

55.25% 67.55% 41.00%

10 GIUGNO 2025

L'epidemiologia dei tumori
tra nuovi bisogni di salute, ricerca
clinica e innovazioni tecnologiche



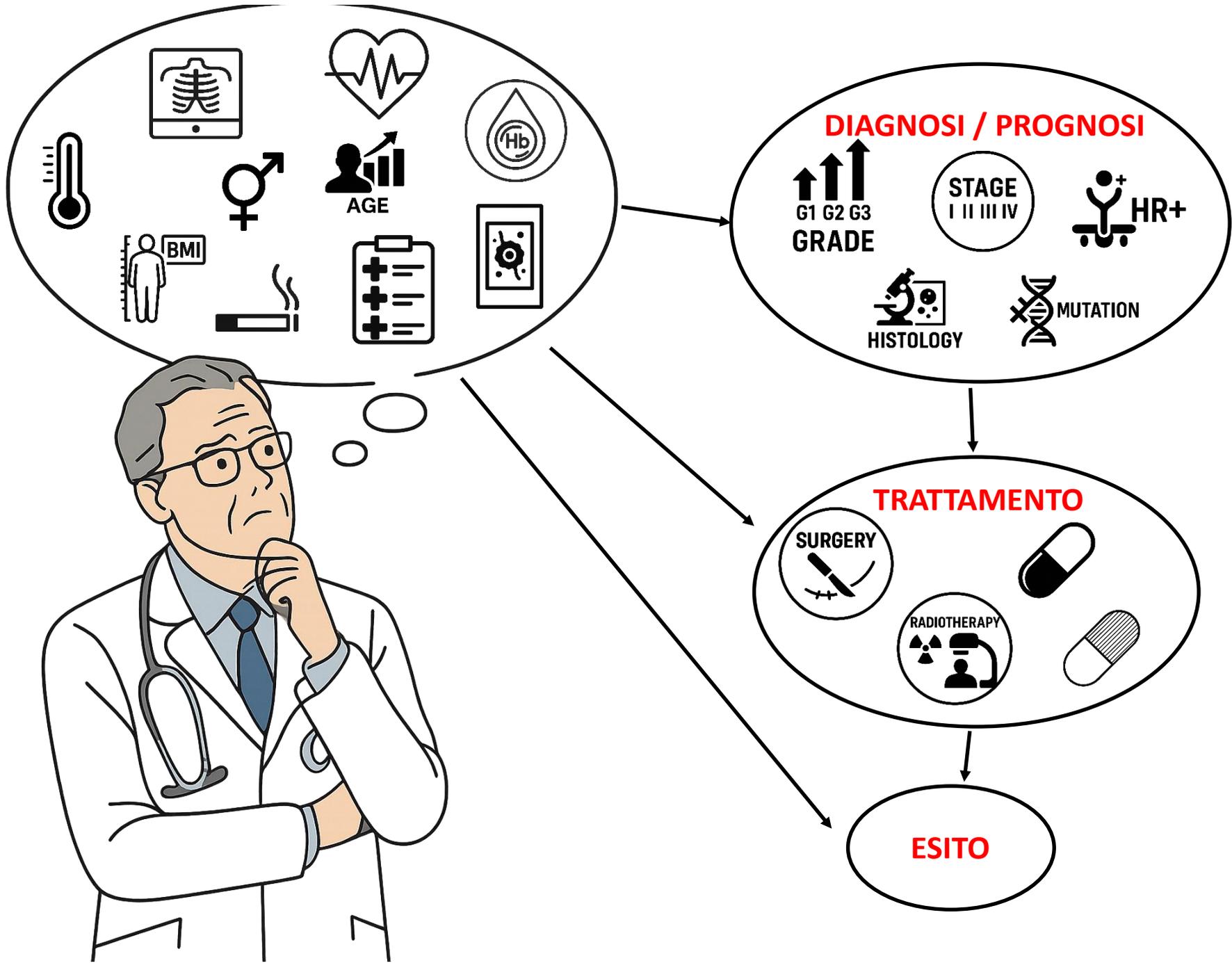
IA in epidemiologia clinica: nuove possibilità, nuove domande, alcuni limiti

