilire le aspettative di trattamento realistiche e prevedere il rischio di eventuali complicazioni chirurgiche.

Il periodo perioperatorio abbraccia l'intero arco temporale che gravita attorno a un intervento chirurgico, dall'iniziale preparazione del paziente al recupero successivo all'operazione.

Da numerosi anni, i simulatori, noti anche come ambienti di realtà virtuale, hanno fatto parte integrante del campo della chirurgia. Tuttavia, sono i simulatori con sistemi chirurgici completi a destare particolare interesse. A differenza dei simulatori laparoscopici, che sono confinati in box addestrativi, i simulatori robotici sono direttamente collegati alle console robotiche reali, offrendo un'esperienza molto più fedele a ciò che il chirurgo sperimenta effettivamente in sala operatoria. Qualunque sia il tipo di simulatore impiegato, la simulazione rappresenta una forma di addestramento assistito dall'IA. Rispetto alle generazioni precedenti di chirurghi, che si esercitavano su cadaveri o animali vivi, la nuova generazione di professionisti ha accesso potenzialmente illimitato, o almeno notevolmente ampliato, ai simulatori e può ripetere le sessioni di pratica senza costi aggiuntivi [23]. Uno studio di meta-analisi condotto nel 2020 ha confrontato la formazione in realtà virtuale con l'apprendistato tradizionale, evidenziando miglioramenti nell'efficienza dei tirocinanti, una maggiore abilità nella manipolazione dei tessuti e una riduzione degli errori rispetto al metodo di addestramento chirurgico tradizionale [24].

Un altro vantaggio dell'IA nella formazione chirurgica è la stampa 3D, una tecnologia all'avanguardia, che utilizza materiali come polvere di metallo o biomateriali adesivi per creare oggetti strato dopo strato, seguendo modelli digitali generati da dati di *imaging* medico, come tomografie computerizzate o risonanze magnetiche. Questi dati clinici vengono importati in *software* avanzati che, con il supporto dell'operatore nella selezione delle regioni di interesse,

sono in grado di generare una rappresentazione virtuale tridimensionale mediante complesse analisi algoritmiche che verrà successivamente stampata [10].

Un esempio di questo tipo viene discusso nell'articolo "*Utility of 3D printed abdominal aortic aneurysm phantoms: a systematic review*" [25] in cui gli scienziati hanno dato vita a "fantasmi" vascolari stampati in 3D che replicano aneurismi aortici. Queste creazioni hanno assunto un ruolo fondamentale nell'ambito delle simulazioni preoperatorie per la riparazione endovascolare degli aneurismi con stent; questi "fantasmi" non sono soltanto semplici modelli tecnici ma possono essere anche utilizzati per educare i pazienti sulla loro condizione medica in modo straordinario. Attraverso queste rappresentazioni tridimensionali, i pazienti possono ottenere una comprensione più chiara dei rischi e dei benefici legati all'intervento di riparazione endovascolare dell'aneurisma, aiutandoli a prendere decisioni consapevoli sulla propria salute.

Il mondo della stampa 3D continua a rivestire un ruolo centrale anche nella fase intraoperatoria. Durante un intervento chirurgico, emergono costantemente necessità cruciali legate alla fissazione interna<sup>1</sup> e all'assoluta precisione richiesta per l'applicazione di ortesi o l'asportazione di tumori. Queste sfide possono includere la determinazione dell'angolazione e della posizione ottimale per la fissazione al fine di massimizzare l'efficacia, nonché la delicata scelta della sede di taglio per preservare al massimo il tessuto sano circostante. Grazie all'utilizzo dei dati di rilevamento preoperatorio, la stampa 3D è in grado di creare guide chirurgiche personalizzate e modelli che forniscono un supporto inestimabile durante l'intervento [10]. Ad esempio, per quanto concerne l'eliminazione del tumore, la tecnologia di stampa 3D gioca un ruolo fondamentale nell'individuare con precisione e convalidare il punto di taglio durante l'intervento chirurgico per il tumore osseo, contribuendo a ottenere un risultato post-operatorio altamente gratificante [26].

Oltre alla stampa 3D, si aprono nuovi orizzonti nell'ambito dell'efficienza in sala operatoria grazie ad altre tecnologie all'avanguardia. La realtà virtuale, la realtà aumentata e le innovazioni delle realtà miste rappresentano nuovi approcci basati su immagini olografiche, che, in un modo simile alla stampa 3D, abbracciano l'IA per reinterpretare i dati clinici. Queste tecnologie rappresentano un'autentica rivoluzione nell'arte della chirurgia, aprendo scenari inediti per procedure sempre più precise ed efficaci [23] [10].

Nell'ambito della riabilitazione post-operatoria, l'IA svolge un ruolo chiave. Ad esempio, nei reparti di terapia intensiva, l'utilizzo di sensori wireless basati sull'IA permette di raccogliere

---

<sup>1</sup> Nel contesto chirurgico, si riferisce alla necessità di stabilizzare parti del corpo durante un intervento. Questo può avvenire mediante l'uso di dispositivi come viti, piastre o altri strumenti che vengono inseriti all'interno del corpo per mantenere una certa parte anatomica in una posizione specifica o per unire ossa danneggiate.

informazioni sul paziente in modo più efficiente, riducendo i falsi allarmi e semplificando le operazioni [27]. Con lo sviluppo dell'IA sono emersi nuovi strumenti nel settore dell'assistenza infermieristica come il monitoraggio e la gestione remota [28].

I dispositivi medici basati sull'IA, inoltre, dimostrano la loro utilità nella riabilitazione, accelerando il processo di recupero dei pazienti. Esempi concreti sono i robot intelligenti che hanno dimostrato di poter collaborare efficacemente con i pazienti per migliorare il recupero degli arti, ottenendo risultati più che positivi [29].

La tecnologia basata sull'IA viene anche impiegata per tracciare la progressione del recupero e monitorare la salute, il che si rivela prezioso nella gestione dei pazienti una volta dimessi dall'ospedale [30].

## 2 Intelligenza artificiale come dispositivo medico

### 2.1 Quadro normativo in Europa

Il Regolamento sui Dispositivi Medici (MDR) è una legislazione dell'Unione Europea che disciplina i dispositivi medici (DM) e la loro commercializzazione e utilizzo all'interno del mercato europeo. Esso è entrato in vigore il 26 maggio 2021, sostituendo la precedente Direttiva sui Dispositivi Medici (MDD).

Il passaggio dalla Direttiva al Regolamento ha comportato alcuni cambiamenti significativi nel quadro normativo per i dispositivi medici, tra cui una maggiore centralizzazione delle procedure di approvazione e un miglioramento della trasparenza e dell'accessibilità alle informazioni sui dispositivi medici per i pazienti, i professionisti sanitari e le autorità di regolamentazione [31].

Secondo il Regolamento (UE) 2017/745 del Parlamento Europeo e del Consiglio dell'Europa del 5 aprile 2017 relativo ai dispositivi medici (MDR 745) [32], un “dispositivo medico è un qualunque strumento, apparecchio, apparecchiatura, software, impianto, reagente, materiale o altro articolo, destinato dal fabbricante a essere impiegato sull'uomo, da solo o in combinazione, per una o più delle seguenti destinazioni d'uso mediche specifiche:

- diagnosi, prevenzione, monitoraggio, previsione, prognosi, trattamento o attenuazione di malattie,
- diagnosi, monitoraggio, trattamento, attenuazione o compensazione di una lesione o di una disabilità,
- studio, sostituzione o modifica dell'anatomia oppure di un processo o stato fisiologico o patologico,
- fornire informazioni attraverso l'esame in vitro di campioni provenienti dal corpo umano, inclusi sangue e tessuti donati, e che non esercita nel o sul corpo umano l'azione principale cui è destinato mediante mezzi farmacologici, immunologici o metabolici, ma la cui funzione può essere coadiuvata da tali mezzi.”

I software basati sull'IA che guidano o influenzano un dispositivo medico sono regolamentati secondo il MDR.

Nell'Allegato VIII del Regolamento sono definite le classi di rischio per i dispositivi medici, suddivise in quattro categorie principali: I, IIa, IIb e III in base al grado di rischio associato all'uso del dispositivo. L'appartenenza ad una determinata classe, quindi, determina il percorso di conformità del dispositivo, ovvero l'insieme dei controlli necessari affinché il dispositivo stesso dimostri la sicurezza e le prestazioni prima di poter essere immesso sul mercato europeo.

Le principali fasi previste per collocare nel mercato un dispositivo medico, dotato o meno di IA, sono:

1. Qualificare e classificare il prodotto: prima di mettere in commercio un DM, è necessario definire l'uso previsto, la classificazione del rischio e l'identificazione del percorso appropriato verso la conformità.
2. Predisporre la documentazione tecnica: il fabbricante dovrà evidenziare la conformità ai Requisiti Generali di Sicurezza e Prestazione (GSPR) applicabili al dispositivo, i quali verranno dimostrati, attraverso test di laboratorio e valutazioni tecniche, durante lo sviluppo, la produzione, la verifica e la convalida del dispositivo. Inoltre, il fabbricante dovrà includere nella documentazione tecnica i risultati della valutazione clinica. Quest'ultima è una parte fondamentale del percorso verso la conformità in quanto si cerca di dimostrare la sicurezza e le prestazioni del dispositivo attraverso studi clinici e la raccolta dei dati reali sia pre- che post-commercializzazione (PMCF).
3. Istituire e aggiornare un sistema di gestione della qualità: dopo l'immissione del dispositivo sul mercato, è necessario avere un piano di sorveglianza post-marketing, comprendente il follow up clinico post-marketing (PMCF). Quest'ultimo è un processo continuo che aggiorna la valutazione clinica, conferma la sicurezza e le prestazioni per tutta la durata della vita del DM, garantisce l'accettabilità dei rischi identificati ed individua quelli emergenti.
4. Presentare la domanda di certificazione ad un Organismo Notificato: a seconda della classe di rischio del dispositivo e del percorso verso la conformità scelto, il fabbricante deve presentare la domanda di certificazione ad un Organismo notificato, autorizzato ad operare per la tipologia di dispositivi interessata, e procedere con la valutazione della conformità.
5. Emettere la dichiarazione di conformità UE e apporre la marcatura CE sul dispositivo.

Nell'aprile 2021, poco prima della data di applicazione del MDR, la Commissione Europea ha proposto il primo quadro normativo dell'UE per l'IA. Attraverso la *AI Act*, “la Commissione propone di stabilire una definizione tecnologicamente neutra di sistemi di IA nel diritto dell'UE e di stabilire una classificazione per i sistemi di IA con requisiti e obblighi diversi fondati su un *'risk-based approach'*” [33].

L'adozione di questo nuovo atto legislativo dell'UE implementerà norme armonizzate sui sistemi di IA e le integrerà nei quadri legislativi dell'UE esistenti, come l'MDR per i dispositivi medici.

## 2.2 Artificial Intelligence Act

Come accennato precedentemente, nell'aprile 2021 la Commissione Europea ha presentato la proposta del quadro normativo dell'UE sull'IA. Il Parlamento Europeo, il 14 giugno 2023, ha dato il via libera all'*Artificial Intelligence Act* la cui approvazione definitiva dovrebbe arrivare alla fine dell'anno 2023. [34]

Giuridicamente parlando, la legge sull'IA è stata considerata il primo regolamento dell'IA al mondo in grado di regolare lo sviluppo, la commercializzazione e l'uso dei sistemi di IA coerentemente con i valori e i diritti dell'UE.

Tra gli obiettivi della nuova legge c'è la volontà di “*stabilire una definizione tecnologicamente neutra di sistemi di IA nel diritto dell'UE* “. Pertanto, il progetto di legge stabilisce che per un ‘Sistema di intelligenza artificiale’ si intende:

“*...software sviluppato con tecniche e approcci [specifici] [elencati nell'Allegato 1] e in grado, per un dato insieme di obiettivi definiti dall'uomo, di generare output come contenuti, previsioni, raccomandazioni o decisioni che influenzano gli ambienti con cui interagiscono.*” [33]

Con “sistemi di IA”, inoltre, la legge fa riferimento a tecnologie basate su software che comprendono l'apprendimento automatico, i sistemi logici basati sulla conoscenza e gli approcci statistici. Questa definizione estesa include sia i sistemi di IA che possono operare autonomamente sia quelli che possono essere impiegati come parti costituenti di un prodotto.

Le caratteristiche dell'IA quali, ad esempio, l'opacità, la complessità, la dipendenza dei dati e il comportamento autonomo, possono influire negativamente sui diritti e sulla sicurezza degli utenti. Pertanto, un altro obiettivo dell'AI Act è quello di “*stabilire una classificazione per i sistemi di IA con requisiti e obblighi diversi fondati su un approccio basato sul rischio*” [33].

A tal fine, la proposta individua, come mostrato in figura 2, quattro diversi livelli di rischio delle applicazioni dell'IA [35]:

### 1. Rischio inaccettabile

I sistemi di IA che rientrano in questo livello sono quelli che costituiscono una minaccia per le persone; pertanto, sono vietati. In questa categoria rientrano la manipolazione comportamentale cognitiva di persone e/o gruppi vulnerabili, la stratificazione sociale delle persone in base al comportamento, al livello socio-economico e alle caratteristiche personali e i sistemi di identificazione biometrica in tempo reale e a distanza.

## 2. Alto rischio

Verranno identificati come "ad alto rischio" quei sistemi di IA che hanno un impatto negativo sulla sicurezza o sui diritti fondamentali. Il Regolamento, in tale contesto, delinea due distinti insiemi:

1. Sistemi di IA impiegati in prodotti soggetti alla direttiva dell'Unione Europea sulla sicurezza generale dei prodotti: questi comprendono giocattoli, settore dell'aviazione, veicoli automobilistici, dispositivi medici e impianti di ascensori;
2. Sistemi di IA rientranti in settori specifici che saranno necessariamente registrati in un database dell'UE. Alcune aree specifiche sono: istruzione e formazione professionale; forze dell'ordine; gestione delle migrazioni, asilo e controllo delle frontiere.

È necessaria un'analisi continua lungo tutto il ciclo di vita e prima dell'introduzione sul mercato per tutti i sistemi di IA considerati ad alto rischio.

## 3. Rischio limitato

In questo caso i sistemi di IA dovranno soddisfare requisiti di trasparenza fondamentali che permettano agli utenti di prendere decisioni basate su informazioni complete. Gli utenti, quindi, dovranno essere costantemente informati nel momento in cui entrano in contatto con l'IA, come ad esempio le chatbots, così da poter decidere se proseguire con l'interazione o ritirarsi. Questo insieme comprende anche i sistemi di IA in grado di creare o manipolare contenuti visivi, sonori o video.

## 4. Rischio minimo o nullo

Per questo tipo di sistemi di IA non sono imposti vincoli obbligatori e l'utilizzo dell'IA è permesso in modo libero. A questa categoria fanno parte le applicazioni come videogiochi dotati di IA o i filtri antispam.



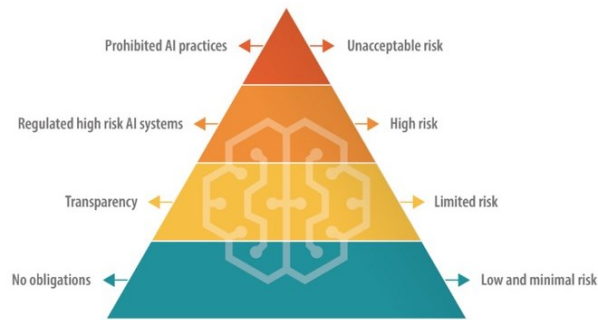


Figura 2: Piramide dei rischi [36]

### 2.3 Etica e governance dell'IA medica

L'espansione e l'accumulo di big data in medicina e salute ha permesso all'IA di addentrarsi ed implementarsi sempre più nell'assistenza sanitaria. Con il progresso dei primi algoritmi basati su regole, seguiti da quelli basati sull'apprendimento automatico e poi su quello profondo, l'IA medica è entrata a far parte di più campi della medicina come la diagnostica delle bioimmagini, lo screening e la previsione delle malattie, il supporto alle decisioni cliniche, la gestione della salute, gli assistenti medici virtuali e la robotica chirurgica.

Il rapido sviluppo della tecnologia dell'IA, però, è accompagnato anche da molti rischi e sfide; gli esperti tendono a considerare gli effetti negativi dell'IA, in particolare quelli legati ai rischi etici. Questi effetti negativi condizionano l'atteggiamento del pubblico e dei professionisti della salute nei confronti dell'IA medica, che risulta essere poco sicura ed affidabile.

Un articolo recente [37] ha analizzato i fattori che influenzano l'affidabilità dell'IA medica sia a livello di progettazione che di applicazione (Fig.3). Secondo quanto riportato in figura 3, il livello di progettazione riguarda l'affidabilità della tecnologia mentre il livello di applicazione riguarda l'impatto sull'uomo dell'utilizzo dell'IA medica.

Nella progettazione dell'IA, la sicurezza e l'affidabilità della tecnologia dipendono in gran parte dai dati e dagli algoritmi: l'aspetto dei dati comprende l'acquisizione, la produzione e l'archiviazione degli stessi; l'aspetto degli algoritmi comprende i difetti e gli errori, le black-box e le vulnerabilità algoritmiche.

Nell'applicazione dell'IA medica, l'impatto sui diritti umani e l'attribuzione di responsabilità influiscono sull'affidabilità dell'IA medica; in particolare, l'impatto sui diritti umani coinvolge l'autonomia umana e la privacy, mentre l'attribuzione di responsabilità riguarda lo status morale dell'IA e chi è responsabile dell'IA.

Sono stati identificate, quindi, cinque tematiche che influenzano l'affidabilità dell'IA medica: qualità dei dati, distorsioni algoritmiche, opacità, sicurezza e attribuzioni di responsabilità.

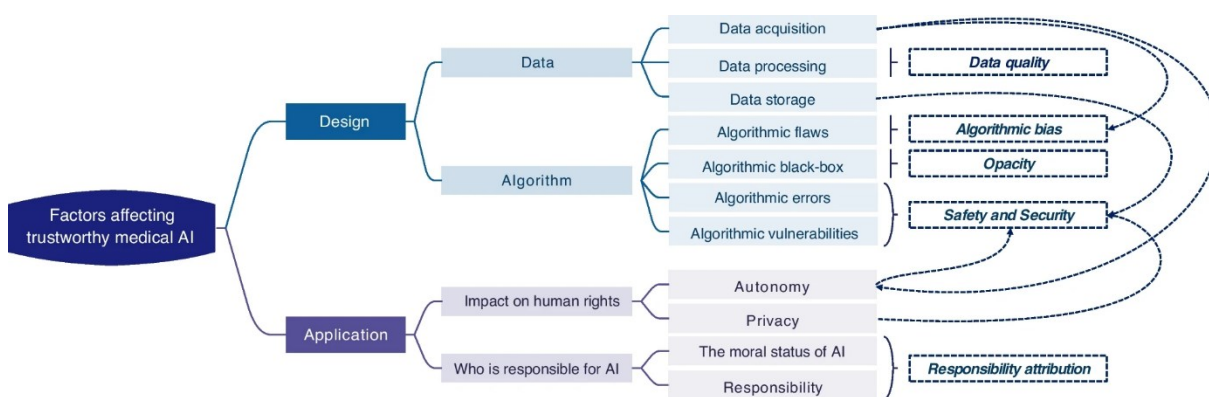


Figura 3: Sistema di governance etica [37]

### 2.3.1 Qualità dei dati

L'espressione "*garbage in, garbage out*" viene utilizzata nel campo dell'informatica per sottolineare il fatto che i computer, e quindi anche l'IA, elaborano in modo acritico anche un insieme di dati in entrata palesemente insensati (*garbage in*) producendo, a loro volta, un risultato insensato (*garbage out*). La qualità dei dati, quindi, determina automaticamente la qualità dell'IA. Nel campo medico, i dati vengono acquisiti da più fonti come, ad esempio, dalla letteratura, dagli studi clinici, dal mondo reale ed anche da un gran numero di persone che indossano apparecchiature intelligenti, come quelle della sfera sportiva o sanitaria. [38]

Di conseguenza, la validità del risultato dell'IA medica è direttamente proporzionale alla correttezza dei dati che vengono considerati. Una scelta approssimativa e superficiale dei dati potrebbe generare gravi danni all'utilizzatore del sistema di IA: è il caso, ad esempio, del sistema Watson di IBM per l'oncologia, il quale era stato addestrato con un piccolo numero di casi di cancro "sintetici" o con dati di pazienti ipotetici piuttosto che reali, generando molteplici esempi di raccomandazioni terapeutiche non sicure ed errate. [39] Garantire dati di alta qualità, quindi, è un prerequisito principale per lo sviluppo dell'IA.