Andrea Evangelista

Unità di Epidemiologia Clinica e Valutativa

AOU Città della Salute e della Scienza di Torino e CPO Piemonte

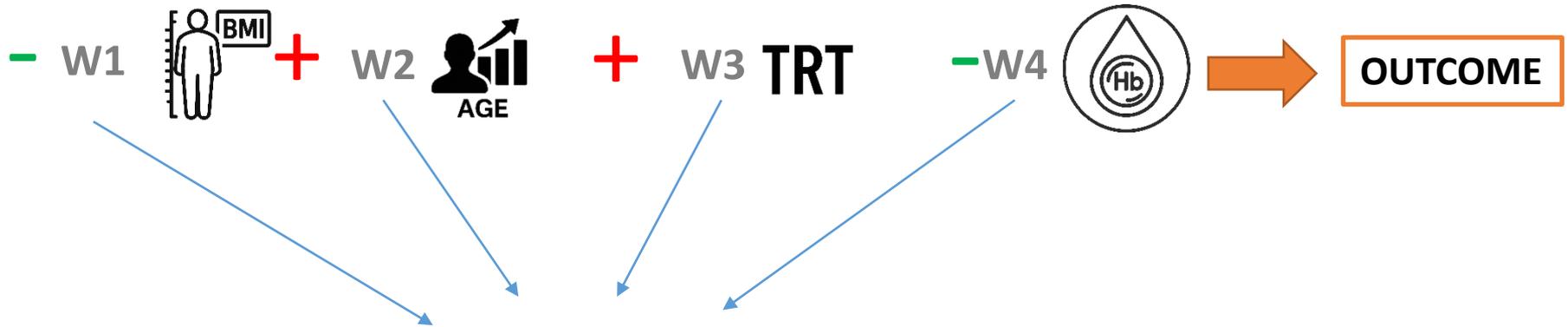




Valutazione Clinica



Valutazione Clinica

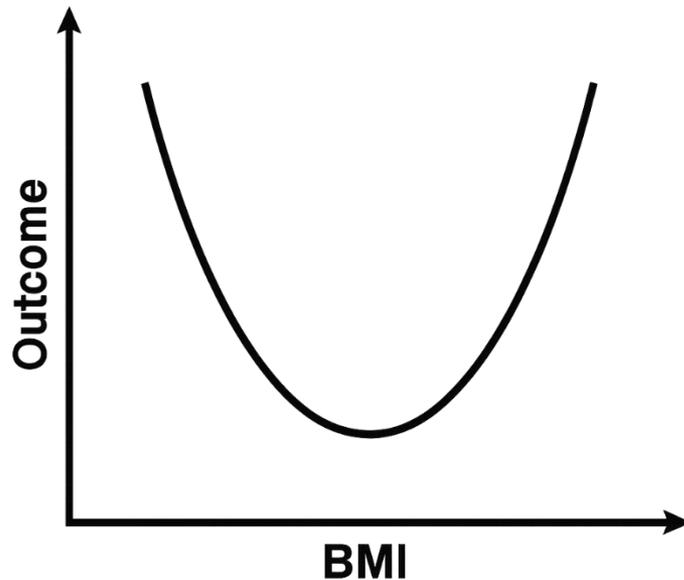


Coefficienti \approx Pesi

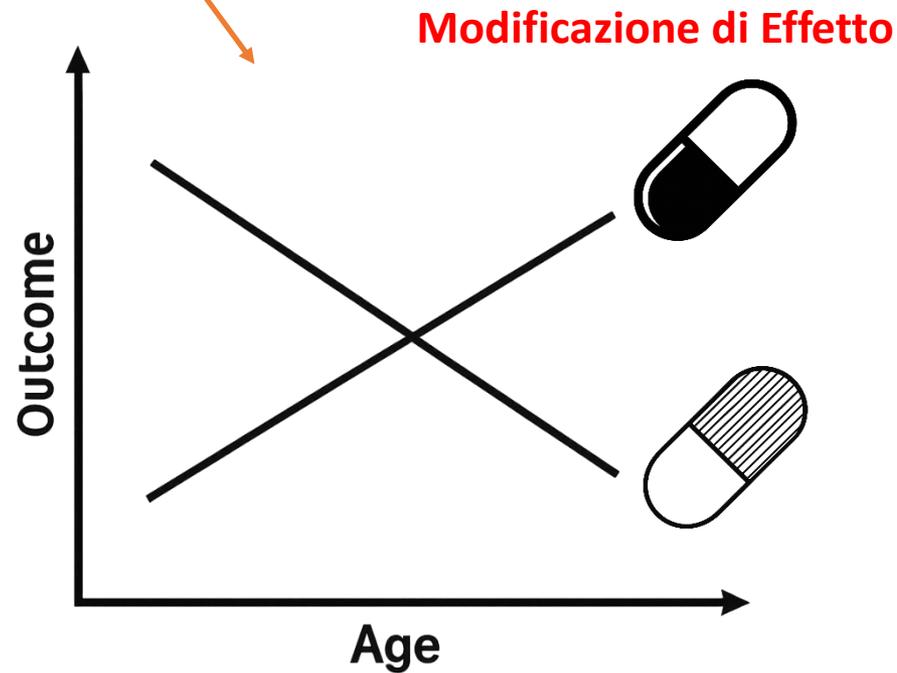