Nel campo dei dati medici, inclusi quelli utilizzati nei dispositivi medici basati sull'IA, spesso c'è una mancanza di standard unificati per i metadati. Questa mancanza può limitare la capacità di combinare dati provenienti da fonti diverse per ottenere una visione completa; ostacolare lo scambio e la condivisione dei dati tra organizzazioni e sistemi diversi; rendere difficile la valutazione sulla provenienza, affidabilità e accuratezza dei dati.

### 2.3.2 Pregiudizio algoritmico

Secondo alcuni studiosi l'IA può ridurre i bias esistenti nel sistema sanitario, come ad esempio gli errori umani e i pregiudizi cognitivi dei medici nel determinare decisioni terapeutiche,

portando a risultati più equi [40]. Tuttavia, gli algoritmi possono portare anche a risultati distorti, dovuti sia a pregiudizi algoritmici indotti dall'uomo sia a quelli indotti dai dati.

I pregiudizi indotti dall'uomo possono essere causati, intenzionalmente o meno, dagli sviluppatori i quali, influenzati dalle proprie percezioni morali e dai propri interessi, influiscono sulla formazione dei dati [41]. Inoltre, gli algoritmi, una volta ereditati i pregiudizi umani, possono rafforzarli ed ampliarli con l'accumulo dei dati e le numerose iterazioni.

Il bias causato dai dati, invece, si ha nel momento in cui i dati di addestramento non sono rappresentativi o sono insufficienti e influenzano il modello dell'algoritmo [42].

Per esempio, un modello prognostico multi-omico che incorpora dati genomici potrebbe non funzionare bene su individui appartenenti a minoranze etniche in quanto la maggior parte degli archivi di riferimento di dati genomici, utilizzati per lo sviluppo del modello, sono costituiti principalmente da individui di origine nordeuropea. Poiché i modelli di IA sono spesso considerati oggettivi è fondamentale divulgare attentamente i dati onde evitare che i pregiudizi incorporati possano essere trascurati. Infatti, è noto che il deep learning è considerato una "scatola nera", opaca e non interpretabile, che rende difficile individuare i pregiudizi. Questi ultimi possono essere replicati continuamente fino a generare risultati di previsione distorta che possono causare un trattamento discriminatorio di alcuni pazienti nelle cure mediche e persino portare ad incidenti di sicurezza medica. Ad esempio, un'IA per la diagnosi del melanoma è difficile da applicare su coloro i quali hanno la pelle nera se nel database delle lesioni cutanee c'è una netta maggioranza di immagini mediche provenienti dalla popolazione bianca e una scarsa minoranza di quelle provenienti dalla popolazione nera [43]

Un altro esempio è quello raccontato in un articolo della rivista "Science" [44] in cui un algoritmo utilizza i costi sanitari (piuttosto che la malattia) per rappresentare il livello dei bisogni sanitari. Poiché viene utilizzato meno denaro per la salute dei pazienti neri che hanno lo stesso livello di bisogno di quelli bianchi, di conseguenza l'algoritmo deduce erroneamente che i pazienti neri sono più sani dei pazienti bianchi.

Da questi esempi ne consegue che sono essenziali l'eterogeneità e rappresentatività dei dati affinché l'IA medica sia generalizzabile e trasferibile e, soprattutto, sia ugualmente applicabile ai pazienti di diverse etnie, aree geografiche, generi e livelli di assistenza.

### **2.3.3 Opacità**

I più recenti modelli di apprendimento automatico hanno un grande difetto che mette in discussione l'affidabilità, specialmente in contesti sensibili come quello medico-sanitario. È il

caso della loro opacità (o il loro essere "black box") che impedisce agli utenti umani di comprenderne e spiegarne appieno il funzionamento.

Le possibili forme di opacità sono [45]: (1) opacità come segreto commerciale delle aziende che tendono a nascondere; (2) opacità inteso come analfabetismo tecnico da parte dei programmatori; (3) opacità che deriva dalla natura complessa degli algoritmi stessi, che risultano incomprensibili per gli esseri umani.

Mentre i primi due tipi di opacità possono essere migliorati istituendo una migliore regolamentazione e formazione, il terzo tipo è un problema intrinseco dell'apprendimento automatico. Tuttavia non tutti gli algoritmi sono scatole nere e il problema dell'opacità si pone soprattutto nell'ultima generazione dell'IA, ovvero l'apprendimento profondo. Quest'ultimo simula il ragionamento umano: infatti è un metodo di calcolo simile al cervello umano che imita la struttura multistrato della rete neurale del cervello umano. È in grado di realizzare l'autoapprendimento e l'estrazione automatica delle caratteristiche senza assistenza umana.

Il deep learning è molto utilizzato nell'IA medica e attualmente offre prestazioni che permettono di migliorare significativamente le capacità diagnostiche e predittive dei medici. Tuttavia, è quasi impossibile spiegare l'IA medica basata sul deep learning sia per l'astrazione matematica delle reti neurali sia per la complessità delle connessioni con altri algoritmi che vanno al di là della comprensione umana. Questa opacità influisce automaticamente sulla fiducia dei pazienti e dei clinici nei confronti degli strumenti dotati di IA. Infatti, la ricerca dimostra che i pazienti sono favorevoli all'utilizzo dell'IA per la gestione di questioni amministrative mentre respingono le applicazioni dell'IA in ambito sanitario [46].

Inoltre, i medici, alle volte, risultano più restii dei pazienti in quanto non saprebbero spiegare il processo decisionale che ha portato ad una determinata diagnosi; tuttavia anche se l'IA medica sapesse dare spiegazioni, queste ultime sarebbero difficili da tradurre in termini medici in quanto si tratta di un'analisi di correlazione statistica piuttosto che di un'analisi di causalità medica [47].

Sebbene il futuro prossimo sarà quello dell'IA spiegabile (XAI), ovvero un'AI comprensibile ma che mantenga allo stesso tempo alte le prestazioni, gli informatici sono concordi nell'affermare l'esistenza di una tensione intrinseca tra le prestazioni e la spiegabilità dell'apprendimento automatico: molto spesso gli algoritmi con rendimento migliore sono i meno trasparenti mentre quelli che forniscono algoritmi esplicativi sono meno accurati [48].

### **2.3.4 Sicurezza e protezione**

La sicurezza nell'ambito dell'IA è un argomento di grande rilevanza e presenta una serie di problemi e sfide. Alcuni principali problemi sono: (1) il controllo umano; (2) la vulnerabilità agli attacchi informatici; (3) la privacy dei dati.

Il controllo umano è fondamentale per l'utilizzo dell'IA in quanto quest'ultima, quando si imbatte in situazioni inimmaginabili, può seguire il proprio istinto ed evolvere in azioni scorrette. Un esempio pratico nell'ambito medico sono le chatbot mediche che, solitamente, forniscono consigli diagnostici e terapeutici per ridurre le visite mediche non necessarie. Tuttavia, se queste non vengono continuamente aggiornate, controllate o regolate possono danneggiare gravemente i pazienti [49]. Un altro esempio può essere un robot di cura che è coinvolto nell'assistenza ai pazienti e agli anziani: se cambia la dose dei farmaci che il paziente deve assumere e l'algoritmo alla base del dispositivo intelligente non viene aggiornato e controllato, il robot stesso può danneggiare la salute del paziente [50].

Gli algoritmi di IA e i sistemi che li ospitano possono essere soggetti ad attacchi informatici. Le vulnerabilità possono compromettere la privacy dei dati, la sicurezza dei pazienti e persino la correttezza dei risultati prodotti dall'IA. Nel seguente articolo [51], infatti, è stato simulato un attacco che potrebbe essere facilmente mercificato utilizzando software moderni. Questa simulazione ha dimostrato come l'aggiunta di un minimo rumore all'immagine originale ha fatto sì che l'IA cambiasse la diagnosi di cancro alla pelle, falsando l'esito. Questo lascia intuire la facilità con cui un hacker malintenzionato potrebbe ingannare un sistema di IA con sottili interferenze danneggiando gli esseri umani.

Il terzo grande problema di sicurezza è garantire la privacy dei dati. Al giorno d'oggi, i dati sono conosciuti come il "nuovo petrolio" per il loro valore economico [52] ma anche di ricerca e business per quanto riguarda i dati sanitari. Proprio per questo si verificano sempre più frequentemente le violazioni ai dati, soprattutto nel settore sanitario che è secondo soltanto a quello finanziario. Molti dati dei pazienti vengono archiviati nel cloud o su server di terzi, rendendoli molto più vulnerabili agli hackeraggi. Pertanto, per la creazione e l'aggiornamento di un sistema di IA che necessita dell'utilizzo dei dati è importante trovare un giusto compromesso tra la condivisione dei dati medici e la protezione della privacy dei pazienti.

### **2.3.5 Attribuzione di responsabilità**

Lo sviluppo di algoritmi sempre più efficienti ha permesso che l'IA medica sia in grado di svolgere alcune attività precedentemente eseguite dai medici. Ciò solleva una serie di domande volte a comprendere l'attribuzione della responsabilità in caso di incidenti medici.

Alla domanda “l'IA può essere oggetto di responsabilità?” si potrebbero identificare due filosofie di pensiero differenti.

La prima ripone molta fiducia nell'IA, tanto da poterla paragonare a quella umana. Il test di Turing<sup>2</sup> suggerisce che le IA complesse possono avere un certo livello di coscienza [53]. Nel "Progetto di relazione con raccomandazioni alla Commissione sulle norme di diritto civile sulla robotica" [54] l'Unione Europea ha conferito ai dispositivi intelligenti lo status di persone elettroniche indipendenti, rendendoli idonei a rispondere dei danni causati da essi stessi. Inoltre, ha introdotto la regola secondo cui "la responsabilità di uno sviluppatore è inversamente proporzionale all'autonomia del robot AI".

La seconda filosofia di pensiero sostiene che, sebbene le IA possano superare gli esseri umani in molti aspetti, non possiedono il libero arbitrio e non hanno una soggettività morale. La visione predominante tra gli studiosi è che, al momento attuale, i robot e l'IA siano considerati prevalentemente strumenti. Nello specifico del contesto medico, l'IA è vista come un ausilio che i medici utilizzano per migliorare le loro capacità diagnostiche e terapeutiche. Tuttavia, l'IA non è considerata in grado di assumere responsabilità come un soggetto autonomo. In altre parole, l'IA è attualmente sottoposta alla guida e alla supervisione umana, e le decisioni e le azioni finali restano di responsabilità umana. Un esempio tangibile è rappresentato dal caso del dottor Watson, il quale dimostra abilità nel rispondere a domande ma è limitato nell'abilità di valutare autonomamente le problematiche. Questa limitazione deriva dalla mancanza di una capacità cognitiva avanzata, necessaria per il pensiero critico e per generare soluzioni diverse e variegata [55].

Pertanto, l'obiettivo primario nel progresso dell'IA è di migliorare il servizio agli esseri umani, evitando che l'IA prevalga sugli individui. L'evoluzione dell'IA non dovrebbe essere eccessivamente orientata verso una "singolarità tecnologica". La tecnologia dovrebbe rimanere

---

<sup>2</sup> Il test si propone di determinare se una macchina è in grado di esibire un comportamento intelligente indistinguibile da quello di un essere umano. Il test di Turing coinvolge tre partecipanti: un essere umano (il giudice), un'altra persona umana (il concorrente umano) e un computer (il concorrente macchina). Il giudice pone delle domande a entrambi i partecipanti, senza sapere quale sia l'essere umano e quale sia il computer. Se il giudice non riesce a distinguere correttamente quale delle risposte provenga dalla macchina e quale dall'essere umano, allora si considera che la macchina ha superato il test di Turing.

un mezzo per servire l'umanità, preservando il valore del talento e della creatività umana, senza sostituirli.

## **2.4 Norma ISO/IEC 24028: spiegabilità e controllabilità**

La spiegabilità e la controllabilità sono due questioni molto importanti che vengono tenute in considerazione per stabilire la fiducia nei sistemi di IA. Di seguito verranno discusse brevemente tali tematiche secondo la Norma ISO/IEC 24028, la quale tratta de “Tecnologia dell’informazione - Intelligenza Artificiale - Panoramica dell’affidabilità nell’intelligenza artificiale” [56].

Come affermato in tale norma, la spiegabilità è una componente importante sebbene non sia sufficiente a garantire la trasparenza in un sistema di IA. La capacità della spiegazione può essere potenziata attraverso la modifica della struttura in modo da adattarla al contesto in cui viene presentata. Questo include la considerazione del pubblico a cui è destinata e del livello di comprensione che si desidera trasmettere. Gli obiettivi dei sistemi di IA spiegabili consistono nel fornire una chiara comprensione dei processi sottostanti che contribuiscono alla veridicità, all'accuratezza e alla coerenza dei risultati ottenuti. Tali spiegazioni fanno sì che il sistema di IA sia più controllabile e contestabile, specialmente per le parti riguardanti le decisioni e le azioni che ne derivano.

Esistono due tipi di spiegazioni, quelle ex-ante ed ex-post, le quali hanno funzioni distinte. Come spiegato nella norma, le spiegazioni ex-ante hanno l’obiettivo di assicurare che il sistema sia ben progettato e che serva allo scopo. Quelle ex-post, invece, spiegano i risultati specifici dell’algoritmo e il circostante in cui sono stati acquisiti. Nell'ottica ideale, un sistema di IA garantirà coerenza tra le sue spiegazioni ex-ante ed ex-post.

È possibile generare spiegazioni nelle fasi dello sviluppo di un modello di IA elencate di seguito [56]:

- Pre-modellazione: in cui si comprendono i dati prima di costruire il modello
- Modellazione: in cui vengono sviluppati modelli di IA in grado di spiegare le decisioni
- Post-modellazione: in cui vengono generate spiegazioni sulle scelte di un modello di IA non interpretabile

Infine, esistono i livelli di spiegabilità che aiutano a selezionare il tipo di IA per un sistema in base al suo potere esplicativo. Ad esempio, i sistemi non interpretabili non sono adatti per essere utilizzati in prodotti che si aspettano un alto livello di spiegabilità.

Quando si parla di “controllabilità” in relazione ai sistemi di IA, invece, si fa riferimento alla capacità di un operatore umano di avere un certo grado di supervisione e controllo sul comportamento e sulle decisioni prese dall’IA. Questo è importante per garantire che il sistema agisca in modo previsto, etico e sicuro. Nella norma viene descritta la “necessità di integrare i punti di controllo nel ciclo di vita del sistema di IA”, in modo da procedere verso la direzione dell’affidabilità del sistema.

I ruoli fondamentali dell’uomo descritti nella norma [56] sono:

- i responsabili delle decisioni che hanno la facoltà e l'autonomia nel prendere decisioni finali;
- gli esperti del settore che hanno la possibilità di fornire un feedback al fine di rivalutare il livello di fiducia del sistema e al contempo contribuire al miglioramento del suo funzionamento.



### **3 Endoscopia Digestiva: Colonscopia**

La colonscopia è un esame diagnostico invasivo che viene eseguito con l'obiettivo di esplorare l'intero intestino crasso e, eventualmente, identificare lesioni sia cancerose che precancerose tramite visualizzazione diretta [57]. Questa procedura non solo rappresenta uno strumento di fondamentale importanza per la diagnosi, ma offre anche l'opportunità di intervenire chirurgicamente sulle lesioni intestinali.

L'alta definizione delle immagini video consente di individuare patologie infiammatorie, formazioni tumorali, ulcere, stenosi e difetti congeniti, permettendo anche interventi diretti su tali condizioni mediante strumenti terapeutici collegati alla sonda ottica, manovrati a distanza dall'operatore medico.

La colonscopia si caratterizza per una sensibilità superiore al 95% nella rilevazione del cancro del colon-retto (CCR), mentre per gli adenomi avanzati (con un diametro  $\geq 10\text{mm}$ ) la sensibilità varia tra l'88% e il 98% [58]. Studi caso-controllo [59] dimostrano una significativa riduzione, compresa tra il 53% e il 72%, nell'incidenza del CRC, accompagnata da una riduzione del 31% della mortalità correlata al CRC, grazie alla colonscopia.

Esistono diverse limitazioni associate alla colonscopia che vanno prese in considerazione. Innanzitutto, la procedura stessa è invasiva e richiede uno o più giorni di dieta priva di scorie e pulizia intestinale spesso poco piacevole e impegnativa in termini di tempo [60]. L'uso di sedazione o anestesia può essere necessario, il che comporta la necessità di un accompagnatore dopo l'intervento, oltre che un giorno di assenza dal lavoro o da altre attività quotidiane. Inoltre, esiste il rischio di perforazione intestinale durante la colonscopia e il rischio di sanguinamento dopo l'intervento, specialmente nei pazienti che hanno subito una polipectomia [61]. Tutte queste limitazioni contribuiscono a una bassa compliance alla campagna di screening del cancro del colon-retto [62]. La qualità dell'esame dipende notevolmente dalla competenza tecnica degli endoscopisti nel rilevare e rimuovere lesioni, soprattutto nel colon prossimale, che possono essere particolarmente difficili da individuare. I polipi piatti o sessili possono rappresentare una sfida aggiuntiva e richiedere l'impiego di tecniche speciali per la rimozione [63]. Un'altra limitazione da considerare è la possibilità dei "cancri intervallo", che sono casi di cancro del colon-retto che si sviluppano nel periodo tra due colonscopie consecutive, contribuendo al 6%-9% di tutti i casi di cancro del colon-retto [64].

#### **3.1 Strumentazione ed esecuzione**

La colonscopia viene condotta attraverso l'uso di uno strumento flessibile noto come colonscopio, una sonda munita di una videocamera ad alta risoluzione e una fonte di luce a fibra

ottica, come mostrato in figura 4. Non si tratta soltanto di una videocamera: il colonscopio può anche essere equipaggiato con diversi strumenti aggiuntivi che trasformano l'endoscopia in una procedura interventistica completa. Questi includono un idropulsore per la pulizia delle pareti intestinali, un aspiratore e, cosa importante, canali di supporto per le pinze per biopsie, pinze operatorie ed elettrocoagulatori. Inoltre il colonscopio è collegato a un terminale video che consente al medico operatore di visualizzare le immagini trasmesse dalla sonda.

Dopo essere stato opportunamente lubrificato con gel anestetico, il colonscopio viene inserito con delicatezza nell'ano del paziente, il quale è posizionato in decubito laterale sinistro, e quindi fatto avanzare attraverso il retto, il sigma, il colon discendente, il colon trasverso e il colon ascendente fino a raggiungere il cieco, permettendo così una mappatura completa dell'intestino crasso. Durante l'esame il colon viene insufflato con aria o CO<sub>2</sub>, per permettere una più accurata visione e successivamente i gas vengono aspirati prima del termine dell'esame. Raggiunto il cieco, il colonscopio viene ritirato con attenzione seguendo lo stesso percorso e successivamente estratto dall'ano. Durante la retrazione dello strumento, nei colon che anatomicamente consentono di farlo, viene effettuata la retroversione al fine di osservare anche aree dietro alle pliche che altrimenti rimarrebbero non visibili.

Il successo dell'esame non è soltanto legato all'abilità dell'endoscopista, ma è altresì attribuibile al personale infermieristico altamente qualificato ed esperto nelle procedure endoscopiche.

Al termine della procedura endoscopica, il colonscopio subisce una fase iniziale di prelavaggio in sala, seguita da una completa pulizia e sterilizzazione in appositi ambienti dedicati.

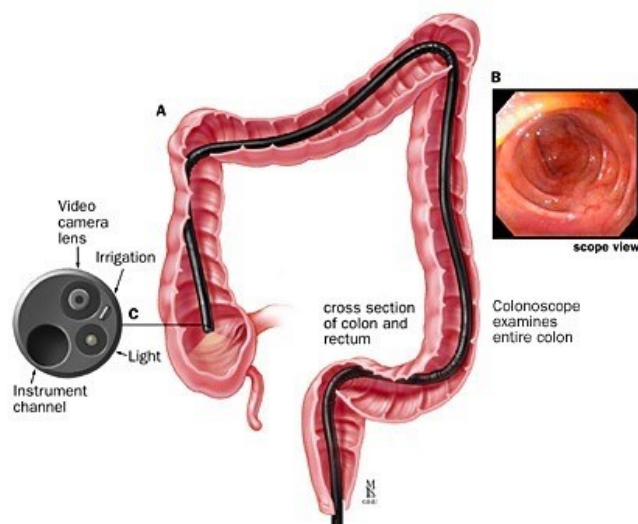


Figura 4: A: colonscopio una volta raggiunta l'intubazione cecale; B: immagine del colon vista dal monitor

## **3.2 Indicatori di qualità**

La colonscopia è considerata il *gold standard* per lo screening del cancro del colon-retto (CRC). Al fine di mitigare le notevoli variazioni di qualità che si riscontrano tra le diverse procedure, sono stati introdotti degli indicatori di qualità. Tra questi, l'indicatore principale è rappresentato dall'ADR (tasso di rilevamento dell'adenoma dell'endoscopista), che misura la capacità dell'endoscopista di individuare adenomi. Vengono, inoltre, presi in considerazione parametri di qualità basati sul processo, come il tempo di retrazione (o tempo di ritiro dell'endoscopio), il tasso di intubazione cecale e la qualità della preparazione intestinale.

### **3.2.1 Tasso di intubazione cecale**

L'intubazione cecale rappresenta il processo di far avanzare la punta del colonscopio fino alla valvola ileocecale in modo da permettere l'ispezione dell'orifizio appendicolare e dell'intera parete mediale del cieco [65]. Un indicatore utilizzato per valutare l'efficacia della colonscopia è il tasso di intubazione cecale (CIR) il quale, se risulta basso, è associato ad un aumento del rischio di cancro del colon retto [66].

Secondo le recenti linee guida, l'intubazione cecale dovrebbe avvenire in almeno il 90% delle colonscopie totali e nel 95% delle colonscopie di screening e, in aggiunta, dovrebbe essere accompagnata da una documentazione fotografica adeguata della valvola ileocecale e, se intubato, dell'ileo terminale [67].

Poiché i livelli di prestazione nell'ottenere un CIR elevato tendono a rimanere stabili o addirittura a migliorare nel tempo, lo stesso CIR è stato messo in discussione in quanto risulta un obiettivo "facile" da raggiungere. Pertanto, l'autore dell'articolo citato [65] solleva un interrogativo riguardo l'efficacia del CIR, ovvero egli sostiene che se la maggior parte dei medici riesce sempre a raggiungere il cieco senza problemi, questo indicatore potrebbe perdere la sua importanza nel tempo, poiché non rappresenta più una sfida significativa. Alcuni suggeriscono che potrebbero essere sufficienti controlli periodici dei tassi di intubazione cecale per documentare prestazioni adeguate, soprattutto quando si tratta di professionisti con un alto volume di colonscopia [68].

### **3.2.2 Qualità della preparazione intestinale**

L'adeguatezza diagnostica della colonscopia è strettamente legata alla qualità della preparazione intestinale. Un'inadeguata pulizia può occultare patologie. Il procedimento di preparazione ideale dovrebbe svuotare completamente il colon da ogni residuo fecale, senza influenzare negativamente l'aspetto della mucosa, richiedere un breve periodo di tempo e

causare il minimo disagio al paziente [69]. Inoltre, dovrebbe massimizzare la rivelazione di patologie coliche, tra cui polipi e carcinomi.

Solitamente, la qualità della pulizia intestinale viene valutata in base alla presenza di feci solide o liquide nel lume dopo la pulizia procedurale. Il miglior indice di valutazione clinica è il *Boston Bowel Preparation Scale* [70], dove un punteggio  $\geq 2$  in ogni segmento del colon è indicativo di una preparazione adeguata. I pazienti con tale punteggio dovrebbero avere un determinato intervallo per il successivo esame di screening o di sorveglianza post-polipectomia, mentre coloro che hanno una preparazione insufficiente dovrebbero ripetere la procedura entro un anno.

Esistono varie pratiche cliniche per migliorare la soddisfazione del paziente legata alla preparazione intestinale. Ad esempio, i pazienti con evacuazioni di colore marrone il giorno della visita dovrebbero essere valutati per ulteriori interventi di pulizia prima della colonscopia, poiché questo colore indica un rischio del 50% di preparazione inadeguata [71]. I pazienti con fattori di rischio noti per una preparazione inadeguata per precedenti interventi chirurgici intestinali, costipazione cronica, uso di farmaci costipanti come gli oppioidi, obesità o diabete, dovrebbero essere valutati per dosi supplementari di preparazione [72].

Purtroppo, numerosi studi [69] [73] indicano che molti centri continuano a registrare tassi di preparazione intestinale inadeguata durante la colonscopia, con una percentuale che si aggira intorno al 20%-25%. Questa situazione ha costi significativi, poiché la necessità di ripetere la procedura aumenta il costo totale della colonscopia di circa l'1% per ogni 1% di esami con preparazione insufficiente [74].

### **3.2.3 Tasso di rilevamento dell'adenoma (ADR) e tempo di retrazione**

Nonostante i diversi indicatori di qualità, l'ADR dell'endoscopista è quello che attualmente definisce la qualità della colonscopia ed è, quindi, considerato il “*gold standard*” per le misure di qualità nello screening della colonscopia [75]. Questo è ulteriormente supportato da uno studio condotto da *Kaminski et al* [76], i quali hanno dimostrato che un aumento dell'ADR durante la colonscopia ha comportato una riduzione del rischio di sviluppare il cancro del colon-retto e di morire a causa di esso. Questo miglioramento è stato osservato in confronto a coloro che non hanno mostrato un aumento dell'ADR durante la colonscopia.

Il valore di riferimento per l'ADR è globalmente del 25%, con una percentuale del 30% negli uomini e del 20% nelle donne [67]. Da un punto di vista statistico, si osserva che, in media, ad ogni incremento dell'1% nell'ADR, è associata una riduzione del 3% del rischio di cancro del colon-retto (CRC) [75]. È importante notare che l'ADR dell'endoscopista è inversamente

correlata al rischio di CCR post-colonscopia del paziente; di conseguenza l'ottimizzazione delle procedure per incrementare notevolmente l'ADR porterà a un miglioramento della qualità delle cure offerte ai pazienti. Un esempio tangibile di ottimizzazione delle procedure per aumentare l'ADR è rappresentato dalla diagnostica assistita da computer (CAD) per la colonscopia. L'integrazione di strumenti basati sull'IA nella pratica clinica offre un notevole potenziale per ridurre l'incidenza futura dei tumori. L'impiego del CAD può significativamente migliorare la capacità di individuare adenomi e altri segni precoci del cancro del colon-retto, contribuendo, di conseguenza, a un trattamento più efficace per i pazienti [77].

Nonostante le evidenze convincenti sulla correlazione tra l'ADR migliorata e la riduzione dell'intervallo CCR, va notato che l'ADR non è esente da limitazioni [75]. Esistono numerose variabili che influenzano la capacità di un endoscopista nell'individuare un polipo adenomatoso, e quindi la sua ADR. Per questo, una variabile correlata all'ADR è il tempo di ritiro dell'endoscopio, il quale può fornire una valutazione supplementare durante la procedura stessa [75]. Il lavoro condotto da *Barclay et al.* [78] ha rivelato una notevole variabilità nei tassi di rilevamento degli adenomi in funzione del tempo di ritiro. Dopo aver suddiviso gli endoscopisti in due gruppi in base al tempo medio di retrazione, che poteva essere superiore o inferiore a 6 minuti, è emerso che il rilevamento degli adenomi era significativamente associato a tempi di retrazione più lunghi. Gli endoscopisti con tempi di retrazione superiori a 6 minuti avevano individuato più del doppio di pazienti con adenomi di dimensioni pari o superiori a 1 cm. Ciò ha dimostrato che aumentare il tempo dedicato all'esame risulta particolarmente rilevante poiché un periodo più lungo per l'osservazione della mucosa porta inevitabilmente alla scoperta di un numero maggiore di polipi che altrimenti potrebbero passare inosservati.

### **3.3 Screening del Cancro del Colon-Retto (CCR)**

Il cancro del colon retto (CCR) rappresenta la terza forma di neoplasia maligna più frequente sia negli uomini che nelle donne. Questo causa un aumento significativo del numero di decessi correlati al cancro in tutto il mondo [79]. È importante notare che circa il 60-70% dei pazienti con CCR manifestano sintomi quando la malattia è già in uno stadio avanzato [1]. Tuttavia, la diagnosi precoce può avere un impatto positivo sui risultati clinici dei pazienti, contribuendo a evitare ritardi nel trattamento e a ridurre la mortalità legata al CCR [1]. La colonscopia riveste un'importanza cruciale nella prevenzione del cancro al colon, soprattutto tra le persone di età superiore ai 50 anni. L'adozione di programmi di screening di routine risulta un passo cruciale nella riduzione dell'incidenza di questa forma di neoplasia [80].

La transizione dalla mucosa intestinale sana alla crescita di tessuti precancerosi e, successivamente, alla formazione di lesioni maligne, è un processo che richiede dai 15 ai 20 anni. La trasformazione dei polipi benigni in tumori maligni si ha gradualmente, talvolta impiegando anche più di 10 anni [81].

Sono stati sviluppati metodi di screening altamente efficaci per individuare tessuti anomali che possono segnalare la presenza di una lesione precancerosa o di un tumore nello stadio iniziale [80]. Tra i metodi utilizzati per individuare la presenza di sangue nelle feci, vi sono il test del sangue occulto fecale al guaiaco (gFOBT) e il test immunochimico fecale altamente sensibile (FIT) [82]. I metodi endoscopici, che fanno uso di approcci ottici per esaminare in modo diretto il retto e il colon, includono la sigmoidoscopia flessibile e la colonscopia [83]. Quest'ultima viene impiegata sia come principale strumento di screening che come procedura di follow-up per individui che hanno ottenuto risultati positivi da screening precedenti. In aggiunta, la colonscopia virtuale (o colon-TAC), una tecnica di imaging basata sulla scansione, è stata sviluppata come un'opzione di visualizzazione meno invasiva per il rilevamento del CCR [84].

Le metodologie più recenti, ancora in fase di consolidamento e non ancora diffuse su larga scala, si fondano su due principali approcci: l'ispezione visiva, come ad esempio l'endoscopia con videocapsula, e l'analisi di biomarcatori presenti nelle feci (come il DNA fecale multitarget), nel sangue (come il DNA Septin9 metilato), o nell'analisi del respiro, che coinvolge la rilevazione di composti organici volatili e una gamma diversificata di marcatori tra cui proteine, RNA e DNA [80].

### **3.3.1 Programma di Screening del tumore del colon retto della Regione Veneto**

L'avvio del Programma di screening nell'ospedale Sant'Antonio risale al 2009 ed ha contribuito progressivamente alla riduzione della mortalità legata al cancro del colon-retto, come mostrato nel grafico. (Fig. 5) Nei primi anni di avvio del programma, si è registrato un aumento delle diagnosi di tumori o lesioni precancerose. Tuttavia, grazie alla rimozione tempestiva di tali lesioni, si è verificato un conseguente declino nelle diagnosi di tumori nei successivi anni.

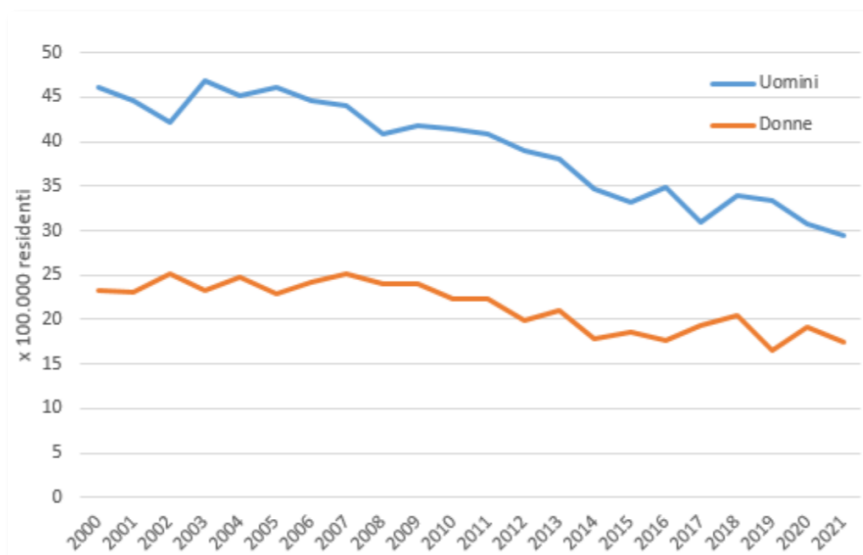


Figura 5: Mortalità per tumore del CR nella Regione del Veneto; Fonte: Registro Tumori del Veneto, SERR - Azienda Zero

Il programma di screening CCR è destinato a tutti i residenti del Veneto di età compresa tra i 50 e i 69 anni, fascia d'età in cui esiste un maggior rischio di sviluppare queste forme tumorali.

Il programma (Fig.6) prevede l'invio di una lettera di invito per sottoporsi al test del sangue occulto nelle feci (screening di I livello) ogni due anni. Questa lettera viene spedita direttamente al domicilio del soggetto dalla ULSS di competenza. Se il risultato del test è negativo, la persona riceverà una comunicazione dalla ULSS e, dopo due anni, verrà invitata a sottoporsi a un nuovo test. Se l'esito è dubbio, la persona sarà invitata a ripetere l'esame subito, seguendo le istruzioni riportate nella lettera di esito. Nel caso in cui il test dia esito positivo, il soggetto riceverà una comunicazione dalla ULSS e un invito a procedere con esami di approfondimento, come la colonscopia (screening di II livello). In seguito, se viene confermata la presenza di una lesione precancerosa o di un tumore, saranno pianificate le analisi e le cure necessarie. Nel caso di un tumore, si procederà con l'asportazione chirurgica, passando così allo screening di III livello.

I PROGRAMMI DI SCREENING DEL CARCINOMA DEL COLON RETTO (FOBT)

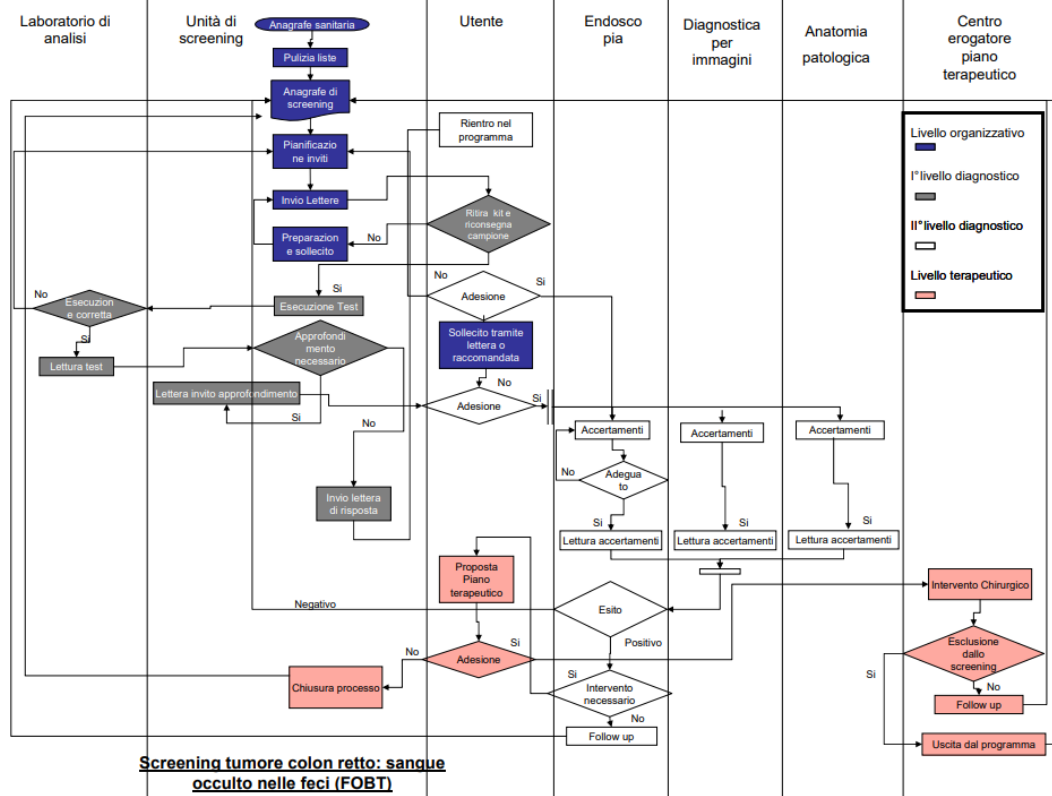


Figura 6: Principali macrofasi del Programmi di Screening CCR; Fonte: ALLEGATO A Dgr n. 4538 del 28/12/2007

Nella figura 7 è possibile osservare l’andamento delle colonscopie di screening eseguite nell’ospedale Sant’Antonio a partire dall’avvio del programma. I numeri mostrano un notevole aumento annuale, indice di una maggiore adesione alla prevenzione del cancro del colon-retto. Il numero totale di colonscopie di screening eseguite fino al 31 dicembre del 2022 è di 28.638<sup>3</sup>.

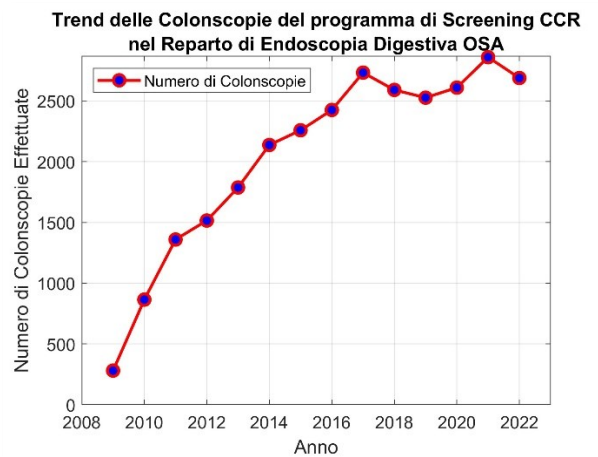


Figura 7: Colonscopie di screening svolte nel periodo luglio 2009 - dicembre 2022, nell'Ospedale Sant'Antonio

<sup>3</sup> Dati forniti dall’amministrazione dell’ospedale Sant’Antonio, Padova.



## 4 Valutazione dell'efficacia di un sistema di IA utilizzato in colonscopia: stesura di un protocollo di ricerca

### 4.1 Introduzione

Il cancro del colon-retto (CCR) è il terzo tipo di tumore più frequentemente diagnosticato in tutto il mondo ed è il secondo in termini di mortalità [85]. Si ritiene, da un po' di anni, che lo screening per il cancro del colon-retto possa avere un impatto sulla riduzione della mortalità legata alla malattia in due modi: uno è la capacità di individuare i tumori in uno stadio iniziale, quando sono ancora trattabili, l'altro è la capacità di individuare ed eliminare gli adenomi, che sono precursori del cancro [86] [87] [88]. È stato dimostrato attraverso studi di screening che sebbene l'individuazione del CCR in uno stadio iniziale sia associata a una diminuzione della mortalità per la malattia stessa, tuttavia, durante gli esami endoscopici di screening è più comune riscontrare polipi adenomatosi. La colonscopia di screening, quindi, riduce l'incidenza e la mortalità del CCR, ma è subordinata alla qualità dell'esame. Quest'ultima ha diversi indicatori, il primo fra tutti è il tasso di rilevamento dell'adenoma (Adenoma Detection Rate, ADR), definito come la percentuale di colonscopie di screening con almeno un adenoma rilevato; un ADR basso è correlato ad un aumento del rischio di "cancro intervallo"<sup>4</sup> [89]. Come sostenuto nell' articolo "*Quality Indicators for Colonoscopy*" [67] l'ADR di riferimento dovrebbe essere del 30% per gli uomini e del 20% per le donne, con un ADR complessivo, quindi, del 25%.

Al giorno d'oggi lo sviluppo tecnologico in medicina è incentrato anche sull'applicazione dell'IA. I sistemi basati sull'IA si sono evoluti in modo significativo grazie allo sviluppo di reti neurali profonde e algoritmi di apprendimento automatico, in particolare nell'ambito della visione artificiale. Le reti neurali convoluzionali (CNN) rappresentano una categoria di potenti reti neurali profonde specializzate nell'elaborazione e nell'analisi di immagini e video.