Modello Statistico



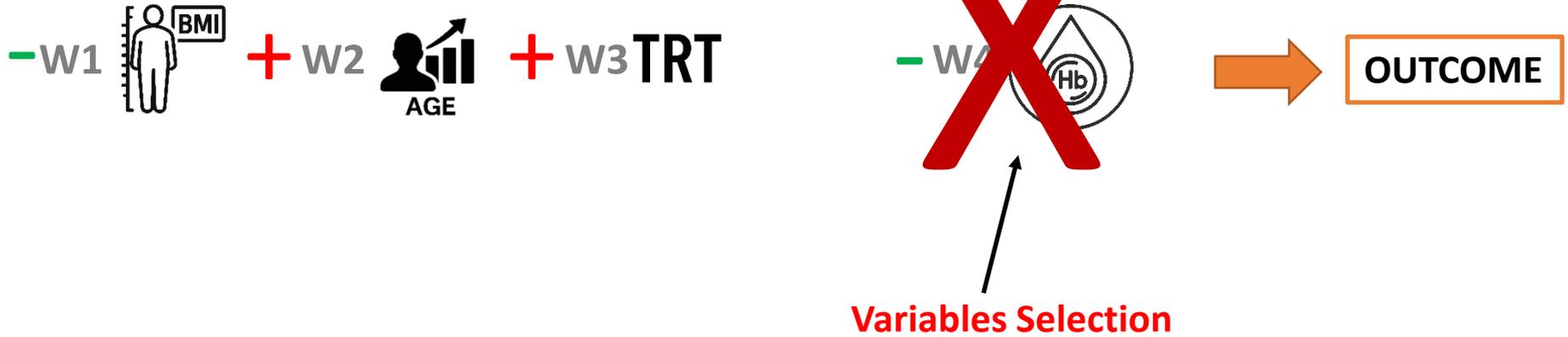
Effetti Non Lineari



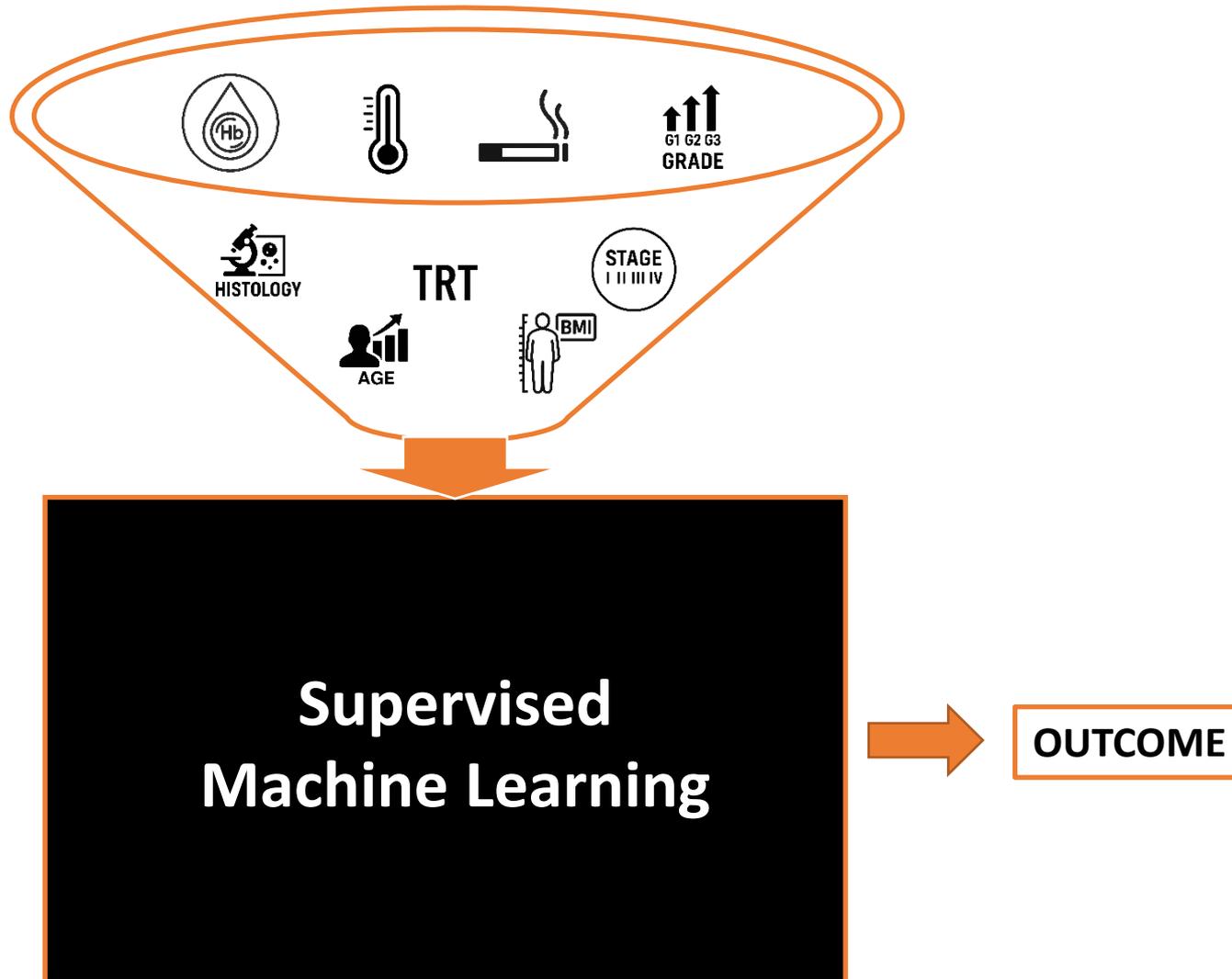
Modello Statistico



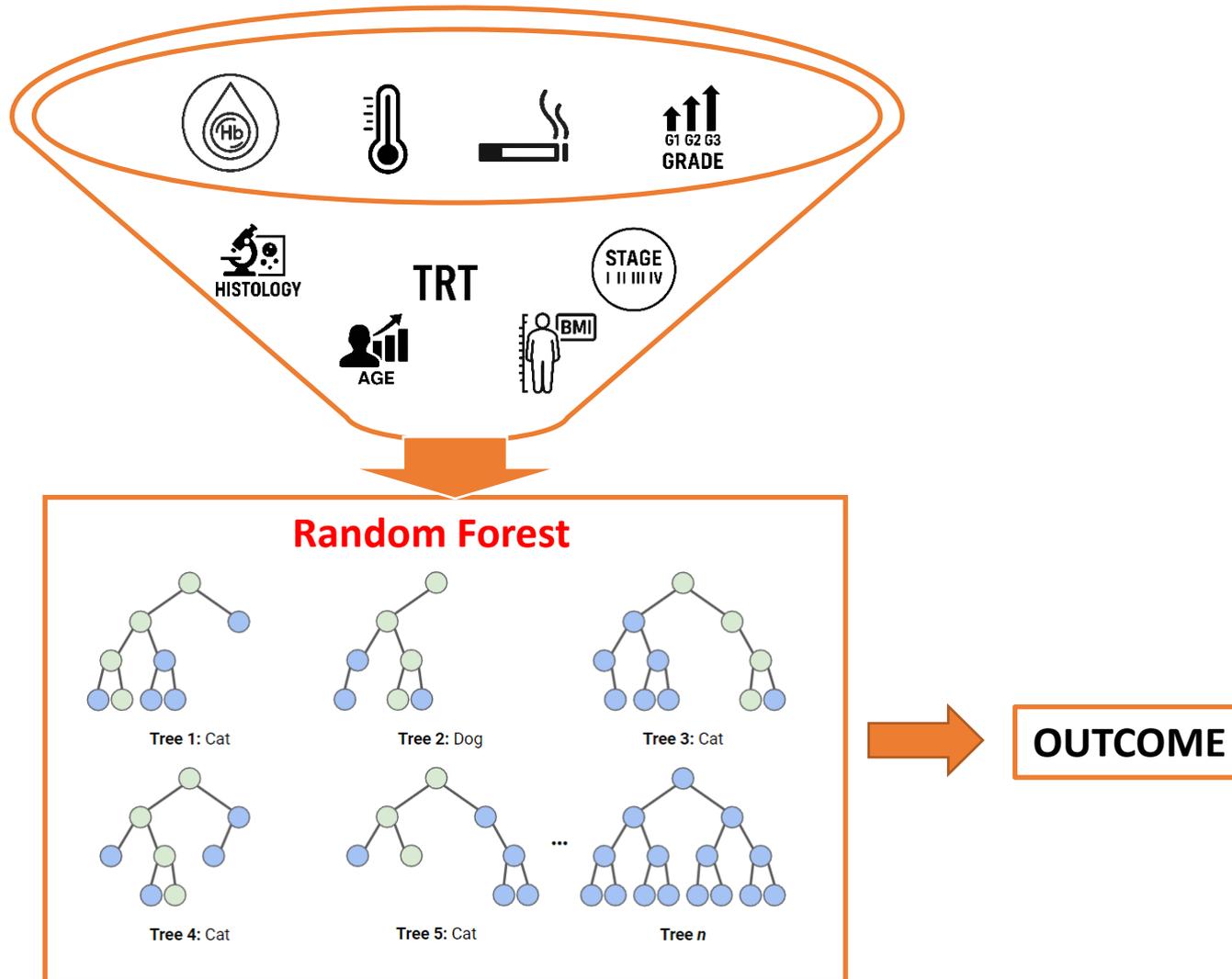
Modello Statistico



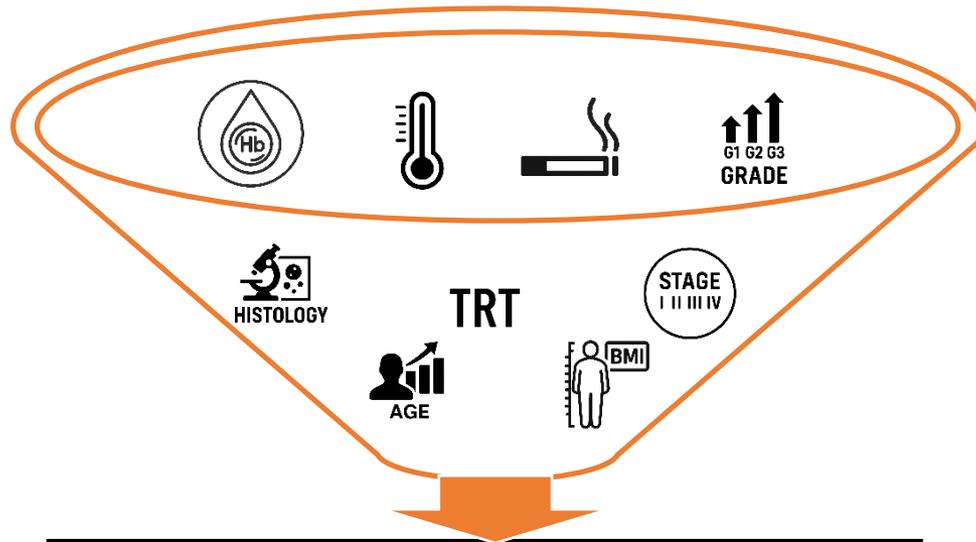
Machine Learning



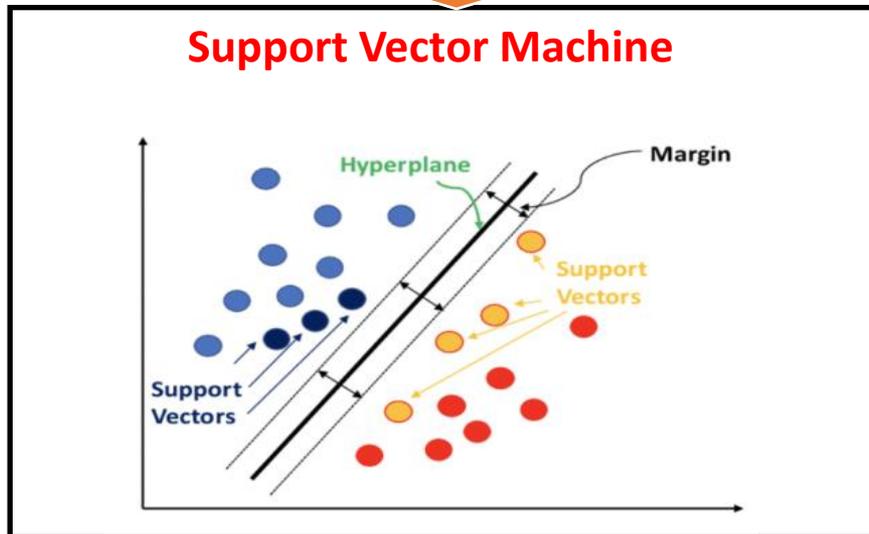