Dopo i primi successi nell'ambito della radiologia, l'IA è stata introdotta in più settori medicali [8], tra cui quello dell'endoscopia gastrointestinale [90] [91].

La gastroenterologia è una disciplina medica che richiede ai professionisti di avere numerose competenze, che comprendono, oltre a quelle cliniche, la capacità di utilizzo di dispositivi endoscopici, il riconoscimento, la capacità di classificazione della lesione e il processo

---

<sup>4</sup> Carcinoma che compare dopo che un paziente ha ricevuto un risultato negativo da uno screening precedente e prima di sottoporsi al successivo screening programmato. La comparsa di questo tipo di cancro può essere attribuita sia ai limiti del test, che possono portare a risultati falsi negativi, sia all'errore umano che può verificarsi in qualsiasi pratica medica.

decisionale clinico sulla base dell'esame effettuato. Gli sviluppi delle tecnologie dotate di IA, in questo campo, hanno come obiettivo quello di assistere i clinici in queste attività. In particolare, la diagnosi assistita da computer (CAD) per la colonscopia è stata l'area più studiata e quella che ha avuto notevoli successi durante lo sviluppo [18] [19] [20].

Dato che la colonscopia è una procedura operatore dipendente in quanto eseguita da esseri umani, lo scopo dell'introduzione dell'IA durante lo svolgimento degli esami è quello di migliorare la qualità degli stessi per quanto riguarda il rilevamento e la caratterizzazione delle lesioni sospette. In particolare, l'obiettivo è quello di aiutare a prevenire i limiti umani degli endoscopisti dovuti a lesioni difficili da rilevare ad occhio nudo e, inoltre, a fornire una diagnosi ottica precisa per i polipi rilevati. In diversi studi è stato dimostrato che il supporto dell'IA durante le colonscopie ha aumentato il tasso di rilevamento dell'adenoma (ADR) ed ha permesso di ridurre il costo della polipectomia per i polipi iperplastici [92] [93] [94] [95].

Una realtà che utilizza l'IA durante le sedute di colonscopia è l'Ospedale Sant'Antonio di Padova dell'Azienda Ospedale-Università Padova che svolge il programma di screening del CCR. La scelta dell'introduzione di questo dispositivo innovativo, in questo specifico ospedale, è dettata dal fatto che, oltre allo svolgimento di numerose sedute di screening giornaliero, vengono svolte anche sedute di endoscopia avanzata.

Questo studio si pone come obiettivo quello di valutare l'efficacia dell'utilizzo di un sistema dotato di IA durante le sedute dello screening del cancro del colon-retto.

## **4.2 Materiali e metodi**

### **4.2.1 Impostazione dello studio e pazienti inclusi**

Lo studio si è svolto nei mesi di giugno, luglio e settembre 2023, presso il reparto di Endoscopia Digestiva dell'Ospedale Sant'Antonio di Padova, preceduto da un periodo dedicato alla revisione della letteratura per individuare le variabili che concorrono alla valutazione dell'efficacia dell'utilizzo dell'IA in endoscopia, e quindi lo sviluppo dell'indagine volta a raccogliere i dati con i quali verificare e misurare l'obiettivo.

Il 07 giugno 2023 è cominciato lo studio ed è stato incluso il primo paziente. L'analisi è stata condotta sui dati aggregati forniti dall'equipe senza accesso a dati personali e sensibili.

Questo studio esplorativo, randomizzato e controllato mira a comprendere quanto l'IA aiuti il medico endoscopista nelle diagnosi e se ci sia un'ottimizzazione dei tempi con l'utilizzo dell'IA.

Questo lavoro è stato condotto seguendo le linee guida stabilite nella Dichiarazione Consolidated Standards of Reporting Trials – Artificial Intelligence (CONSORT-AI) per la presentazione di ricerche che coinvolgono studi randomizzati e controllati basati sull'IA.

Allo studio hanno partecipato 308 pazienti, residenti nella provincia di Padova, di età compresa tra i 50 e gli 80 anni. In questo range di età sono rientrati pazienti per: accesso spontaneo, screening del CCR di II livello (risultanti positivi all'esame del sangue occulto fecale [SOF+]), follow-up ad 1 anno, follow-up a 3 anni e follow-up a 5 anni.

Sono stati esclusi dallo studio i pazienti che: si recavano a svolgere l'esame da "esterni"; avevano una preparazione intestinale (*Boston Bowel Preparation Scale* [BBPS] [96]) inadeguata (definita come punteggio BBPS <2 in almeno un segmento del colon); avevano stenosi non valicabili o conformazioni del colon particolari (es. dolicosigma); non rientravano nella fascia di età considerata. Infine, non sono stati considerati anche i pazienti che dovevano svolgere soltanto la rettosigmoidoscopia.

#### **4.2.2 Progettazione dello studio**

I pazienti inclusi nello studio sono stati randomizzati in misura paritetica (1:1) in due bracci. La randomizzazione è stata creata utilizzando un programma informatico che ha generato una sequenza di assegnazioni casuali in blocchi di dieci pazienti ciascuno. L'endoscopista non era a conoscenza né delle dimensioni del blocco né della sequenza della lista randomizzata.

Nel braccio di controllo, i pazienti erano sottoposti a colonscopia standard a luce bianca (WL) mentre nel braccio di intervento i pazienti sono stati sottoposti a colonscopia standard a luce bianca con l'ausilio dell'IA (WL+AI) che è stata attivata soltanto nella fase d'uscita dell'endoscopio. Ogni paziente ha quindi avuto il medesimo trattamento come da prassi clinica.

Lo studio è stato condotto a singolo cieco: il paziente non era a conoscenza del trattamento assegnato mentre il medico endoscopista veniva a conoscenza del braccio di randomizzazione poco prima dell'inizio dell'esame.

#### **4.2.3 Obiettivi dello studio**

L'obiettivo primario è stato quello di analizzare e comprendere l'efficacia dell'IA utilizzata durante le colonscopie: in particolare, evidenziare i benefici e le limitazioni dell'IA in un'Azienda Universitaria Ospedaliera che svolge il programma di screening del Cancro del Colon Retto (CCR). Viene quindi verificato se l'utilizzo dell'IA durante l'esame aumenti il tasso di rilevamento degli adenomi (ADR) rispetto alle colonscopie standard. Viene valutata anche l'ottimizzazione dell'IA riguardo i tempi della durata dell'esame e, in generale,

dell'occupazione della sala. Infine, viene valutato se i fattori di rischio come il sesso, l'età del paziente, l'utilizzo dell'IA e il tempo di retrazione influiscono sull'ADR.

#### 4.2.4 Procedura endoscopica

La prima acquisizione è stata l'ora di ingresso del paziente in sala.

Tutti i pazienti erano preparati con lassativi osmotici: come mostrato in figura 8 la maggior parte dei pazienti ha utilizzato Selgesse (69,80%) come farmaco per la preparazione colica; in percentuali minori sono stati utilizzati Clensia (13,06%), Pursennid + Citrafleet (4,08%), Citrafleet (2,86%), Moviprep (2,45%), Plenvu (2,45%), Isocolan (2,04%) ed altre combinazioni di farmaci (3,27%).

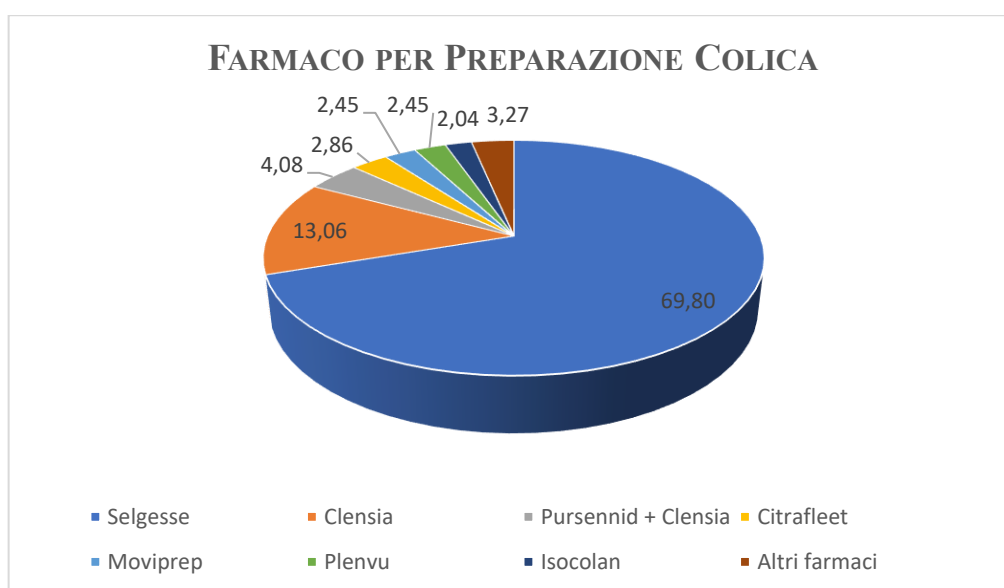


Figura 8: Rappresentazione delle percentuali riguardo i farmaci utilizzati per la preparazione colica

La procedura è stata eseguita con sedazione cosciente dopo che il paziente aveva letto e firmato il consenso informato.

Una volta cominciato l'esame, è stato misurato il periodo di tempo dall'ingresso della sonda nell'ano fino a quando l'endoscopio è avanzato all'ileo (neo) terminale. Dopo aver raggiunto quest'ultimo, è iniziata la misurazione del tempo di ritiro dello strumento ed è stata valutata la preparazione intestinale.

Ai pazienti appartenenti al gruppo WL+AI, il sistema di IA per il rilevamento e la caratterizzazione dei polipi veniva attivato durante il ritiro dello strumento, mentre nel gruppo WL il sistema AI rimaneva disattivato. Inoltre, l'endoscopista, in base alla forma fisica del paziente e al quadro clinico, sceglieva se utilizzare un colonscopio standard di diametro 3,8 cm oppure uno pediatrico di diametro 3,2 cm.

Gli endoscopisti partecipanti allo studio hanno un'esperienza avanzata sia nel seguire i programmi di screening del colon-retto sia nella conoscenza e nell'utilizzo della strumentazione endoscopica. I 6 endoscopisti avevano tutti svolto più di 8000 colonscopie durante la loro carriera.

Ciascun polipo rilevato e rimosso veniva descritto per dimensione (misurata indicativamente con una pinza o un righello), posizione anatomica e forma del polipo secondo la classificazione di Parigi. I clinici hanno utilizzato *Linked Color Imaging (LCI)*, *Blue Laser Imaging* e la modalità di caratterizzazione dell'IA per valutare le lesioni che erano più difficili da identificare con la sola luce bianca. Successivamente le lesioni (fig.9) venivano resecate endoscopicamente utilizzando tecniche di polipectomia standard come: polipectomia con pinza biottica<sup>5</sup>, polipectomia con ansa "a freddo"<sup>6</sup> e mucosectomia endoscopica<sup>7</sup>.

Le lesioni che presentavano già uno stadio avanzato venivano marcate con l'inchiostro di china che facilitava il successivo intervento laparoscopico.

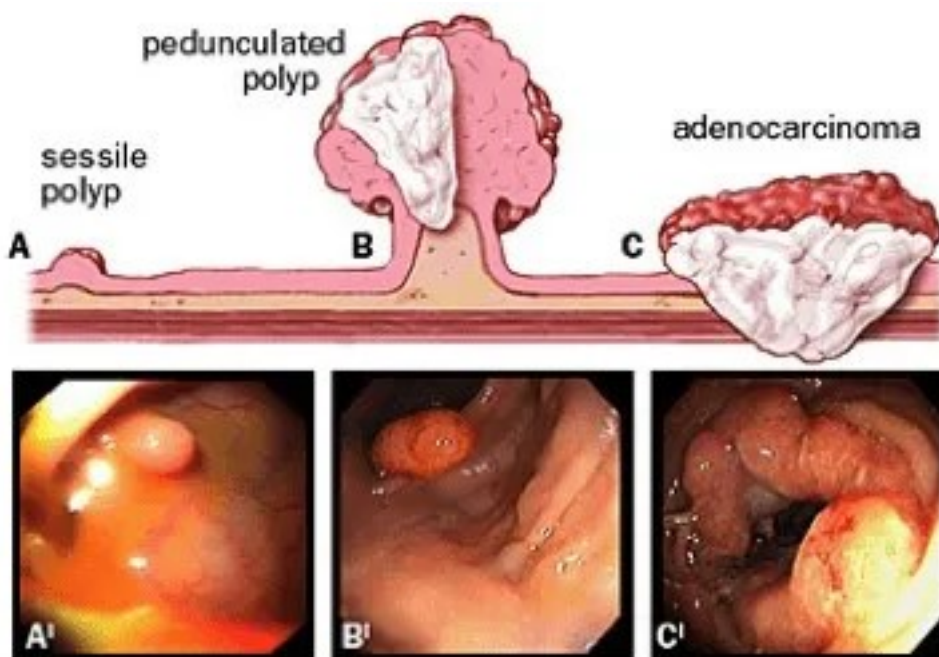


Figura 9: A: polipo sessile; B: polipo peduncolato; C: adenocarcinoma [97]

<sup>5</sup> La pinza utilizzata è la stessa che viene impiegata per le biopsie. Tale procedura è relativamente scevra di rischi nei pazienti con una normale coagulazione ematica.

<sup>6</sup> Tecnica utilizzata per polipi diminutivi; viene utilizzata un'ansa in cui non scorrono né correnti di taglio né di coagulo.

<sup>7</sup> Nota come "lift and cut technique", viene utilizzata per lesioni, superiori al centimetro, sia piatte che peduncolate. Con questa tecnica la base della lesione viene infiltrata mediante un ago utilizzando una soluzione contenente acqua fisiologica, adrenalina e indaco di carminio. Una volta sollevata la lesione si procede con l'asportazione attraverso un'ansa diatermica.

Dopo la resezione, il materiale cellulare veniva inserito nelle provette contenenti la formalina neutra tamponata al 10% che permetteva la fissazione necessaria per mantenere inalterata la morfologia cellulare e l'architettura dei tessuti. Le lesioni, poi, venivano raccolte in un unico contenitore istopatologico ed inviate per la diagnosi istopatologica. I reperti istologici sono stati riportati secondo la classificazione di Vienna.

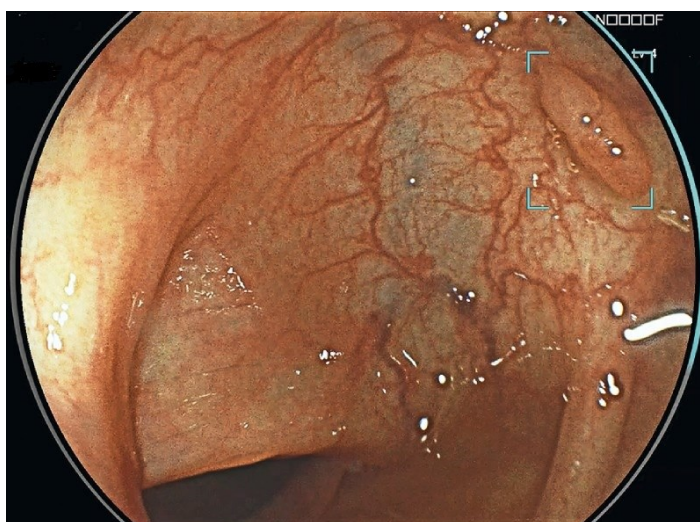
A tutti gli esami è stata eseguita la retroversione in ampolla rettale.

Il tempo complessivo dell'esame è stato calcolato come la somma tra il tempo di salita e quello di ritiro dello strumento, privato del tempo impiegato per le eventuali polipectomie. Non è stato escluso, invece, il tempo impiegato per la pulizia dell'intestino.

#### **4.2.5 Sistema di intelligenza artificiale**

Il sistema di IA utilizzato è basato su una CNN per il rilevamento dei polipi del colon-retto. L'IA è integrata in un box esterno collegato ad un monitor video standard. Questo sistema di IA, a differenza di altri, permette sia di rilevare in tempo reale la presenza di polipi sia di suggerire la caratterizzazione degli stessi fornendo ai medici una previsione istologica, ovvero se si tratta di un polipo neoplastico o iperplastico. Nel momento in cui nell'immagine endoscopica viene rilevato un polipo sospetto appaiono contemporaneamente due segnali: uno di tipo visivo, ovvero attorno al bordo dell'immagine circolare si illumina di azzurro un pezzo di circonferenza per indicare la direzione di visione del polipo sospetto mentre un riquadro azzurro delimita il polipo stesso (fig. 10); l'altro di tipo acustico per attirare l'attenzione dell'endoscopista.

In questo studio non era obbligatorio attivare la caratterizzazione per ogni polipo rilevato tramite l'IA. Tuttavia, sono stati raccolti i risultati tutte le volte che veniva attivata.



*Figura 10: Acquisizione del video endoscopico durante il rilevamento di una lesione attraverso il sistema di IA [98]*

#### **4.2.6 Analisi statistica e calcolo della dimensione del campione**

I confronti dei due gruppi di studio, WL e WL+AI, sono stati eseguiti mediante il test del chi quadro e il t-test, a seconda del tipo di variabile considerata. In tutti i confronti, un *valore*  $P \leq 0,05$  è stato considerato statisticamente significativo. Sono stati eseguiti confronti tra proporzioni campionarie e regressioni logistiche sia univariate che multivariate.

L'analisi statistica è stata eseguita utilizzando Microsoft Excel e MATLAB (Version: 9.14.0.2306882).

La dimensione del campione è stata scelta considerando di avere almeno un centinaio di individui per braccio, per un'analisi consistente. Tenendo conto del numero di esami giornalieri svolti nel reparto di Endoscopia Digestiva, del numero medio di pazienti con una preparazione intestinale inadeguata e della variabilità di esami giornalieri, si è deciso di analizzare circa 300 dati. Inoltre, la dimensione del campione è stata condizionata anche dal tempo destinato allo studio.

#### **4.2.7 Acquisizione delle variabili**

La scelta delle variabili da acquisire per lo svolgimento dello studio è stata preceduta da una ricerca bibliografica. Sono state acquisite le variabili standard quali il sesso e l'anno di nascita del paziente, quest'ultima importante per individuare l'appartenenza alla fascia di età dei pazienti ammessi allo screening del CCR. I dati dei pazienti erano aggregati secondo tre classi di quesito diagnostico: accostamento spontaneo, screening di II livello e follow up. In questo modo è stato possibile confrontare le classi di appartenenza. È stato segnato anche il farmaco che il paziente aveva utilizzato per la preparazione colica in modo da verificare l'efficacia in base ad un'altra variabile importante, ovvero il Boston Scale. Quest'ultima è considerata una variabile discriminante in quanto permette di determinare l'adeguatezza della colonscopia. Un'altra variabile acquisita è stata la sospensione di farmaci effettuata prima dell'esame per comprendere se ci fossero particolari correlazioni con lo svolgimento e l'esito della colonscopia. Per lo svolgimento dello studio è stata generata una lista randomizzata che assegnava un'ulteriore variabile al paziente riguardo la tipologia di colonscopia, ovvero se standard oppure dotata dell'ausilio dell'IA.

Dalla letteratura è emerso che alcuni tempi dello svolgimento dell'esame rappresentavano un indice di qualità: pertanto, in questo studio sono stati acquisiti il tempo di salita, ovvero quello necessario all'endoscopista per raggiungere la valvola ileocecale, il tempo di retrazione, ovvero il tempo di ritiro dell'endoscopio al quale venivano sottratti i minuti necessari per svolgere eventuali polipectomie e/o biopsie. Sono stati rilevati anche il tempo di refertazione e richiesta

dell'esame istologico e la durata dell'occupazione della sala, ovvero da quando il personale allestiva la sala fino alla chiusura completa della cartella clinica.

Infine è stato segnato il numero di asportazioni effettuate, la posizione e l'istologia della lesione. Quest'ultima veniva inserita nel database direttamente dal medico in quanto l'unico a conoscenza del codice identificativo associato a ciascun paziente. In questo modo veniva garantito l'anonimato dei pazienti.

## **4.3 Risultati**

### **4.3.1 Studio della popolazione**

Nei mesi di giugno, luglio e settembre 2023, 308 pazienti, che si sono recati in ospedale per visite di tipo endoscopico, sono stati considerati idonei per lo studio. Dopo l'esclusione di 20 pazienti, le cui ragioni e il flusso dei pazienti sono riportati in figura 11, gli individui sono stati randomizzati 1:1 in due gruppi. C'è stata una divisione equa di individui nei due bracci: 144 nel primo, quello riguardante la colonscopia standard effettuata a luce bianca (WL), e 144 nel secondo, quello riguardante la colonscopia standard effettuata a luce bianca supportata dall'ausilio dell'IA (WL+AI). Altri 43 pazienti (26 del primo braccio (WL) e 17 del secondo (WL+AI)) sono stati esclusi dall'analisi dopo la randomizzazione a causa di una preparazione intestinale inadeguata o dell'incompleta intubazione cecale. Infine, 245 pazienti (118 del gruppo della colonscopia standard (WL) e 127 del gruppo della colonscopia standard dotata di IA (WL+AI)) sono stati inclusi nell'analisi dello studio.



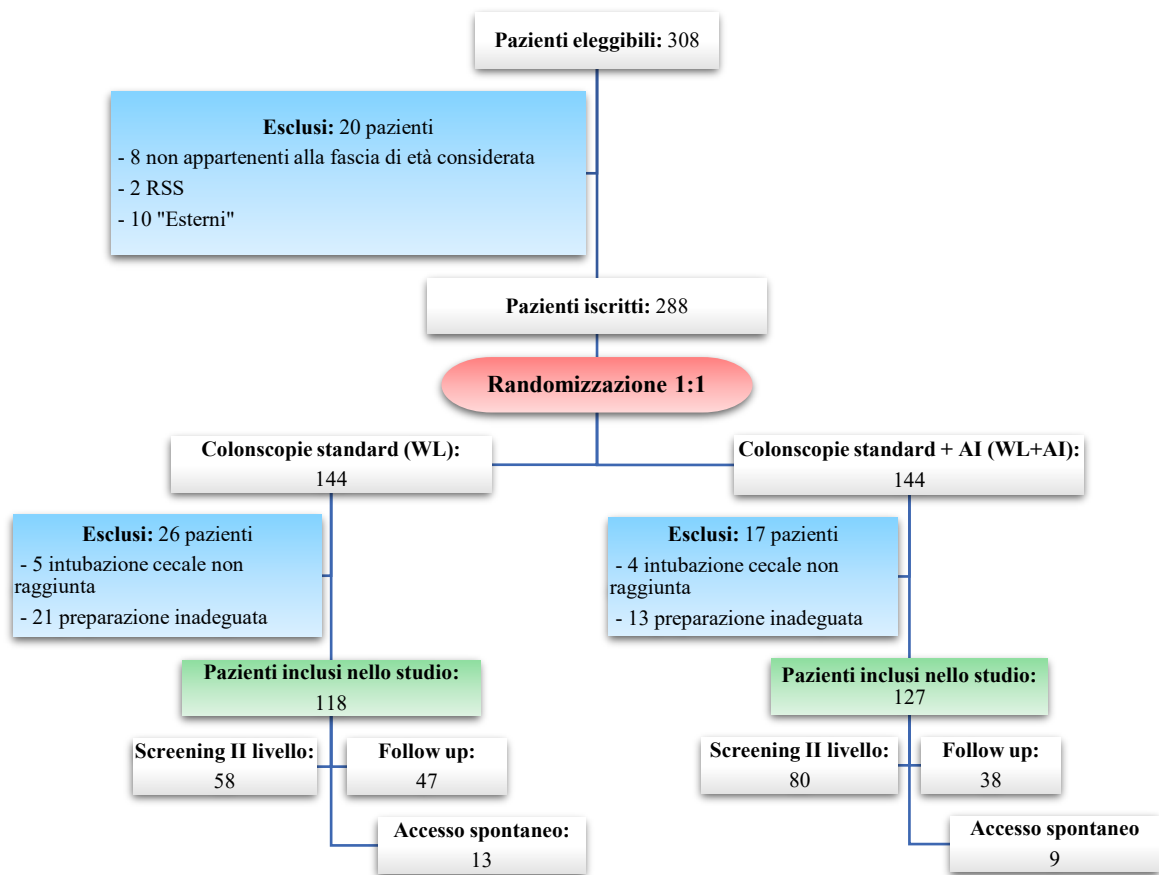


Figura 11: Flusso dei pazienti idonei allo studio. WL, white light; WL+AI, white light + artificial intelligence

Dei pazienti inclusi, il 52,65% era di sesso maschile e il restante 47,35% di sesso femminile; l'età media era di 62,69 anni (IQR 58-68). Come mostrato nella figura 12, per quanto riguarda il quesito diagnostico, il 56,33% dei partecipanti apparteneva allo screening di II livello, il 34,69% appartenevano al follow up e l'8,98% aveva effettuato un accesso spontaneo.

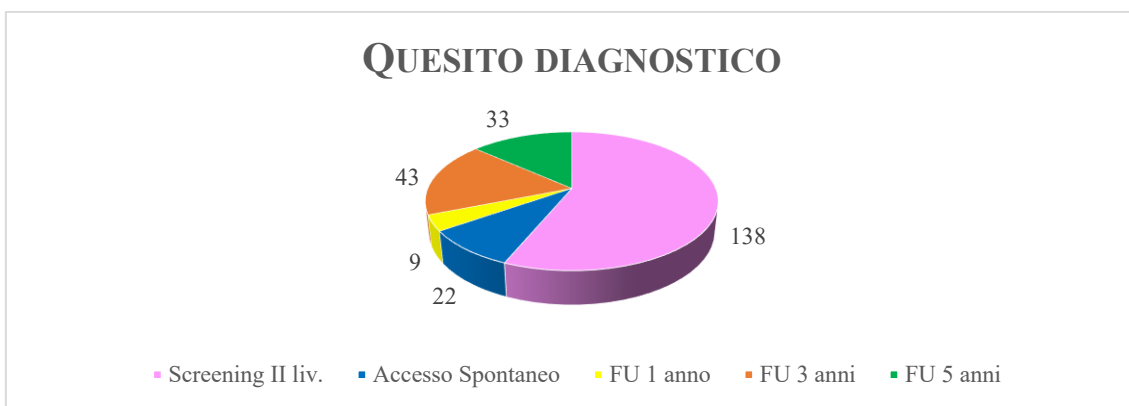


Figura 12: Categorie del quesito diagnostico dei pazienti

I pazienti inclusi nei due bracci erano comparabili riguardo le caratteristiche demografiche, quali sesso ed età, la qualità della preparazione colica, la durata dell'occupazione della sala e il tempo medio sia dell'inserimento che della retrazione dell'endoscopio (Tab. 1).

Tabella 1: Caratteristiche generali

	<b>Colonscopia standard (WL)</b>	<b>Colonscopia standard con AI (WL+AI)</b>	<b>Valore P</b>
<b>Pazienti, n</b>	118	127	-
<b>Età, mediana (IQR), anni</b>	63 (58-67,75)	62,5 (58-68)	0,64
<b>Sesso femminile, n (%)</b>	55 (46,61)	61 (48,03)	0,82
<b>Sesso maschile, n (%)</b>	63 (53,39)	66 (51,97)	
<b>Screening di II livello (SOF+), n (%)</b>	58 (42,03)	80 (57,97)	-
<b>Follow up, n (%)</b>	47 (55,29)	38 (44,71)	-
<b>Accostamento spontaneo, n (%)</b>	13 (59,09)	9 (40,91)	-
<b>Pulizia complessiva del colon, n (%)</b>			
- <b>BBPS 6 – 7</b>	109 (92,37)	114 (89,76)	0,47
- <b>BBPS 8 – 9</b>	9 (7,63)	13 (10,24)	0,47

#### 4.3.2 Analisi del tasso di rilevamento di adenomi per ciascuna popolazione

In 91 pazienti, sul totale di 245, è stato rilevato almeno un adenoma classificato come displasia di basso o alto grado. Il tasso complessivo di rilevamento dell'adenoma è stato del 37,14%. Nel braccio WL, gli adenomi sono stati rilevati in 38/118 pazienti rispetto a 53/127 del braccio WL+AI, corrispondente ad un ADR più elevato (32,2% vs 41,73%, p=0,123) nella colonscopia assistita dall'IA (Tab. 2).

Tabella 2: ADR complessivo dello studio

<b>ADR complessivo dello studio = 37,14%</b>	
<b>Colonscopia standard (WL) (%)</b>	<b>Colonscopia standard con AI (WL+AI) (%)</b>
32,20	41,73

Per quanto riguarda la popolazione dello screening di II livello, sul totale di 138 pazienti, è stato rilevato almeno un adenoma in 55 individui, ottenendo così un tasso di rilevamento dell'adenoma del 39,86%. In particolare, nel braccio WL, gli adenomi sono stati rilevati in 22/58 pazienti rispetto a 33/80 del braccio WL+AI, corrispondente, anche in questo caso, ad un ADR più elevato (37,93% vs 41,25%, p=0,69) nella colonscopia assistita dall'IA (Tab. 3).

Tabella 3: ADR della popolazione di screening di II livello

<b>ADR della popolazione Screening di II livello = 39,86%</b>	
Colonscopia standard (WL) (%)	Colonscopia standard con AI (WL+AI) (%)
37,93	41,25

Riguardo la popolazione del follow up, sul totale di 85 individui, è stato rilevato almeno un adenoma in 36 pazienti, ottenendo così un tasso di rilevamento dell'adenoma del 42,35%. In particolare, nel braccio WL, gli adenomi sono stati rilevati in 17/47 pazienti rispetto a 19/38 del braccio WL+AI, corrispondente, anche in questo caso, ad un ADR più elevato (36,17% vs 50,00%,  $p=0,20$ ) nella colonscopia assistita dall'IA (Tab. 4).

Tabella 4: ADR della popolazione del follow up

<b>ADR della popolazione Follow up = 42,35%</b>	
Colonscopia standard (WL) (%)	Colonscopia standard con AI (WL+AI) (%)
36,17	50,00

L'analisi della popolazione degli accessi spontanei è stata tralasciata poiché di questa fanno parte soltanto 22 individui, ovvero un numero poco consistente.

### 4.3.3 Analisi delle istologie

Durante lo studio sono state effettuate 195 asportazioni (Tab. 5). Gli istologici sono stati raggruppati in quattro classi: adenoma con displasia di basso grado, adenoma con displasia di alto grado, polipi sessili serrati e polipi iperplastici.

Il totale degli adenomi rilevati è 139; di questi 123 sono risultati adenomi con displasia di basso grado e soltanto 16 sono risultati adenomi con displasia di alto grado. Dei 123, 53 appartenevano al braccio WL e 70 a quello WL+AI (53/118 [44,91%] vs 70/127 [55,12%],  $p=0,11$ ); degli adenomi con displasia di alto grado, 4 appartenevano al braccio WL e 12 a quello WL+AI (4/118 [3,39%] vs 12/127 [4,45%],  $p=0,055$ ).

Sebbene non si osservino alcune differenze significative nei due bracci riguardo il tasso di rilevamento degli adenomi, quest'ultimo, sia per gli adenomi con displasia di basso grado sia quelli con displasia di alto grado, è risultato più elevato nel braccio della colonscopia dotata di IA rispetto a quella standard.

I polipi rimossi sono stati, in totale, 43. Di questi 11 sono risultati polipi sessili serrati e 32 iperplastici. Dei polipi sessili serrati, 4 appartenevano al gruppo WL mentre 7 a quello

WL+AI (4/118 [3,39%] vs 7/127 [5,51%], p=0.42). I polipi iperplastici, invece, erano così divisi: 15 appartenevano al braccio WL e 17 a quello WL+AI (15/118 [12,71%] vs 17/127 [13,38%], p=0.87).

Anche in questo caso, sebbene non si osservino alcune differenze significative nei due bracci riguardo il tasso di rilevamento dei polipi, quest'ultimo è risultato maggiore nel braccio che utilizzava l'IA piuttosto che in quello delle colonscopie standard.

È risultato che 13 asportazioni erano campioni di mucosa mentre a 3 pazienti è stato rilevato un adenocarcinoma.

Tabella 5: Lesioni asportate

	<i>Colonscopia standard (WL)</i> (n=118)	<i>Colonscopia standard con AI (WL+AI)</i> (n=127)	<i>Valore P</i>
<i>Asportazioni effettuate (n)</i>	82	113	0.0002
<i>Adenomi rilevati (n)</i>			
- <i>Con displasia di basso grado</i>	53	70	0.1105
- <i>Con displasia di alto grado</i>	4	12	0.0551
<i>Polipi rilevati (n)</i>			
- <i>Sessili serrati</i>	4	7	0.4229
- <i>Iperplastici</i>	15	17	0.8757
<i>Campioni di mucosa (n)</i>	4	6	0.5978
<i>Adenocarcinomi (n)</i>	2	1	0.5187

#### 4.3.4 Analisi dei tempi

Durante lo studio sono stati acquisiti i seguenti tempi: intubazione cecale, retrazione ed occupazione della sala. Il tempo di intubazione cecale medio è risultato di 00:07:15 per il braccio della colonscopia standard e di 00:07:00 per quello che ha utilizzato l'IA. Il tempo di retrazione medio è stato di 00:07:18 per il gruppo WL e di 00:07:51 per il gruppo WL+AI. Riguardo il tempo medio di occupazione della sala, nel primo gruppo (WL) è stato di 34 minuti mentre nel secondo (WL+AI) di 35 minuti (Tab. 6). In tutti e tre i casi, il tempo medio non differiva significativamente tra i due bracci.

Tabella 6: Tempi acquisiti

	<i>Colonscopia standard (WL)</i>	<i>Colonscopia standard con AI (WL+AI)</i>	<i>Valore P</i>
<i>Tempo di intubazione cecale medio</i>	00:07:15	00:07:00	0,62
<i>Tempo di retrazione medio</i>	00:07:18	00:07:51	0,14
<i>Tempo di occupazione sala medio</i>	00:34:00	00:35:00	0,19

#### 4.3.5 Analisi descrittiva dei dati dello Screening di II livello

In questa analisi è stato svolto il confronto tra due proporzioni campionarie, ovvero il sesso biologico maschile e quello femminile, riguardo i pazienti appartenenti al gruppo dello screening di II livello. Nella seguente tabella (Tab. 7) sono riportati il numero degli individui di entrambi i sessi (M e F) a cui è stato o meno rilevato almeno un adenoma (rispettivamente 1 e 0).

Tabella 7: M=sesso biologico maschile; F=sesso biologico femminile; 0=nessun adenoma rilevato; 1=almeno un adenoma rilevato

	<i>0</i>	<i>1</i>	
<i>M</i>	41	29	70
<i>F</i>	42	26	68
	83	55	

Le proporzioni campionarie sono state calcolate nel seguente modo:

$$\text{Proporzione campionaria } M = \frac{M \text{ con almeno un adenoma}}{\text{popolazione totale maschile}} = \frac{29}{70} = 0,4143$$

$$\text{Proporzione campionaria } F = \frac{F \text{ con almeno un adenoma}}{\text{popolazione totale femminile}} = \frac{26}{68} = 0,3824$$

Per confrontare le due proporzioni campionarie e determinare se sono statisticamente significative è stato utilizzato sia il test del chi-quadrato che il t-test. Nella tabella di seguito (Tab. 8) sono riportati i risultati ottenuti:

Tabella 8: Valore p

	<i>Test Chi-quadrato</i>	<i>T- Test</i>
<i>Valore P</i>	0.702	0.77

Entrambi i test sottolineano che la differenza tra le due proporzioni campionarie non è statisticamente significativa.

È stata inoltre svolta una regressione logistica per valutare la relazione tra l'età e la probabilità degli adenomi, utilizzando il modello di regressione lineare generalizzato con una distribuzione binomiale e la funzione *logit*. Di seguito sono riportate le statistiche del modello (Tab. 9) e il grafico ottenuto (fig. 13).

Tabella 9: Statistiche del modello

	<i>Coefficienti stimati (<math>\beta</math>)</i>	<i>Errore standard</i>	<i>t-Statistico</i>	<i>Valore P</i>
<i>Intercetta</i>	-2,95	1,93	-1,53	0,12
<i>Età</i>	0,04	0,03	1,32	0,18

- **Z – Score** = 1.33
- **Odds Ratio** = 1.04

Il valore dello z-score ottenuto, 1.33, è compreso nell'intervallo  $\pm 1.96$ . Di conseguenza l'odds ratio non è statisticamente diverso da zero e, in aggiunta, è possibile rilevare che l'età non è un fattore che può migliorare statisticamente il tasso di rilevamento degli adenomi.

In questo grafico (Fig. 13) ciascun pallino blu rappresenta uno o più pazienti, di età pari al numero indicato sull'asse delle ascisse, al quale è stato o meno rimosso almeno un adenoma, a seconda che il pallino si trovi rispettivamente sull'asse  $y=1$  o  $y=0$ . La regressione logistica ottenuta mostra che la probabilità di rilevare adenomi aumenta all'aumentare dell'età del paziente.

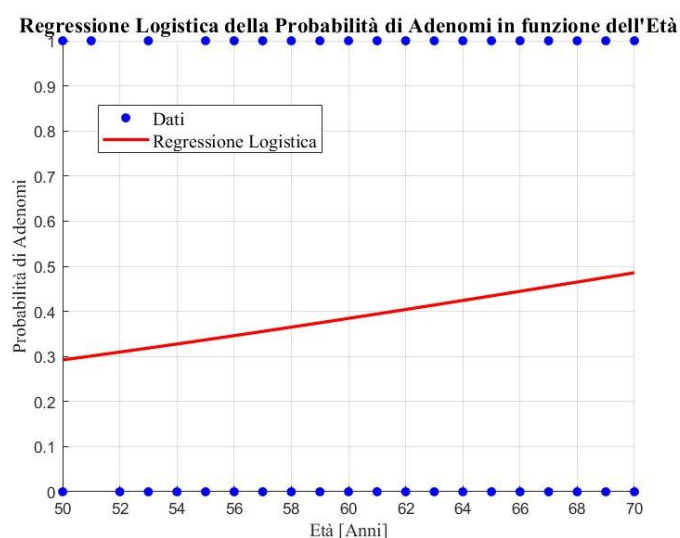


Figura 13: Regressione logistica della probabilità di rilevare adenomi in funzione dell'età

Infine è stata svolta una regressione logistica multivariata della probabilità di rilevare adenomi in funzione dell'età e del sesso del paziente. Il modello di regressione è lo stesso dell'analisi precedente con le statistiche riportate in tabella 10.

Tabella 10: Statistiche del modello

	<i>Coefficienti stimati (<math>\beta</math>)</i>	<i>Errore standard</i>	<i>t-Statistico</i>	<i>Valore P</i>
<i>Intercetta</i>	-2,95	1,93	-1,53	0,13
<i>Età</i>	0,04	0,03	1,36	0,17
<i>Sesso</i>	-0,17	0,35	-0,49	0,62

In questo caso è stata tracciata la curva ROC (fig. 14) e calcolata l'area sotto la curva stessa: AUC=0,5766. La sensibilità e la specificità sono risultate rispettivamente pari a 0,71 e 0,47.

Il grafico sottostante mostra la relazione tra il tasso di falsi positivi e il tasso di veri positivi al variare della soglia di classificazione che in questo caso è stata calcolata utilizzando l'indice di Youden. Dalla Curva ROC e dal AUC=0,5766 si deduce che il modello è in grado di discriminare casualmente la prevalenza dell'adenoma in funzione delle classi considerate, ovvero, in altre parole, questo modello non è diverso da un modello senza covariate.

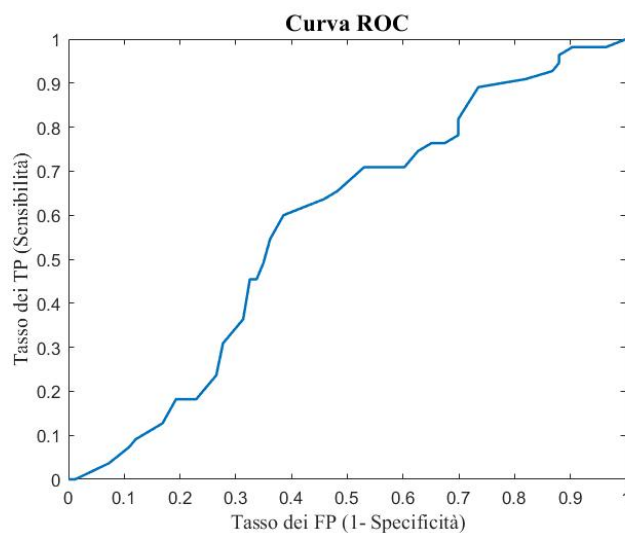


Figura 14: Curva ROC del modello

Dagli z-score riportati in tabella 10 di entrambe le covariate si evince che sia il sesso che l'età sono fattori di rischio che non influenzano statisticamente il tasso di rilevamento degli adenomi nella popolazione presa in considerazione.