Machine Learning



Machine Learning

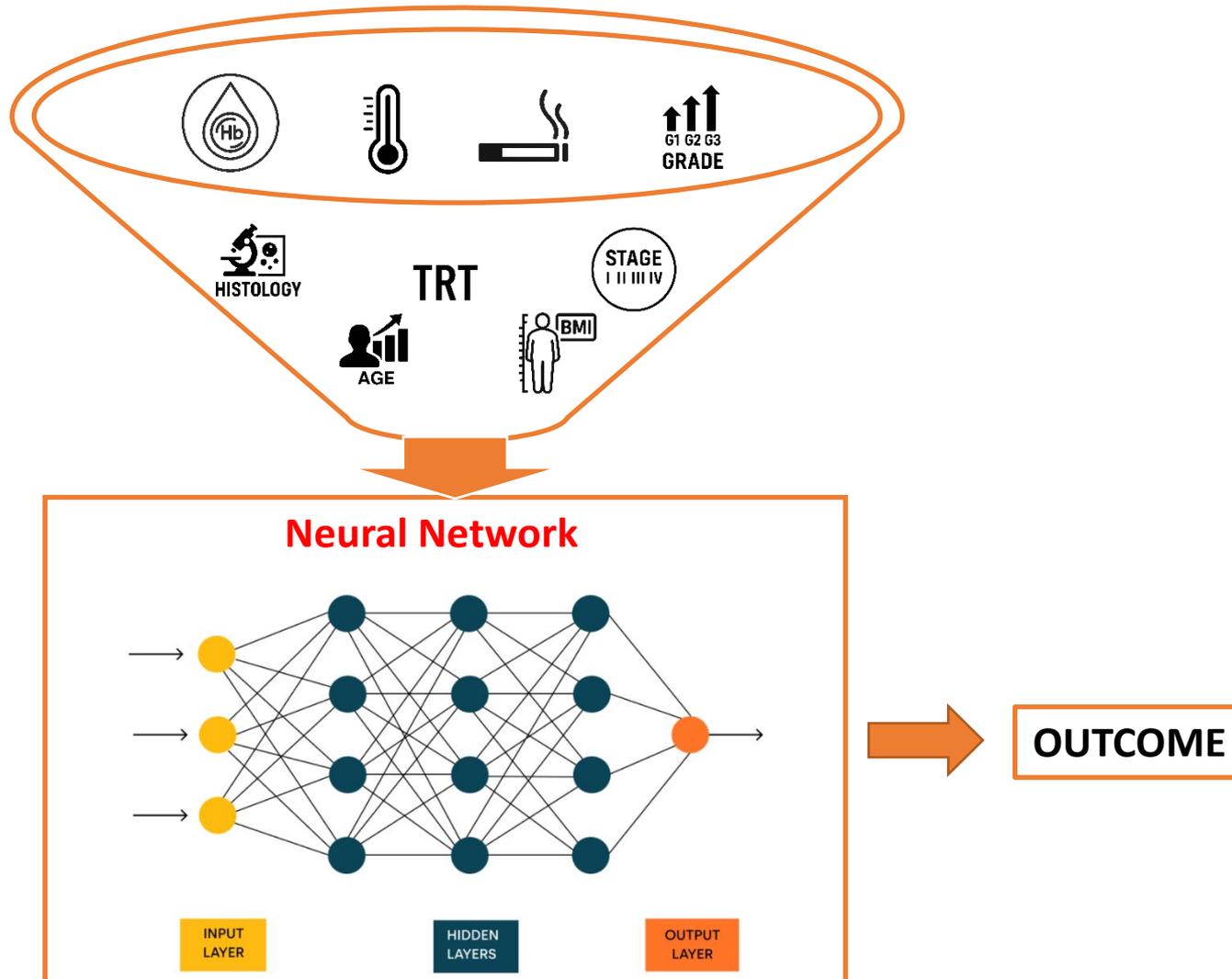


Support Vector Machine



OUTCOME

Machine Learning



Vantaggi e Svantaggi del Machine Learning



Vantaggi

Gestione di dati complessi

Ottimo con dati eterogenei e non strutturati (es. immagini, testo)

Flessibilità del modello