#### 4.3.6 Analisi descrittiva dei dati del follow up

Questa analisi è uguale a quella fatta per la popolazione dello screening di II livello.

È stato svolto il confronto tra le due proporzioni campionarie riguardanti il sesso biologico dei pazienti appartenenti al gruppo del follow up. Nella seguente tabella (Tab. 11) vengono riportati il numero degli individui di entrambi i sessi a cui è stato o meno rilevato almeno un adenoma (rispettivamente 1 e 0).

Tabella 11: *M*=sesso biologico maschile; *F*=sesso biologico femminile; 0=nessun adenoma rilevato; 1=almeno un adenoma rilevato

	<b>0</b>	<b>1</b>	
<b>M</b>	29	18	47
<b>F</b>	31	7	38
	60	25	

Le proporzioni campionarie sono:

$$\text{Proporzione campionaria } M = \frac{18}{47} = 0,3830$$

$$\text{Proporzione campionaria } F = \frac{7}{38} = 0,1842$$

Anche in questo caso, per confrontare le due proporzioni campionarie e determinare se sono statisticamente significative, è stato utilizzato sia il test del chi-quadrato che il t-test. Nella tabella di seguito (Tab. 12) sono riportati i risultati ottenuti:

Tabella 12: Valore P

	<b>Test Chi-quadrato</b>	<b>T- Test</b>
<b>Valore P</b>	0.02	0.26

Il test del chi-quadrato mostra che c'è una differenza statisticamente significativa tra le due proporzioni campionarie.

La regressione logistica, svolta per valutare la relazione tra l'età e la probabilità degli adenomi, è stata calcolata utilizzando il modello di regressione lineare generalizzato con una distribuzione binomiale e la funzione *logit*. Di seguito sono riportate le statistiche del modello (tab. 13) e il grafico ottenuto (fig. 15).



Tabella 13: Statistiche del modello

	<i>Coefficienti stimati (<math>\beta</math>)</i>	<i>Errore standard</i>	<i>t-Statistico</i>	<i>Valore P</i>
<i>Intercetta</i>	-1,81	2,19	-0,83	0,41
<i>Età</i>	0,02	0,03	0,69	0,49

- *Z – Score* = 0,69
- *Odds Ratio* = 1,02

Il valore dello z-score ottenuto è compreso nell'intervallo  $\pm 1.96$ . Di conseguenza l'odds ratio non è statisticamente diverso da zero e, in aggiunta, è possibile sostenere che l'età non è un fattore che può migliorare statisticamente il tasso di rilevamento degli adenomi.

Nel grafico sottostante (fig. 15) è rappresentata la regressione logistica che mostra l'aumento della probabilità di rilevare adenomi all'aumentare dell'età del paziente.

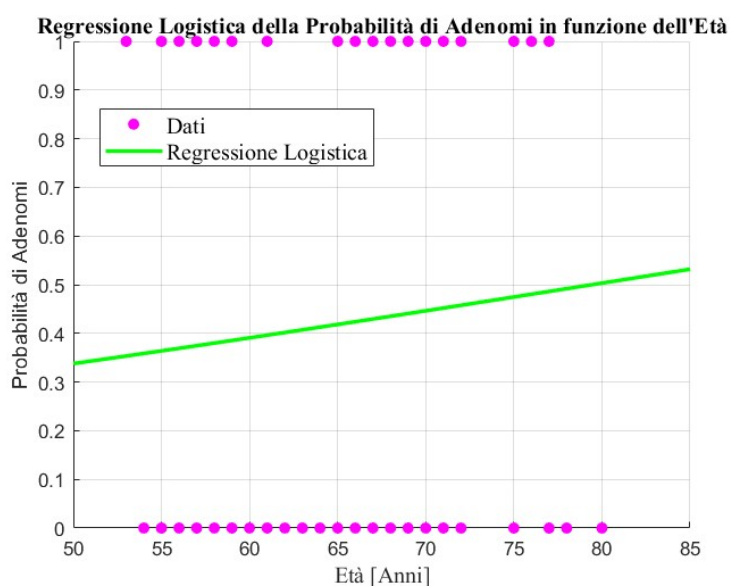


Figura 15: Regressione logistica della probabilità di rilevare adenomi in funzione dell'età

Infine la regressione logistica multivariata della probabilità di rilevare adenomi in funzione dell'età e del sesso del paziente è stata calcolata utilizzando lo stesso modello di regressione precedente descritto e le statistiche ottenute sono state riportate in tabella 14.

Tabella 14: Statistiche del modello

	<i>Coefficienti stimati (<math>\beta</math>)</i>	<i>Errore standard</i>	<i>t-Statistico</i>	<i>Valore P</i>
<i>Intercetta</i>	-0,89	2,29	-0,39	0,70
<i>Età</i>	0,01	0,03	0,45	0,65
<i>Sesso</i>	-1,00	0,46	-2,16	0,03

Dalla Curva ROC e dal  $AUC=0,6460$  si deduce che il modello non funziona ovvero non è diverso da uno senza covariate.

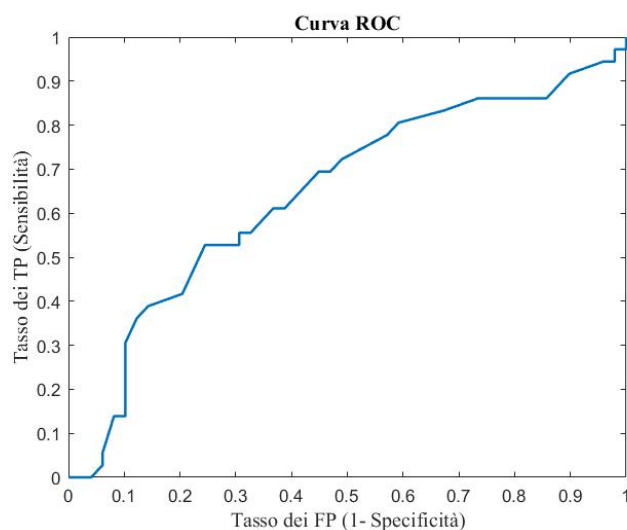


Figura 16: Curva ROC del modello

Dagli z-score riportati nella tabella 14 si evince che il sesso è un fattore di rischio che influenza statisticamente la prevalenza di adenomi a differenza dell'età, in accordo con quanto ottenuto dal test chi-quadrato.

Dai risultati ottenuti, sia nell'analisi dello screening di II livello che in quella del follow up, è possibile sostenere che:

- le due popolazioni, in questo studio, non hanno un tasso di adenoma diverso;
- la regressione logistica della probabilità di adenomi in funzione dell'età evidenzia che ci sono diverse probabilità di adenomi a seconda dell'età del paziente, anche se il legame con l'età non è statisticamente significativo;
- dalla regressione multivariata della prevalenza di adenomi in funzione del sesso e dell'età si evince che, per lo screening di II livello, il modello discrimina casualmente la prevalenza di adenomi nelle due classi (sesso ed età), mentre per il follow up soltanto la variabile del sesso è un fattore che può migliorare statisticamente il tasso di rilevamento degli adenomi.

#### 4.3.7 Confronto delle proporzioni campionarie (WL / WL+AI)

In questa analisi è stato svolto il confronto delle proporzioni campionarie riguardo i due bracci dello studio. Nella tabella (tab. 15) vengono riportati il numero degli individui appartenenti a

ciascun gruppo (WL e WL+AI) a cui è stato o meno rilevato almeno un adenoma (rispettivamente 1 e 0).

Tabella 15: WL=colonscopia standard; WL+AI=colonscopia standard con l'ausilio dell'IA; 0=nessun adenoma rilevato; 1=almeno un adenoma rilevato

	<b>0</b>	<b>1</b>	
<b>WL</b>	80	38	118
<b>WL+AI</b>	74	53	127
	154	91	

Le proporzioni campionarie sono state calcolate nel seguente modo:

$$Pz. \text{ senza AI con almeno 1 adenoma} = \frac{pz \text{ senza AI con almeno un adenoma}}{popolazione \text{ totale senza AI}} = 0,3220$$

$$Pz. \text{ con AI con almeno 1 adenoma} = \frac{pz \text{ con AI con almeno un adenoma}}{popolazione \text{ totale con AI}} = 0,4173$$

Il confronto delle due proporzioni è stato svolto con il test del chi-quadrato e con il t-test. I risultati ottenuti sono riportati in tabella (tab. 16).

Tabella 16: Valore P

	<b>Test Chi-quadrato</b>	<b>T- Test</b>
<b>Valore P</b>	0.12	0.36

Entrambi i test sottolineano che la differenza tra le due proporzioni campionarie non è statisticamente significativa.

#### 4.3.8 Regressioni logistiche

Sono state svolte due regressioni logistiche per analizzare l'influenza che il tempo di retrazione ha nei confronti della probabilità di rilevare adenomi.

La regressione logistica è stata calcolata utilizzando il modello di regressione lineare generalizzato con una distribuzione binomiale e la funzione *logit*. Le statistiche del modello (tab. 17) e il grafico ottenuto (fig. 17) sono riportati di seguito:

Tabella 17: Statistiche del modello

	<i>Coefficienti stimati (<math>\beta</math>)</i>	<i>Errore standard</i>	<i>t-Statistico</i>	<i>Valore P</i>
<i>Intercetta</i>	-2,31	0,42	-5,46	$4,66 \cdot 10^{-8}$
<i>T. di retrazione</i>	0,23	0,05	4,53	$5,77 \cdot 10^{-6}$

- *Z – Score* = 4,53
- *Odds Ratio* = 1,26

Il valore dello z-score ottenuto è al di fuori nell'intervallo  $\pm 1.96$ . Di conseguenza l'odds ratio è statisticamente diverso da zero e, in aggiunta, è possibile sostenere che il tempo di retrazione è un fattore di rischio che può migliorare statisticamente il tasso di rilevamento degli adenomi. Infatti la figura sottostante (fig. 17) evidenzia che tanto più veloce è il tempo di retrazione, tanto meno si riescono a identificare gli adenomi.

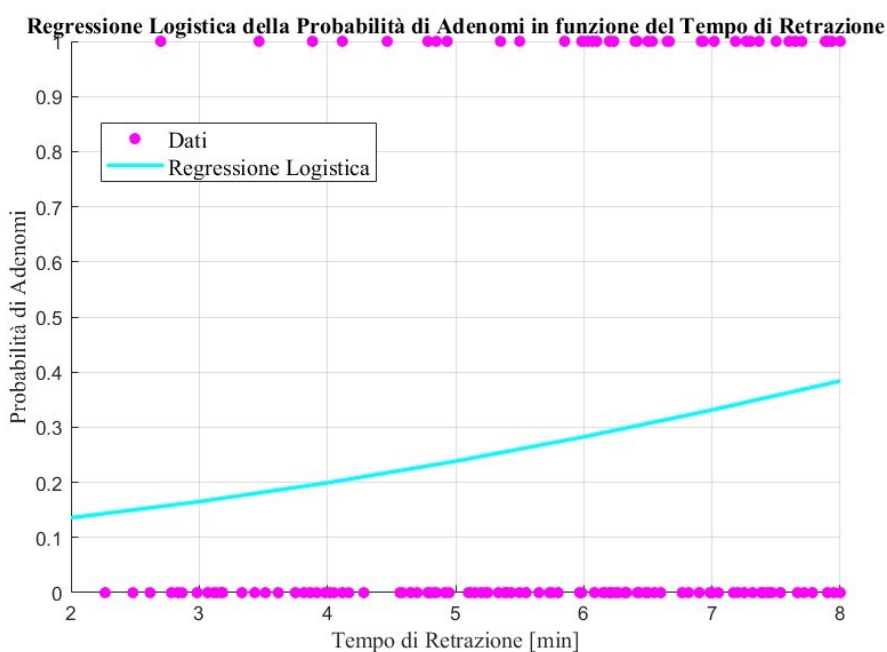


Figura 17: Regressione logistica della probabilità di rilevare adenomi in funzione del tempo di retrazione

L'ultima analisi è composta da due regressioni logistiche che calcolano la relazione tra la probabilità di rilevare adenomi in funzione del tempo di ritiro impiegato sia durante le colonscopie standard (WL) sia in quelle dotate di IA (WL+AI).

Il modello della regressione è lo stesso utilizzato precedentemente e di seguito sono riportati i risultati:

Tabella 18: Statistiche del modello

<i>AI OFF</i>	<i>Coefficienti stimati (<math>\beta</math>)</i>	<i>Errore standard</i>	<i>t-Statistico</i>	<i>Valore P</i>
<i>Intercetta</i>	-2,25	0,58	-3,89	$9,81 \cdot 10^{-5}$
<i>T. di retrazione</i>	0,20	0,07	2,84	0,004

Tabella 19: Statistiche del modello

<i>AI ON</i>	<i>Coefficienti stimati (<math>\beta</math>)</i>	<i>Errore standard</i>	<i>t-Statistico</i>	<i>Valore P</i>
<i>Intercetta</i>	-2,38	0,62	-3,84	$1,23 \cdot 10^{-4}$
<i>T. di retrazione</i>	0,26	0,07	3,53	$4,13 \cdot 10^{-4}$

- **Odds Ratio (AI OFF) = 1,22**
- **Odds Ratio (AI ON) = 1,30**

Entrambi gli z-score sono al di fuori del range  $\pm 1,96$ : di conseguenza anche i coefficienti OD sono statisticamente diversi da zero. Da ciò si deduce che il tempo di retrazione è un fattore di rischio che influenza la probabilità di rilevare adenomi.

In questo grafico (fig.18) è possibile osservare che all'aumentare del tempo di ritiro, la probabilità di trovare adenomi è leggermente superiore nei casi in cui si svolge la colonscopia standard con l'ausilio dell'IA.

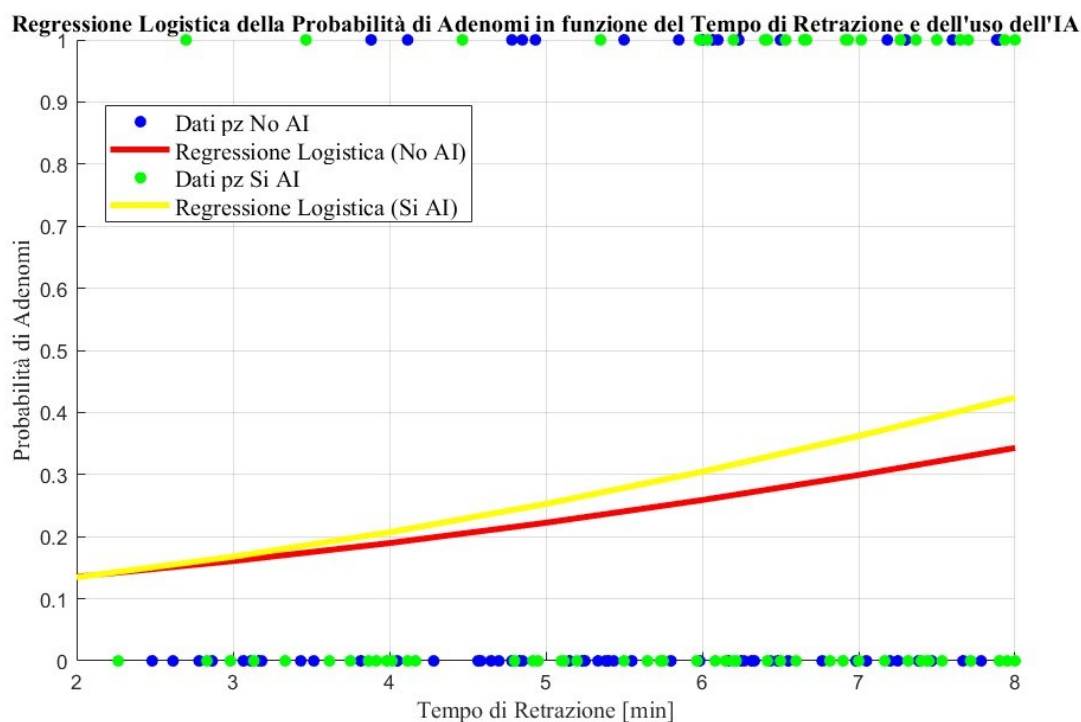


Figura 18: Regressione logistica della probabilità di rilevare adenomi in funzione del tempo di retrazione e dell'uso dell'IA

## 5 Discussione

L'obiettivo primario di questo studio randomizzato controllato monocentrico è quello di verificare se l'assistenza dell'IA durante le colonscopie contribuisse a migliorare le prestazioni dell'endoscopista e, in generale, la qualità dell'esame in una popolazione di screening mista. Lo studio ha coinvolto 6 endoscopisti, ciascuno dei quali aveva molta esperienza (>500 colonscopie l'anno) e un ADR personale confrontabile con quello medio nazionale degli endoscopisti. I risultati confermano che, sebbene non significativo, l'uso dell'IA ha comportato un aumento dell'ADR sia per la popolazione complessiva, come mostrato nella tabella 2, sia per le sottocategorie considerate, come mostrato nelle tabelle 3 e 4. Il risultato viene confermato anche nell'articolo di *Hüneburg, R. et al.* [99].

La letteratura [100] [67] sostiene che l'ADR è l'indicatore di qualità più frequentemente utilizzato e, idealmente, dovrebbe essere  $\geq 25\%$ . Lo studio effettuato riporta un ADR complessivo del 37,14% in linea con la letteratura. È utile precisare che questa percentuale include, oltre alle popolazioni di screening e di follow up, anche la popolazione degli accessi spontanei.

Si sostiene, inoltre, che nei programmi di screening del CCR basati su FIT, in cui il numero di adenomi rilevati è più elevato, l'ADR deve essere almeno del 40% [101]. Lo studio riporta una percentuale minimamente inferiore (39,86%) ma ugualmente accettabile. A differenza di quanto emerso dalla letteratura [102] [103], l'ADR nella popolazione dei follow up è superiore rispetto a quella dello screening (42,35% vs 39,86%). Questa discrepanza potrebbe dipendere dalla fascia di età dei pazienti in follow up che si estende fino agli 80 anni ed è, quindi, più ampia rispetto a quella dei pazienti in screening. Di conseguenza, poiché dalle analisi di regressione logistica è emerso che per entrambe le popolazioni il tasso di rilevamento degli adenomi aumenta all'aumentare dell'età (fig. 13 e fig. 15), considerare anche i pazienti ultrasettantenni nel follow up potrebbe influire sull'aumento dell'ADR. Anche se non vi è alcuna prova diretta che l'ADR nei follow-up aumenti con l'età della popolazione, i seguenti articoli [104] [105] [106] suggeriscono che l'ADR potrebbero aumentare, in generale, con l'età. Inoltre è importante considerare che le cause principali per cui i pazienti vengono inseriti nel follow up sono il rilevamento di adenomi e/o polipi e l'esecuzione di una pulizia intestinale appena sufficiente. Queste cause potrebbero dipendere da diverse situazioni: il rilevamento di adenomi potrebbe riflettere la predisposizione individuale, come ad esempio lo stile di vita o la storia familiare del paziente [107]; il livello di pulizia intestinale potrebbe aver influenzato esami pregressi con limitazioni visive che non avrebbero consentito una completa individuazione

degli adenomi [108]. Questi risultati potrebbero essere oggetto di studi futuri per comprendere se esista una relazione tra l'ADR nei pazienti follow up e l'età degli stessi.

Successivamente è stata svolta l'analisi delle proporzioni campionarie riguardo il sesso biologico maschile e femminile, sia della popolazione di screening di II livello che di quella del follow up, dalla quale è emerso che i tassi di rilevamento degli adenomi nella popolazione maschile e femminile sono conformi alle linee guida della colonscopia [67] ( $\geq 30\%$  negli uomini,  $\geq 20\%$  nelle donne). I risultati del test chi-quadrato della popolazione in follow up mostrano differenze statisticamente significative tra le popolazioni (tab. 12), come confermato dalla letteratura [109] [110] [111]. I risultati ottenuti dalla popolazione dello screening, invece, mostrano che il tasso di rilevamento dell'adenoma negli uomini e nelle donne non ha differenze statisticamente significative (tab. 8). Ciò può essere causato dal numero simile di partecipanti di entrambe le popolazioni (70 pazienti maschi vs 68 femmine). Non sono presenti in letteratura evidenze scientifiche in merito a questo caso; tuttavia una dimensione ridotta del campione potrebbe limitare la capacità di individuare differenze reali nell'ADR.