Può modellare relazioni non lineari e interazioni complesse

Assunzioni sui dati

Non richiede assunzioni forti (es. normalità, linearità)

Svantaggi

Interpretabilità

Spesso considerato una “black box”

Velocità di calcolo

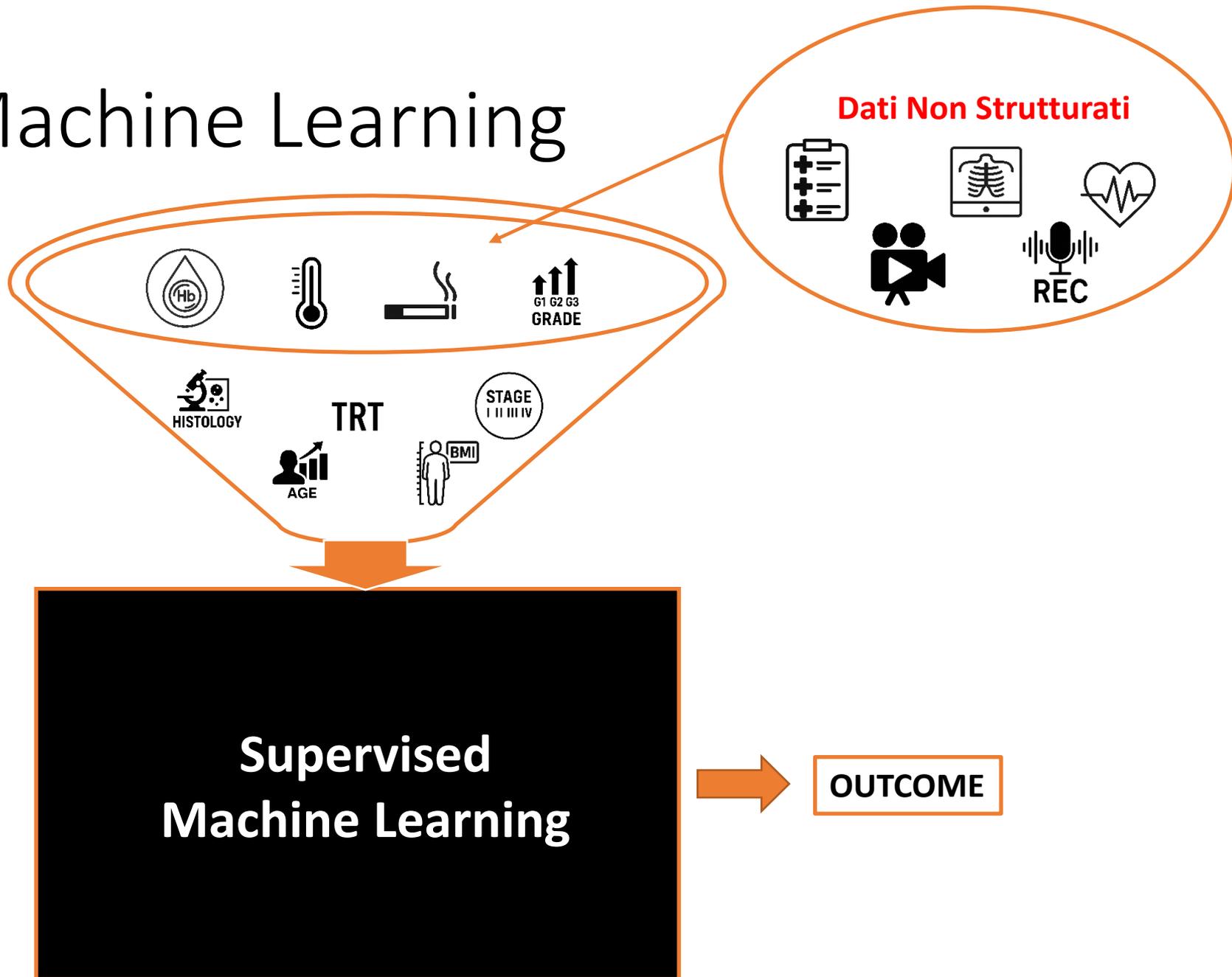
Può essere computazionalmente molto oneroso