L'obiettivo secondario dello studio di tesi è quello di valutare se l'introduzione dell'ausilio dell'IA potesse influire sui tempi di retrazione e sulla durata di occupazione della sala. Lo studio ha dimostrato che il tempo medio di retrazione delle colonscopie dotate di IA è minimamente superiore (33 secondi) rispetto a quello delle colonscopie standard; il risultato è in linea con alcune ricerche e metanalisi [112] [113] [114]. I prolungamenti nel ritiro dell'endoscopio possono derivare da diversi fattori quali l'efficace miglioramento nella capacità di individuare adenomi, il tempo necessario per rimuoverli, una maggiore attenzione alla sicurezza e l'analisi approfondita delle segnalazioni del modulo di IA, che può talvolta generare falsi positivi. Con la crescente abilità e familiarità nell'utilizzo dell'IA è possibile che i tempi di retrazione si riducano fino a diventare equiparabili a quelli delle procedure standard. Gli endoscopisti saranno in grado di riconoscere, caratterizzare e scartare rapidamente i segnali non neoplastici individuati dall'IA, migliorando così l'efficienza complessiva del processo.

L'ultimo obiettivo del lavoro svolto è quello di verificare se l'età, il sesso, l'uso dell'IA e il tempo di retrazione sono fattori che influenzano il tasso di rilevamento degli adenomi. Nel caso della popolazione di screening di II livello è emerso che sia il sesso che l'età non sono fattori che influenzano l'ADR, in contrasto con la letteratura. Riguardo la popolazione di follow up, invece, è emerso che soltanto il sesso è un fattore che influisce sull'ADR. L'incoerenza con la letteratura è giustificabile considerando che la dimensione del campione è ridotta e lo studio, in quanto monocentrico, si basa soltanto su una popolazione appartenente ad una piccola area

territoriale. Un'analisi che consideri più variabili demografiche e geografiche potrebbe rivelare differenze più significative.

I seguenti articoli [115] [116] [117] [118] [119] [120] [121] sostengono che l'aumento del tempo di retrazione può migliorare l'ADR e che è probabilmente necessario un tempo di retrazione minimo, compreso tra 6 e 9 minuti, per una efficace colonscopia di screening. I risultati ottenuti dalla regressione logistica, riguardo l'ADR in funzione del tempo di retrazione, sono in linea con la letteratura: in particolare è emerso che all'aumentare del tempo di retrazione, aumenta la probabilità di rilevare adenomi (fig. 17).

Dalla letteratura emerge che l'uso dell'IA nelle colonscopie porta a un aumento dei tassi di rilevamento degli adenomi rispetto alle colonscopie standard e questo aumento è più evidente quando si prolunga il tempo di retrazione in entrambi i tipi di procedure [19] [122] [123]. Lo studio mostra quanto affermato negli articoli citati, anche se come mostrato nella figura 18 la differenza tra i due casi è minima. È fondamentale sottolineare che, nonostante gli endoscopisti partecipanti allo studio fossero altamente esperti, l'IA ha comunque rappresentato un valore aggiunto, seppur modesto.

Sarebbe opportuno riflettere sull'efficacia della pulizia intestinale dei pazienti. La maggioranza di essi ha ricevuto una valutazione pari a 6 sulla *Boston Scale*, il che rappresenta un livello appena sufficiente per considerare l'esame adeguato. Questo studio ha evidenziato che l'impiego dell'IA nella colonscopia ha il vantaggio di migliorare la rilevazione degli adenomi, e ciò è stato possibile anche grazie all'*expertise* degli endoscopisti nel neutralizzare gli effetti di una preparazione intestinale non ottimale.

Questo studio presenta alcune limitazioni rilevanti. Come già accennato, la dimensione ridotta del campione ha influenzato in modo significativo i risultati ottenuti. Inoltre il fatto che lo studio sia condotto in un unico centro limita la possibilità di confrontarsi su una scala geografica più ampia. Va notato che, in centri di screening di alto livello come quello coinvolto nello studio, dove ogni endoscopista esegue oltre 500 colonscopie l'anno e si raggiungono volumi annui superiori a 4000 esami, il contributo dell'IA può apparire limitato.

È importante sottolineare che l'IA potrebbe avere un impatto più significativo in contesti diversi, come centri con volumi di esami inferiori, medici meno esperti nel riconoscere adenomi e tempi di retrazione meno adeguati. L'IA può contribuire a livellare l'indice di rilevamento degli adenomi nella media o persino ad aumentarlo nei contesti in cui è naturalmente basso.



Sono emerse anche alcune limitazioni di natura tecnica nell'applicazione del sistema di IA. In particolare vi è la necessità di ulteriori miglioramenti nel funzionamento dell'IA per ottimizzare il riconoscimento di specifiche anomalie. Ne emerge che l'IA non è ancora completamente matura come strumento diagnostico.

## 6 Conclusioni

In conclusione, la presente ricerca conferma quanto emerso da studi pregressi, evidenziando che l'impiego dell'IA nelle colonscopie di screening rappresenta una promettente innovazione nell'ambito della diagnostica endoscopica. I risultati indicano che l'IA può contribuire in modo significativo al rilevamento di un numero più elevato di adenomi, soprattutto nei pazienti compresi tra i 50 e gli 80 anni, un gruppo demografico a maggior rischio di sviluppare lesioni precancerose.

Va sottolineato che, sebbene l'ausilio dell'IA possa comportare un aumento del tempo di retrazione dell'endoscopio, questo incremento è giustificato da una combinazione positiva con l'indice di qualità dell'esame. Quindi il leggero allungamento della procedura risulta essere un “investimento” in termini di accuratezza diagnostica, concretizzatosi in un aumento dell'ADR.

Tuttavia è importante notare che l'attuale tecnologia di IA utilizzata nello studio ha evidenziato alcune limitazioni tecniche. Quest'ultime sottolineano la necessità di ulteriori aggiornamenti e miglioramenti al fine di ottimizzare la diagnostica in tempo reale. Ciò mette in luce l'importanza della ricerca continua e dello sviluppo tecnologico nell'ambito dell'IA applicata all'endoscopia digestiva.

In prospettiva, la combinazione di competenze mediche e soluzioni avanzate di IA promette di rivoluzionare la pratica endoscopica migliorando la capacità diagnostica, riducendo gli errori e contribuendo a una diagnosi più precoce e accurata delle patologie gastrointestinali. Tali risultati fanno strada a ulteriori ricerche e all'implementazione dell'IA in ambito clinico, con l'obiettivo di migliorare la salute e la qualità di vita dei pazienti.

Un interessante tema di ricerca futuro potrebbe riguardare l'indagine sulla sinergia tra l'esperienza dei medici e l'impiego dell'IA per ottimizzare l'efficienza delle colonscopie. Questo studio mirerebbe a trovare un equilibrio tra accelerare la procedura e mantenere elevati standard di qualità, al fine di ridurre le liste d'attesa e rendere più accessibili gli esami di screening a una più ampia gamma di pazienti.

## 7 Bibliografia

- [1] M. Maida, F. S. Macaluso, G. Ianiro, F. Mangiola, E. Sinagra, G. Hold, C. Maida, G. Cammarota, A. Gasbarrini e G. Scarpulla, «Screening of colorectal cancer: present and future,» *Expert Review of Anticancer Therapy*, vol. 17, n. 12, pp. 1131-1146, 2017.
- [2] «Healthcare remains the hottest AI category for deals.,» *CB Insights Research*, 2017.
- [3] S. L. Andresen, «John McCarthy: father of AI,» *IEEE Intelligent Systems*, vol. 17, n. 5, pp. 84-85, 2002.
- [4] M. Tripathi, «EHR evolution: policy and legislation forces changing the EHR.,» *J AHIMA.*, vol. 83, n. 10, pp. 24-29, 2012.
- [5] A. Malva e V. Zurlo, «La medicina nell'era dell'Intelligenza Artificiale: applicazioni in Medicina Generale,» *Rivista Società Italiana di Medicina Generale*, vol. 26, n. 4, 2019.
- [6] V. Buch, I. Ahmed e M. Maruthappu, «Artificial intelligence in medicine: current trends and future possibilities,» *Br J Gen Pract.*, vol. 68, pp. 143-144, 2018.
- [7] M. Burgess, «L'NHS sta sperimentando un chatbot AI per rispondere alle tue domande mediche,» *Cablato*, 2017.
- [8] A. Esteva, B. Kuprel, R. Novoa, J. Ko, S. Swetter, H. Blau e S. Thrun, «Dermatologist-level classification of skin cancer with deep neural networks.,» *Nature*, vol. 542, pp. 115-118, 2017.
- [9] N. Asada, K. Doi, H. MacMahon, S. Montner, M. Giger, C. Abe e Y. Wu, «Potential usefulness of an artificial neural network for differential diagnosis of interstitial lung diseases: Pilot study,» *Radiology*, vol. 177, n. 3, p. 857 – 860, 1990.
- [10] P. Liu, L. Lu, J. Zhang, T. Huo, S. Liu e Z. Ye, «Application of Artificial Intelligence in Medicine: An Overview.,» *Curr Med Sci*, vol. 41, pp. 1105-1115, 2021.
- [11] F. Sorrentino, G. Jurman, K. De Nadai, C. Campa, C. Furlanello e F. Parmeggiani, «Application of Artificial Intelligence in Targeting Retinal Diseases.,» *Curr Drug Targets.*, pp. 1208-1215., 2020.
- [12] J. Gong, J. Liu, X. Sun, B. Zheng e S. Nie, «Computer-aided diagnosis of lung cancer: the effect of training data sets on classification accuracy of lung nodules.,» *Phys Med Biol.*, 2018.
- [13] A. Rodriguez-Ruiz, K. Lång, A. Gubern-Merida, M. Broeders, G. Gennaro, P. Clauser, T. Helbich, M. Chevalier, T. Tan, T. Mertelmeier, M. Wallis, I. Andersson, S. Zackrisson, R. Mann, I. Sechopoulos e Stand-Alone, «Artificial Intelligence for Breast Cancer Detection in Mammography: Comparison With 101 Radiologists,» *J Natl Cancer Inst.*, 2019.

- [14] S. Wang, D. Yang, R. Rong, X. Zhan e G. Xiao, «Pathology Image Analysis Using Segmentation Deep Learning Algorithms.,» *Am J Pathol.*, pp. 1686-1698, 2019.
- [15] I. S. Komura D, «Machine learning approaches for pathologic diagnosis,» *Virchows Arch*, vol. 475, pp. 131-138, 2019.
- [16] O. Iizuka, F. Kanavati, K. Kato e e. al., «Deep Learning Models for Histopathological Classification of Gastric and Colonic Epithelial Tumours,» *Sci Rep*, vol. 10, p. 1504, 2020.
- [17] S. Wang, D. Yang, R. Rong, X. Zhan, J. Fujimoto, H. Liu, J. Minna, I. Wistuba, Y. Xie e G. Xiao, «Artificial Intelligence in Lung Cancer Pathology Image Analysis.,» *Cancers (Basel).*, vol. 11, p. 1673, 2019.
- [18] K. Namikawa, T. Hirasawa, T. Yoshio, J. Fujisaki, T. Ozawa, S. Ishihara, T. Aoki, A. Yamada, K. Koike, H. Suzuki e T. Tada, «Utilizing artificial intelligence in endoscopy: a clinician's guide.,» *Expert Rev Gastroenterol Hepatol.*, vol. 14, pp. 689-706, 2020.
- [19] P. Wang, T. Berzin, J. Glissen Brown, S. Bharadwaj, A. Becq, X. Xiao, P. Liu, L. Li, Y. Song, D. Zhang, Y. Li, G. Xu, M. Tu e X. Liu, «Real-time automatic detection system increases colonoscopic polyp and adenoma detection rates: a prospective randomised controlled study.,» *Gut*, vol. 68, pp. 1813-1819, 2019.
- [20] Y. Hwang, H. Lee, C. Park, B. Tama, J. Kim, D. Cheung, W. Chung, Y. Cho, K. Lee, M. Choi, S. Lee e B. Lee, «Improved classification and localization approach to small bowel capsule endoscopy using convolutional neural network.,» *Dig Endosc.*, vol. 33, pp. 598-607, 2021.
- [21] A. Mitsala, C. Tsalikidis, M. Pitiakoudis, C. Simopoulos e A. Tsaroucha, «Artificial Intelligence in Colorectal Cancer Screening, Diagnosis and Treatment. A New Era.,» *Curr Oncol.*, vol. 28, n. 3, pp. 1581-1607, 2021.
- [22] Z. Perkins, B. Yet, A. Sharrock, R. Rickard, W. Marsh, T. Rasmussen e N. Tai, «Predicting the Outcome of Limb Revascularization in Patients With Lower-extremity Arterial Trauma: Development and External Validation of a Supervised Machine-learning Algorithm to Support Surgical Decisions.,» *Ann Surg.*, vol. 272, pp. 564-572, 2020.
- [23] A. Gumbs, S. Perretta, B. d'Allemagne e E. Chouillard, «What is Artificial Intelligence Surgery?,» *Art. Int. Surg.*, vol. 1, pp. 1-10, 2021.
- [24] M. Portelli, S. Bianco, T. Bezzina e J. Abela, «Virtual reality training compared with apprenticeship training in laparoscopic surgery: a meta-analysis.,» *Ann R Coll Surg Engl*, vol. 102, pp. 672-684, 2020.
- [25] J. Coles-Black, D. Bolton, D. Robinson e J. Chuen, «Utility of 3D printed abdominal aortic aneurysm phantoms: a systematic review.,» *ANZ Journal of Surgery*, vol. 91, pp. 1673-1681, 2021.

- [26] J. Park, H. Kang, J. Kim e H. Kim, «The application of 3D-printing technology in pelvic bone tumor surgery.,» *J Orthop Sci.*, vol. 26, pp. 276-283, 2021.
- [27] A. Poncette, L. Mosch, C. Spies, M. Schmieding, F. Schiefenhövel, H. Krampe e F. Balzer, «Improvements in Patient Monitoring in the Intensive Care Unit: Survey Study.,» *J Med Internet Res.*, vol. 22, 2020.
- [28] Z. Angehrn, L. Haldna, A. Zandvliet, E. Gil Berglund, J. Zeeuw, B. Amzal, S. Cheung, T. Polasek, M. Pfister, T. Kerbusch e N. Heckman, «Artificial Intelligence and Machine Learning Applied at the Point of Care.,» *Frontiers in pharmacology*, vol. 11, p. 759, 2020.
- [29] Y. Zhao, C. Liang, Z. Gu, Y. Zheng e Q. Wu, «A New Design Scheme for Intelligent Upper Limb Rehabilitation Training Robot.,» *Int J Environ Res Public Health.*, vol. 17, 2020.
- [30] R. Ramezani, W. Zhang, Z. Xie, J. Shen, D. Elashoff, P. Roberts, A. Stanton, M. Eslami, N. Wenger, M. Sarrafzadeh e A. Naeim, «A Combination of Indoor Localization and Wearable Sensor-Based Physical Activity Recognition to Assess Older Patients Undergoing Subacute Rehabilitation: Baseline Study Results.,» *JMIR Mhealth Uhealth.*, 2019.
- [31] F. Zanca, C. Brusasco, F. Pesapane, Z. Kwade, R. Beckers e M. Avanzo, «Regulatory Aspects of the Use of Artificial Intelligence Medical Software.,» *Seminars in Radiation Oncology*, 2022.
- [32] «Regolamento (UE) 2017/745 del Parlamento Europeo e del Consiglio dell'Europa del 5 aprile 2017 relativo ai dispositivi medici (MDR 745).».
- [33] E. Parliament, «Artificial intelligence act.,» 2023.
- [34] F. Miriam,  
«[https://www.europarl.europa.eu/RegData/etudes/BRIE/2021/698792/EPRS\\_BRI\(2021\)698792\\_EN.pdf](https://www.europarl.europa.eu/RegData/etudes/BRIE/2021/698792/EPRS_BRI(2021)698792_EN.pdf),» *Altalex*, 2023.
- [35] «AI Act: il Regolamento Europeo sull'Intelligenza Artificiale.,» *Namirial Focus*, 2023.
- [36] P. Marinoni, «AI Act e tutela del copyright.,» [Online]. Available: <https://ainews.it/ai-act-e-tutela-del-copyright/>.
- [37] J. Zhang e Z. Zhang, «Ethics and governance of trustworthy medical artificial intelligence.,» *BMC Med Inform Decis Mak*, 2023.
- [38] S. Gerke, T. Minssen, H. Yu e e. al., «Ethical and legal issues of ingestible electronic sensors.,» *Nat Electron*, 2019.
- [39] C. Ross e I. Swetlitz, «IBM's Watson supercomputer recommended 'unsafe and incorrect' cancer treatments, internal documents show.,» *STAT+*, 2018.

- [40] D. Char, N. Shah e D. Magnus, «Implementing machine learning in health care-addressing ethical challenges.,» *The New England Journal of Medicine*, 2018.
- [41] A. Howard e J. Borenstein, «The Ugly Truth About Ourselves and Our Robot Creations: The Problem of Bias and Social Inequity.,» *Sci Eng Ethics*, vol. 24, p. 1521–1536 , 2018.
- [42] S. J. Mooney e V. Pejaver, «Big Data in Public Health: Terminology, Machine Learning, and Privacy.,» *Annual Review of Public Health*, 2018.
- [43] A. Adamson e A. Smith, «Machine Learning and Health Care Disparities in Dermatology.,» *JAMA Dermatol.*, vol. 154, n. 11, pp. 1247-1248, 2018.
- [44] Z. Obermeyer, B. Powers, C. Vogeli e S. Mullainathan, «Dissecting racial bias in an algorithm used to manage the health of populations.,» *Science*, vol. 366, n. 6464, pp. 447-453, 2019.
- [45] J. Burrell, «How the machine ‘thinks’: Understanding opacity in machine learning algorithms.,» *Big Data & Society*, 2016.
- [46] P. Esmailzadeh, T. Mirzaei e S. Dharanikota, «Patients' Perceptions Toward Human-Artificial Intelligence Interaction in Health Care: Experimental Study,» *J Med Internet Res.*, vol. 23, n. 11, 2021.
- [47] K. Sarah, «Artificial Intelligence and Artificial Intelligence and Technology in Health Care technology in Health Care: Overview,» *DePaul Journal of Health Care Law*, 2020.
- [48] A. Holzinger, C. Biemann, C. Pattichis e D. Kell, «What do we need to build explainable AI systems for the medical domain?,» *arXiv*, 2017.
- [49] G. Katznelson e S. Gerke, «The need for health AI ethics in medical school education.,» *Adv in Health Sci Educ*, 2021.
- [50] W. Price, S. Gerke e I. Cohen, «Potential Liability for Physicians Using Artificial Intelligence.,» *JAMA*, 2019.
- [51] S. Finlayson, J. Bowers, J. Ito e e. al., «Adversarial attacks on medical machine learning,» *Science.*, 2019.
- [52] «The World's Most Valuable Resource is No Longer Oil, But Data.,» *The Economist*, 2021.
- [53] K. Warwick e H. Shah, «Passing the turing test does not mean the end of humanity.,» *Cogn Comput*, 2016.
- [54] E. Commission., «European Parliament, Committee on Legal Affairs: Draft Report with Recommendations to the Commission on Civil Law Rules on Robotics.,» 2016.
- [55] J. Chung e A. Zink, «Hey Watson, can i sue you for malpractice? Examining the liability of artificial intelligence in medicine.,» *Asia Pac J Health Law Eth.*, 2018.