Dimensioni dataset

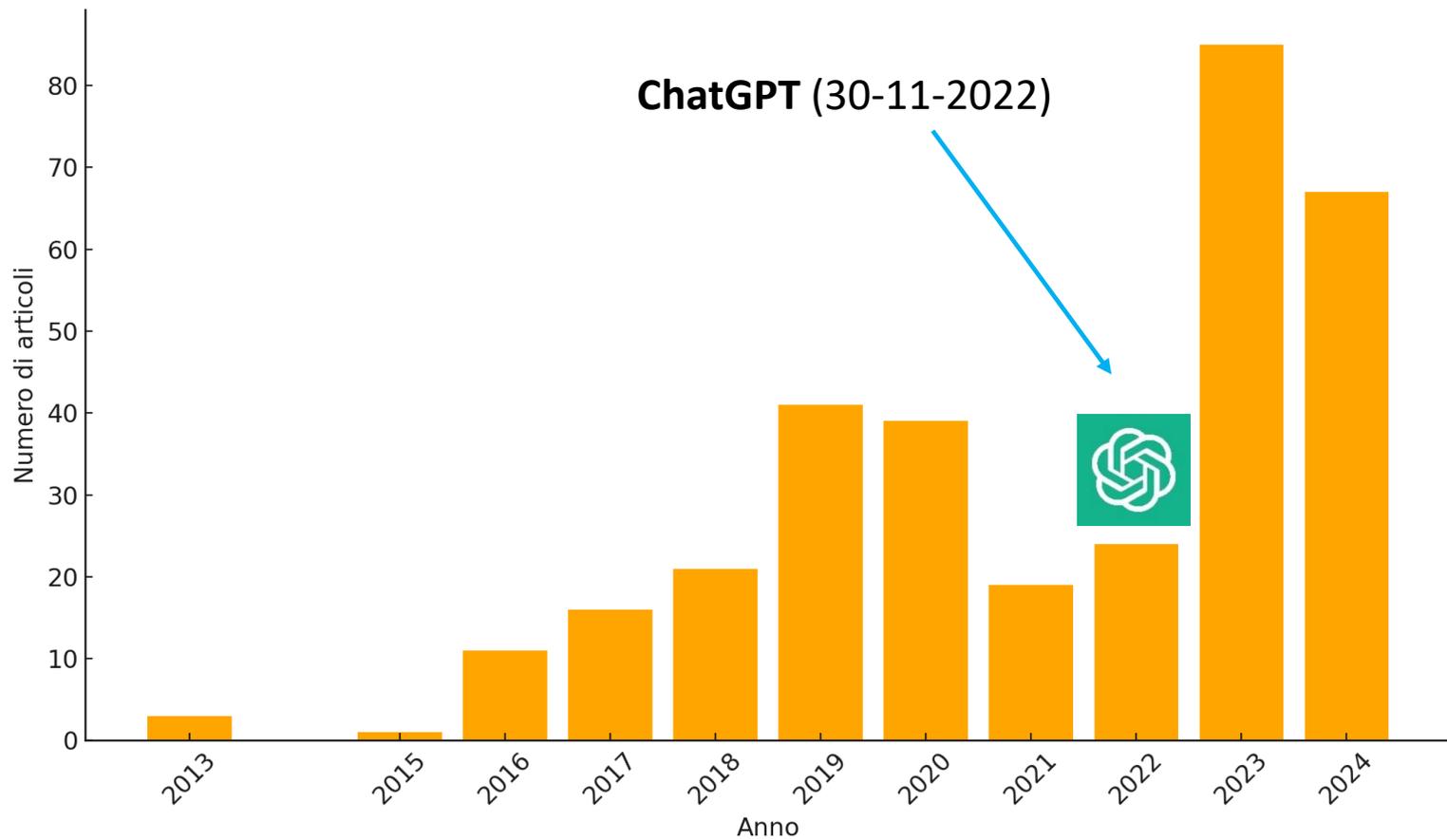
Richiede molti dati per garantire prestazioni affidabili



Machine Learning



Articoli su *Artificial Intelligence* pubblicati su JAMA, NEJM e Lancet (2013-2024)



AI in epidemiologia clinica: ambiti di ricerca

Modelli predittivi

- Modelli **AI** per **prevedere** il rischio di eventi clinici **futuri**.
- *Esempio:* AI per prevedere il rischio di riammissione ospedaliera.

Diagnosi

- **AI** usata per **diagnosticare** condizioni cliniche, spesso confrontata con i medici.
- *Esempio:* AI per rilevare tumori cutanei da immagini.

Validazione

- Test di modelli **AI** su dati **esterni** per valutarne la **generalizzabilità**.
- *Esempio:* AI sviluppata a livello universitario testata in ospedali periferici.

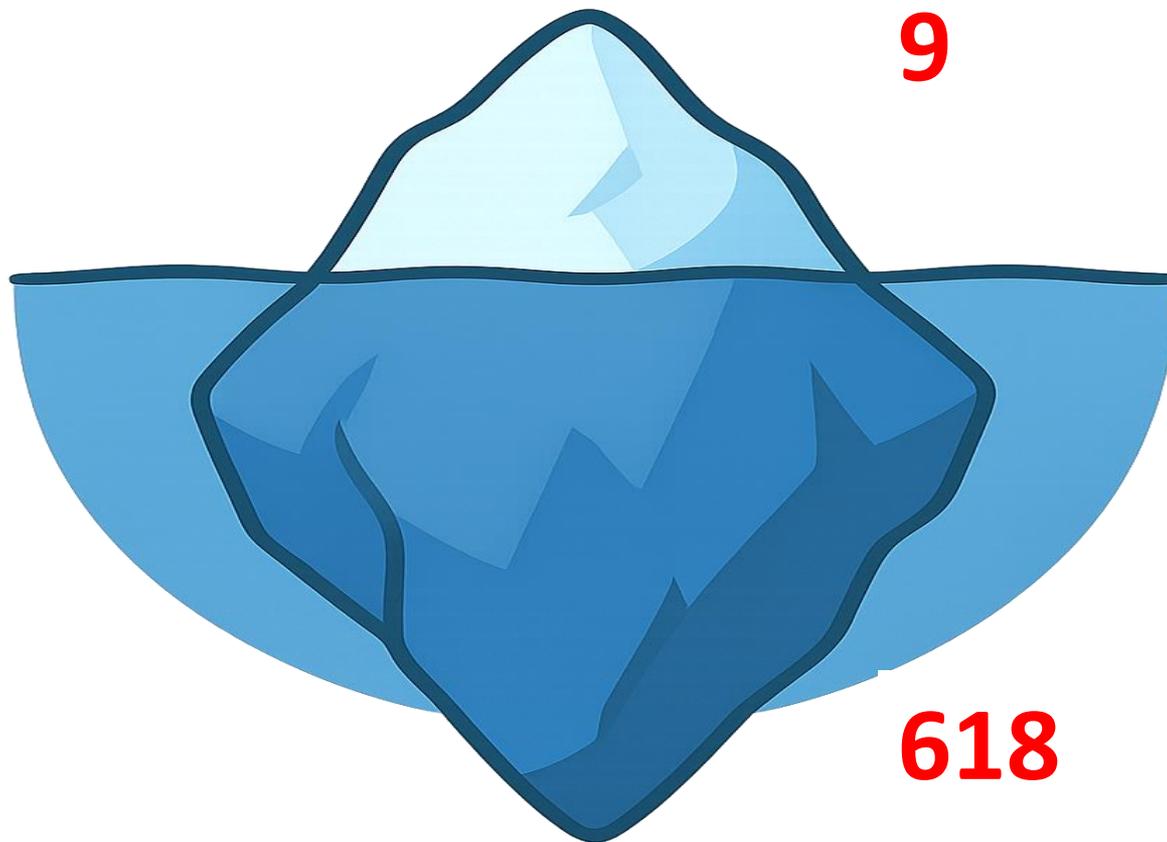
Valutazioni di efficacia

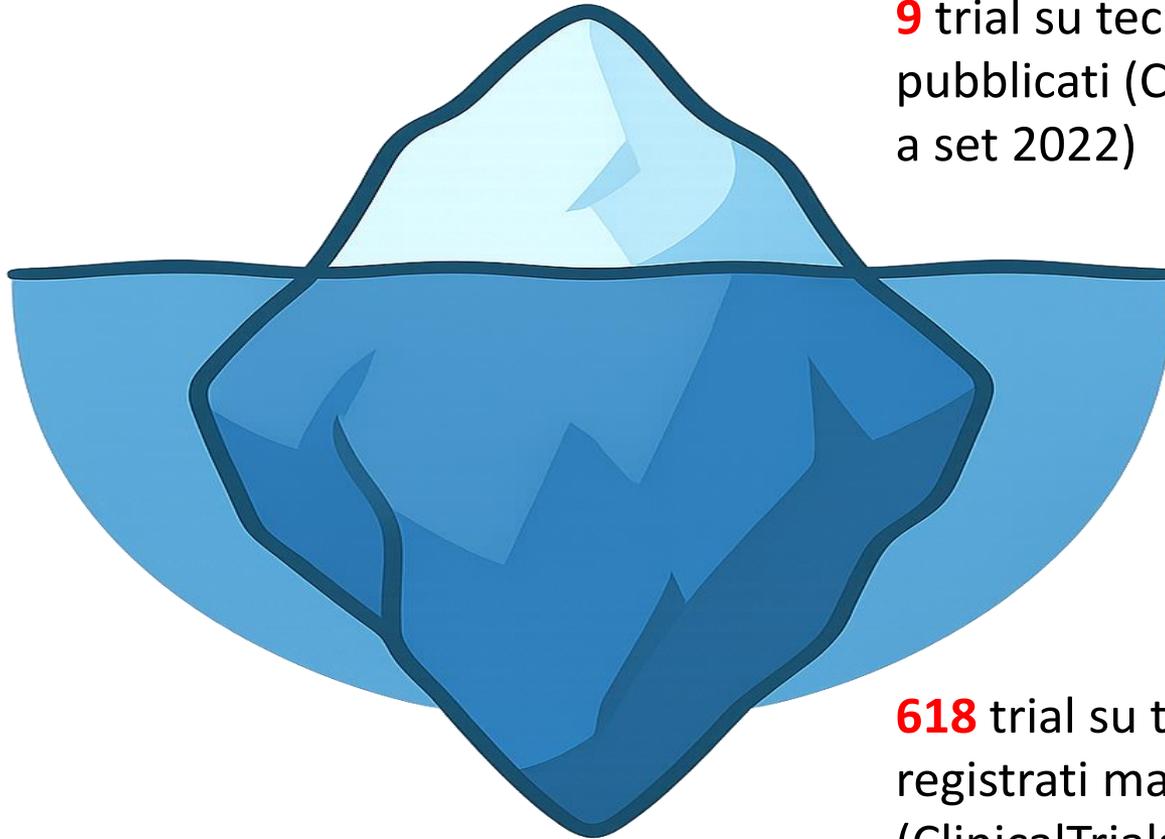
- Studi che valutano l'**impatto** dell'**AI** su esiti e decisioni cliniche.
- *Esempio:* RCT sull'efficacia dell'AI nello screening del tumore colon-rettale.

TRIPOD - AI

SPIRIT - AI
CONSORT - AI







9 trial su tecnologie AI — risultati pubblicati (ClinicalTrials.gov, fino a set 2022)

618 trial su tecnologie AI — registrati ma non pubblicati (ClinicalTrials.gov, fino a set 2022)

	Statistically significant improvement	No statistically significant effect	Showed non-inferiority	Statistically significant deterioration	Total
Care management	15	1	2	..	18
Clinical decision making	6	1	7
Diagnostic yield or performance	34	10	1	1	46
Patient behaviour and symptoms	10	3	2	..	15
Total	65	15	5	1	86

Data are n.