- [56] «Information technology - Artificial Intelligence - Overview of trustworthiness in artificial intelligence,» in *Norma ISO/IEC 24028*, 2020.
- [57] J. Regula, M. Rupinski, E. Kraszewska, M. Polkowski, J. Pachlewski, J. Orłowska, M. Nowacki e E. Butruk, «Colonoscopy in colorectal-cancer screening for detection of advanced neoplasia.,» *N Engl J Med.*, vol. 355, n. 18, pp. 1863-72, 2006.
- [58] K. Simon, «Colorectal cancer development and advances in screening.,» *Clin Interv Aging.*, vol. 11, pp. 967-76, 2016.
- [59] «Prevention of Colorectal Cancer by Flexible Endoscopy and Polypectomy: A Case-Control Study of 32 702 Veterans,» *Annals of Internal Medicine*, vol. 123, n. 12, pp. 904-910, 1995.
- [60] B. Levin e e. al., «Screening and Surveillance for the Early Detection of Colorectal Cancer and Adenomatous Polyps, 2008: A Joint Guideline From the American Cancer Society, the US Multi-Society Task Force on Colorectal Cancer, and the American College of Radiology,» *Gastroenterology*, vol. 134, n. 5, pp. 1570-1595, 2008.
- [61] G. Cooper, T. Kou e D. Rex, «Complications following colonoscopy with anesthesia assistance: a population-based analysis.,» *JAMA Intern Med.*, vol. 173, n. 7, pp. 551-556, 2013.
- [62] J. Inadomi, S. Vijan, N. Janz, A. Fagerlin, J. Thomas, Y. Lin, R. Muñoz, C. Lau, M. Somsouk, N. El-Nachef e R. Hayward, «Adherence to colorectal cancer screening: a randomized clinical trial of competing strategies.,» *Arch Intern Med.*, vol. 172, n. 7, pp. 575-82, 2012.
- [63] F. Stracci, M. Zorzi e G. Grazzini, «Colorectal cancer screening: tests, strategies, and perspectives.,» *Front Public Health*, pp. 2-210, 2014.
- [64] D. Corley, C. Jensen, A. Marks, W. Zhao, J. Lee, C. Doubeni, A. Zauber, J. de Boer, B. Fireman, J. Schottinger, V. Quinn, N. Ghai, T. Levin e C. Quesenberry, «Adenoma detection rate and risk of colorectal cancer and death,» *N Engl J Med.*, vol. 370, n. 14, pp. 1298-306, 2014.
- [65] D. Rex, «Key quality indicators in colonoscopy,» *Gastroenterol Rep (Oxf).*, vol. 11, 2023.
- [66] N. Baxter, R. Sutradhar e e. al., «Analysis of Administrative Data Finds Endoscopist Quality Measures Associated With Postcolonoscopy Colorectal Cancer,» *Gastroenterology*, vol. 140, pp. 65-72, 2011.
- [67] D. K. Rex, P. S. Schoenfeld, J. Cohen, I. M. Pike, D. G. Adler, B. M. Fennerty, J. G. Lieb, W. G. Park, M. K. Rizk, M. S. Sawhney, N. J. Shaheen, S. Wani e D. S. Weinberg, «Quality Indicators for Colonoscopy,» *American Journal of Gastroenterology*, vol. 110, pp. 72-90, 2015.

- [68] K. Vemulapalli, S. Wilder, C. Kahi e D. Rex, «Long-Term Assessment of the Cecal Intubation Rates in High-Performing Colonoscopists: Time for Review.,» *Clin Transl Gastroenterol.*, vol. 11, n. 3, 2020.
- [69] F. Froehlich, V. Wietlisbach, J.-J. Gonvers, B. Burnand e J.-P. Vader, «Impact of colonic cleansing on quality and diagnostic yield of colonoscopy: the European Panel of Appropriateness of Gastrointestinal Endoscopy European multicenter study,» *Gastrointestinal Endoscopy*, vol. 61, n. 3, pp. 378-384, 2005.
- [70] A. Calderwood e B. Jacobson, «Comprehensive validation of the Boston Bowel Preparation Scale,» *Gastrointest Endosc*, vol. 74, n. 4, pp. 686-92, 2010.
- [71] H. Fatima, C. S. Johnson e D. K. Rex, «Patients' description of rectal effluent and quality of bowel preparation at colonoscopy,» *Gastrointestinal Endoscopy*, vol. 71, n. 7, pp. 1244-1252, 2010.
- [72] D. K. Rex, «Optimal bowel preparation—a practical guide for clinicians,» *Nature Reviews Gastroenterology & Hepatology*, vol. 11, pp. 419-425, 2014.
- [73] G. C. Harewood, V. K. Sharma e P. de Garmo, «Impact of colonoscopy preparation quality on detection of suspected colonic neoplasia,» *Gastrointestinal Endoscopy*, vol. 58, n. 1, pp. 76-79, 2003.
- [74] D. K. Rex, T. F. Imperiale, D. Latinovich e L. Bratcher, «Impact of bowel preparation on efficiency and cost of colonoscopy},,,» vol. 97, n. 7, pp. 1696-1700, 2002.
- [75] B. Liem e N. Gupta, «Adenoma detection rate: the perfect colonoscopy quality measure or is there more?,» *Transl Gastroenterol Hepatol.*, pp. 3-19, 2018.
- [76] M. Kaminski, P. Wieszczy, M. Rupinski, U. Wojciechowska, J. Didkowska, E. Kraszewska, J. Kobiela, R. Franczyk, M. Rupinska, B. Kocot, A. Chaber-Ciopinska, J. Pachlewski, M. Polkowski e J. Regula, «Increased Rate of Adenoma Detection Associates With Reduced Risk of Colorectal Cancer and Death.,» *Gastroenterology*, pp. 98-105, 2017.
- [77] M. Aslam, S. Bano, M. Khalid, Z. Sarfraz, A. Sarfraz, M. Sarfraz, K. Robles-Velasco, M. Felix, K. Deane e I. Cherrez-Ojeda, «The effectiveness of real-time computer-aided and quality control systems in colorectal adenoma and polyp detection during colonoscopies: a meta-analysis.,» *Ann Med Surg (Lond).*, vol. 82, n. 2, pp. 80-91, 2023.
- [78] R. Barclay, J. Vicari, A. Doughty, J. Johanson e R. Greenlaw, «Colonoscopic withdrawal times and adenoma detection during screening colonoscopy,» *N Engl J Med*, vol. 355, n. 24, pp. 2533-41, 2006.
- [79] F. Bray, J. Ferlay, I. Soerjomataram, R. Siegel, L. Torre e A. Jemal, «Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries.,» *CA: A Cancer Journal for Clinicians*, vol. 68, 2018.



- [80] B. Lauby-Secretan, N. Vilahur, F. Bianchini, N. Guha, K. Straif e G. I. A. f. R. o. C. H. Working, «The IARC Perspective on Colorectal Cancer Screening.,» *N Engl J Med*, vol. 378, n. 18, pp. 1734-1740, 2018.
- [81] B. Morson, «The polyp-cancer sequence in the large bowel.,» *Proc R Soc Med.*, vol. 67, n. 6, pp. 451-457, 1974.
- [82] J. Lee, E. Liles, S. Bent, T. Levin e D. Corley, «Accuracy of fecal immunochemical tests for colorectal cancer: systematic review and meta-analysis.,» *Ann Intern Med.*, vol. 160, n. 3, p. 171, 2014.
- [83] C. Bray, L. Bell, H. Liang, D. Collins e S. Yale, «Colorectal Cancer Screening,» *WMJ*, vol. 116, n. 1, pp. 27-33, 2017.
- [84] C. Coin, F. Wollett, J. Coin, M. Rowland, R. DeRamos e R. Dandrea, «Computerized radiology of the colon: a potential screening technique.,» *Comput Radiol*, vol. 7, n. 4, pp. 215-21, 1983.
- [85] F. Bray, J. Ferlay, I. Soerjomataram, R. Siegel, L. Torre e A. Jemal, «Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries.,» *CA: A Cancer Journal for Clinicians*, vol. 68, pp. 394-424, 2018.
- [86] B. Levin, D. A. Lieberman, B. McFarland, K. S. Andrews, D. Brooks, J. Bond, C. Dash e e. al, «Screening and Surveillance for the Early Detection of Colorectal Cancer and Adenomatous Polyps, 2008: A Joint Guideline From the American Cancer Society, the US Multi-Society Task Force on Colorectal Cancer, and the American College of Radiology,» *Gastroenterology*, vol. 134, n. 5, pp. 1570-1595, 2008.
- [87] J. Zhang, G. Chen, Z. Li, P. Zhang, X. Li, D. Gan, X. Cao, H. Du, J. Zhang, L. Zhang e Y. Ye, «Colonoscopic screening is associated with reduced Colorectal Cancer incidence and mortality: a systematic review and meta-analysis,» *J Cancer.*, vol. 11, n. 20, pp. 5953-5970, 2020.
- [88] K. Hviding, L. Juvet e F. A. Vines D, *Colorectal Cancer Screening – Effect on Mortality and Incidence Rate of Colorectal Cancer: Overview of Documentation and International Recommendations [Internet].*, 2008.
- [89] M. Kaminski, J. Regula, E. Kraszewska, M. Polkowski, U. Wojciechowska, J. Didkowska, M. Zwierko, M. Rupinski, M. Nowacki e E. Butruk, «Quality Indicators for Colonoscopy and the Risk of Interval Cancer,» *New England Journal of Medicine*, vol. 362, n. 19, pp. 1795-1803, 2010.
- [90] S.-e. M. Y. M. M. T. K. K. T. I. H. O. M. a. M. K. Kudo, «Artificial intelligence and colonoscopy: Current status and future perspectives,» *Digestive Endoscopy*, vol. 31, pp. 363-371, 2019.
- [91] M. Alagappan, J. Brown, Y. Mori e T. Berzin, «Artificial intelligence in gastrointestinal endoscopy: The future is almost here.,» *World J Gastrointest Endosc.*, vol. 10, n. 10, pp. 239-249, 2018.

- [92] E. Sinagra, M. Badalamenti, M. Maida, M. Spadaccini, R. Maselli, F. Rossi, G. Conoscenti, D. Raimondo, S. Pallio, A. Repici e A. Anderloni, «Use of artificial intelligence in improving adenoma detection rate during colonoscopy: Might both endoscopists and pathologists be further helped.,» *World J Gastroenterol.*, vol. 26, n. 39, pp. 5911-5918, 2020.
- [93] C. Schauer, M. Chieng, M. Wang, M. Neave, S. Watson, M. Van Rijnsoever, R. Walmsley e A. Jafer, «Artificial intelligence improves adenoma detection rate during colonoscopy.,» *N Z Med J.*, vol. 135, n. 1561, pp. 22-30, 2022.
- [94] S. Larsen e Y. Mori, «Artificial intelligence in colonoscopy: A review on the current status.,» *DEN Open.*, vol. 2, n. 1, 2022.
- [95] B. Harshavardhan, B. Nandakumar e e. al, «The role of artificial intelligence based systems for cost optimization in colorectal cancer prevention programs.,» *Front. Artif. Intell.*, vol. 5, 2022.
- [96] A. H. Calderwood e B. C. Jacobson, «Comprehensive validation of the Boston Bowel Preparation Scale.,» *Gastrointestinal Endoscopy*, vol. 72, n. 4, pp. 686-692, 2010.
- [97] «POLIPI: CLASSIFICAZIONE.,» [Online]. Available: <https://polipi-cancro-colon.iannetti.it/polipi/>.
- [98] E. Rondonotti, D. Di Paolo, E. R. Rizzotto, C. Alvisi, E. Buscarini, M. Spadaccini, G. Tamanini, S. Paggi, A. Amato, G. Scardino, S. Romeo, S. Alicante, F. Ancona, E. Guido, V. Marzo, F. Chicco, S. Agazzi, C. Rosa, L. Correale e A. Repici, «Efficacy of a computer-aided detection system in a fecal immunochemical test-based organized colorectal cancer screening program: a randomized controlled trial (AIFIT study),» *Endoscopy*, vol. 54, n. 12, pp. 1171-1179, 2022.
- [99] R. Hüneburg, K. Bucksch, F. Schmeißer, D. Heling, T. Marwitz, S. Aretz e e. al., «Real-time use of artificial intelligence (CADEYE) in colorectal cancer surveillance of patients with Lynch syndrome—A randomized controlled pilot trial (CADLY).,» *United European Gastroenterol J.*, vol. 11, n. 1, pp. 60-68, 2023.
- [100] M. Yousaf, S. Shafqat, R. Gill, A. Khurshed e O. Parkash, «Adenoma detection rate as a quality indicator for colonoscopy: a descriptive cross-sectional study from a tertiary care hospital in Pakistan.,» *Endosc Int Open.*, 2020.
- [101] R. Jover, M. Herráiz, O. Alarcón, E. Brullet, L. Bujanda, M. Bustamante, R. Campo, R. Carreño, A. Castells, J. Cubiella, P. García-Iglesias, A. J. Hervás, P. Menchén, A. Ono, A. Panadés, A. Parra-Blanco, M. Pellisé, M. Ponce e E. Quintero, «Clinical practice guidelines: quality of colonoscopy in colorectal cancer screening.,» *Endoscopy*, vol. 44, n. 4, pp. 444-451, 2012.
- [102] E. J. Nwankwo, J. Lines, S. Trehan, M. Marsh, A. Trehan, K. Banwait, S. Pathapati, S. Misra e I. Obokhare, «Improving Adenoma Detection Rates: The Role of the Fecal Immunochemical Test.,» *Cureus*, vol. 13, n. 4, p. 14382, 2021.

- [103] M. J. C. Sanjuan e R. e. al., «Variation in Colonoscopy Performance Measures According to Procedure Indication,» *Clinical Gastroenterology and Hepatology*, vol. 18, n. 5, pp. 1216-1223, 2020.
- [104] X. He, X. Lv, B. Zhang, X. Ying, C. Hu, X. Zhou e J. Hu, «Adenoma Detection Rate in Average-Risk Population: An Observational Consecutive Retrospective Study,» *Cancer Control.*, vol. 30, 2023.
- [105] A. J. Cross, E. C. Robbins, B. P. Saunders, S. W. Duffy e K. Wooldrage, «Higher Adenoma Detection Rates at Screening Associated With Lower Long-Term Colorectal Cancer Incidence and Mortality,» *Clinical Gastroenterology and Hepatology*, vol. 20, n. 2, pp. 148-167, 2022.
- [106] N. Brant, D. Apple, M. Downs e e. al., «Adenoma Detection Rate on Colonoscopy: an Argument for Focused Screening?,» *J Gastrointest Surg*, n. 26, p. 1954–1956, 2022.
- [107] J. Park, S. Kim, J. Hyun, K. Han, B. Kim, C. Hong, S. Lee e D. Sohn, «Correlation Between Bowel Preparation and the Adenoma Detection Rate in Screening Colonoscopy,» *Ann Coloproctol.*, vol. 33, n. 3, pp. 93-98, 2017.
- [108] J. Kim, S. Kang, H. Moon e e. al., «Impact of Bowel Preparation Quality on Adenoma Identification During Colonoscopy and Optimal Timing of Surveillance.,» *Dig Dis Sci*, vol. 60, pp. 3092-3099, 2015.
- [109] S. Wernly, B. Wernly, G. Semmler, S. Bachmayer, D. Niederseer, F. Stickel, U. Huber-Schönauer, E. Aigner e C. Datz, «A sex-specific propensity-adjusted analysis of colonic adenoma detection rates in a screening cohort.,» *Sci Rep*, vol. 11, n. 1, p. 17785, 2021.
- [110] S. Li, J. Maas, M. Spencer, P. Panagiotakis, D. Brokl e M. Rahman, «Differenze di genere nel tasso di rilevamento dell'adenoma del colon: uno studio comunitario: 2066,» *American Journal of Gastroenterology*, vol. 109, 2014.
- [111] S. G. Coe e e. al., «Assessment of adenoma detection rate benchmarks in women versus men,» *Gastrointestinal Endoscopy*, vol. 77, n. 4, pp. 631-635, 2013.
- [112] A. Gimeno-García, A. Hernández-Pérez, D. Nicolás-Pérez e M. Hernández-Guerra, «Artificial Intelligence Applied to Colonoscopy: Is It Time to Take a Step Forward?,» *Cancers (Basel)*, vol. 15, n. 8, 2023.
- [113] M. Ashat, J. Klair, D. Singh, A. Murali e R. Krishnamoorthi, «Impact of real-time use of artificial intelligence in improving adenoma detection during colonoscopy: A systematic review and meta-analysis,» *Endosc Int Open*, vol. 9, n. 4, pp. 1341-1457, 2021.
- [114] B. Mohan, A. Facciorusso, S. Khan, S. Chandan, L. Kassab, P. Gkolfakis, G. Tziatzios, K. Triantafyllou e D. Adler, «Real-time computer aided colonoscopy versus standard colonoscopy for improving adenoma detection rate: A meta-analysis of randomized-controlled trials.,» *EClinicalMedicine*, pp. 29-30, 2020.

- [115] D. Rex, «Colonoscopic withdrawal technique is associated with adenoma miss rates.,» *Gastrointest Endosc*, vol. 51, n. 1, pp. 33-36, 2000.
- [116] R. Lee, R. Tang, V. Muthusamy, S. Ho, N. Shah, L. Wetzel, A. Bain, E. Mackintosh, C. A. Paek AM, L. Saraf, D. Kalmaz e T. Savides, «Quality of colonoscopy withdrawal technique and variability in adenoma detection rates (with videos),» *Gastrointest Endosc*, vol. 74, n. 1, pp. 128-134, 2011.
- [117] R. I. Barclay e al, «Variation in Adenoma Detection Rates and Colonoscopic Withdrawal Times During Screening Colonoscopy,» *Gastrointestinal Endoscopy*, vol. 61, n. 5, 2005.
- [118] W. M. Sanchez, G. C. M. Harewood e B. Petersen, «Evaluation of Polyp Detection in Relation to Procedure Time of Screening or Surveillance Colonoscopy,» *American Journal of Gastroenterology*, vol. 99, n. 10, pp. 1941-1945, 2004.
- [119] D. Rex, «Colonoscopy withdrawal times and adenoma detection rates,» *Gastroenterol Hepatol (N Y)*, vol. 3, n. 8, pp. 609-610, 2007.
- [120] M. Desai e e. al, «Impact of withdrawal time on adenoma detection rate: results from a prospective multicenter trial,» *Gastrointestinal Endoscopy*, vol. 97, n. 3, pp. 537-543, 2023.
- [121] T. Lee, R. Blanks, C. Rees e e. a. on behalf of Northern Region Endoscopy Group (NREG), «Colonoscopy withdrawal time and adenoma detection rate in screening colonoscopy: the optimum average withdrawal time is 10 min,» *Gut*, vol. 60, 2011.
- [122] Z. Lu, L. Zhang, L. Yao, D. Gong, L. Wu, M. Xia, J. Zhang, W. Zhou, X. Huang, C. He, H. Wu, C. Zhang, X. Li e H. Yu, «Assessment of the Role of Artificial Intelligence in the Association Between Time of Day and Colonoscopy Quality,» *JAMA Netw Open.*, vol. 6, n. 1, 2023.
- [123] R. Richter, J. Bruns, W. Obst, V. Keitel-Anselmino e J. Weigt, «Influence of Artificial Intelligence on the Adenoma Detection Rate throughout the Day.,» *Dig Dis*, vol. 41, n. 4, pp. 615-619, 2023.
- [124] S. R. Jayadev C, «Artificial intelligence in laser refractive surgery - Potential and promise!,» *Indian J Ophthalmol.*, vol. 68, pp. 2650-2651, 2020.
- [125] A. Eckardt, J. Kheder, A. Basil, T. Silverstein, K. Patel, M. Mahmoud, Y. Al-Azzawi, D. Ellis, W. Gillespie, Y. Carrasquillo Vega, S. Person e J. Levey, «Trainee participation during screening colonoscopy does not affect ADR at subsequent surveillance, but may result in early follow-up.,» *Endosc Int Open*, vol. 8, n. 12, pp. 1732-1740, 2020.
- [126] A. Do, J. Weinberg, A. Kakkar e B. C. Jacobson, «Reliability of adenoma detection rate is based on procedural volume,» *Gastrointestinal Endoscopy*, vol. 77, n. 3, pp. 376-380, 2013.
- [127] J. Marsano, S. Johnson, S. Yan, L. Alli-Akintade, M. Wilson, A. Al-Juburi, J. Stondell e S. Tejaswi, «Comparison of colon adenoma detection rates using cap-assisted and

Endocuff-assisted colonoscopy: a randomized controlled trial.,» *Endosc Int Open*,  
vol. 7, n. 12, pp. 1585-1591, 2019.