Table 1: Primary endpoints and types for randomised controlled trials of artificial intelligence in clinical practice

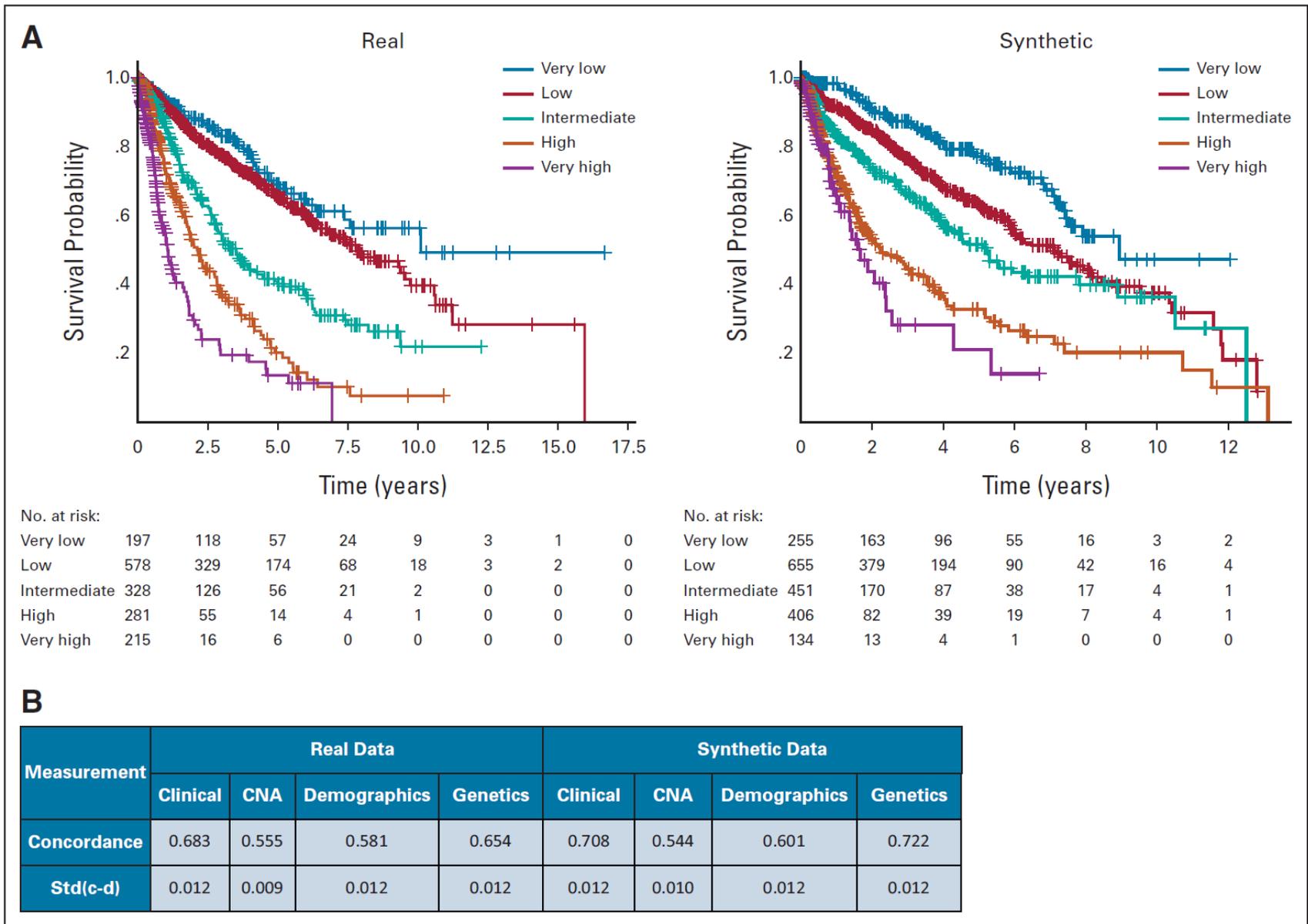
Han R. *et al.* Randomised controlled trials on AI in clinical practice. *Lancet Digit Health* 2024; 6: e367–73.

	Statistically significant improvement	No statistically significant effect	Showed non-inferiority	Statistically significant deterioration	Total
AI vs clinician	3	1	3	1	8
AI vs routine care	16	4	20
AI-assisted clinician vs unassisted clinician	46	10	2	..	58
Total	65	15	5	1	86

Data are n. AI=artificial intelligence.

Table 2: Primary endpoint results and group comparisons for randomised controlled trials of AI in clinical practice

Han R. *et al.* Randomised controlled trials on AI in clinical practice. *Lancet Digit Health* 2024; 6: e367–73.



Possibile impiego dei dati sintetici negli studi di fase II



- Potenziali Vantaggi dei Dati Sintetici
 - **Creazione di un Braccio di Controllo:** I dati sintetici possono essere utilizzati per costruire un braccio di controllo che rispecchi le caratteristiche del gruppo sperimentale
 - **Miglioramento dell'Efficienza:** Potrebbero consentire una valutazione preliminare rapida e possono ridurre i costi e i tempi dello studio.
- Sfide e Limiti
 - **Problema del Drift:** I dati sintetici non risolvono il problema del drift (variazione nel tempo delle caratteristiche della popolazione o del trattamento), che può influire sulla comparabilità dei risultati.

Studio di fase II a braccio singolo: il problema del drifting

- **Drift:** Cambiamenti gradualmente e non intenzionali nelle caratteristiche dello studio nel tempo, dovuti a variazioni nel trattamento, nella popolazione o nelle valutazioni degli esiti.
 - **Drift del trattamento:** Cambiamenti nelle terapie di supporto o miglioramenti nelle cure che alterano i risultati indipendentemente dal trattamento sperimentale.
 - **Drift nella popolazione dei pazienti:** Selezione di pazienti con caratteristiche prognostiche diverse rispetto agli studi storici.
 - **Drift nelle valutazioni degli esiti:** Miglioramenti nelle tecniche diagnostiche (come imaging) o delle tempistiche dei controlli che possono sovrastimare o sottostimare gli effetti del trattamento.

Best Practices (?) per l'Uso Trasparente dei Dati Sintetici

- **Identificare la Base Dati:** Definire chiaramente la base dati di origine per la generazione dei dati sintetici, assicurando consistenza e **trasparenza**.
- **Generazione Pre-Arruolamento:** I dati sintetici dovrebbero essere generati prima dell'arruolamento dei pazienti per evitare **aggiustamenti retroattivi** che potrebbero introdurre bias.
- **Repository Pubblico:** Idealmente, i dati sintetici e i dettagli delle tecniche utilizzate dovrebbero essere archiviati in un repository pubblico accessibile agli stakeholder, promuovendo trasparenza e **riproducibilità**.
- **Esplicitare le Tecniche di Analisi:** Documentare chiaramente le tecniche e gli algoritmi impiegati per creare e analizzare i dati sintetici prima dell'inizio dello studio.