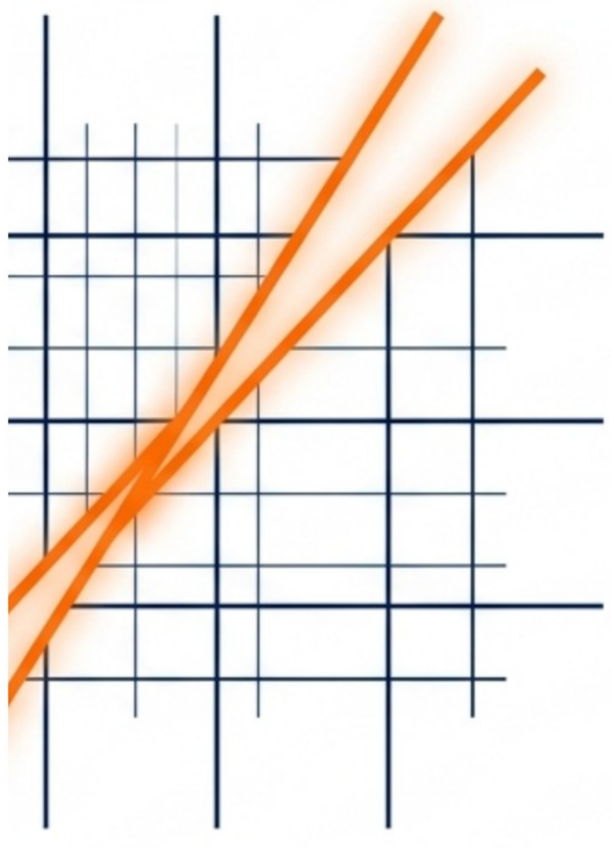
INNOVATIVITÀ TERAPEUTICA
AIFA 2007 vs 2025

Due criteri, la stessa parola.

L'algoritmo AIFA del 2007 applicato retrospettivamente alle valutazioni prodotte con i criteri vigenti. Un esperimento su 23 schede di innovatività.

23 schede · 2 criteri · 1 domanda



I totali sono vicini. Quasi lo stesso numero?

A livello macro, il cambio di algoritmo sembra aver spostato poco.

2007

10

4 Importante
+ 6 Moderata

2025

8

Innovativi

LA DOMANDA: Stiamo contando la stessa idea di innovazione?

Lo stesso materiale, due criteri.

Un esercizio controfattuale su 26 schede AIFA. L'AI ha agito come valutatore costante, con regole fissate a priori: ogni assegnazione doveva essere verificabile risalendo alla frase e alla pagina della scheda.



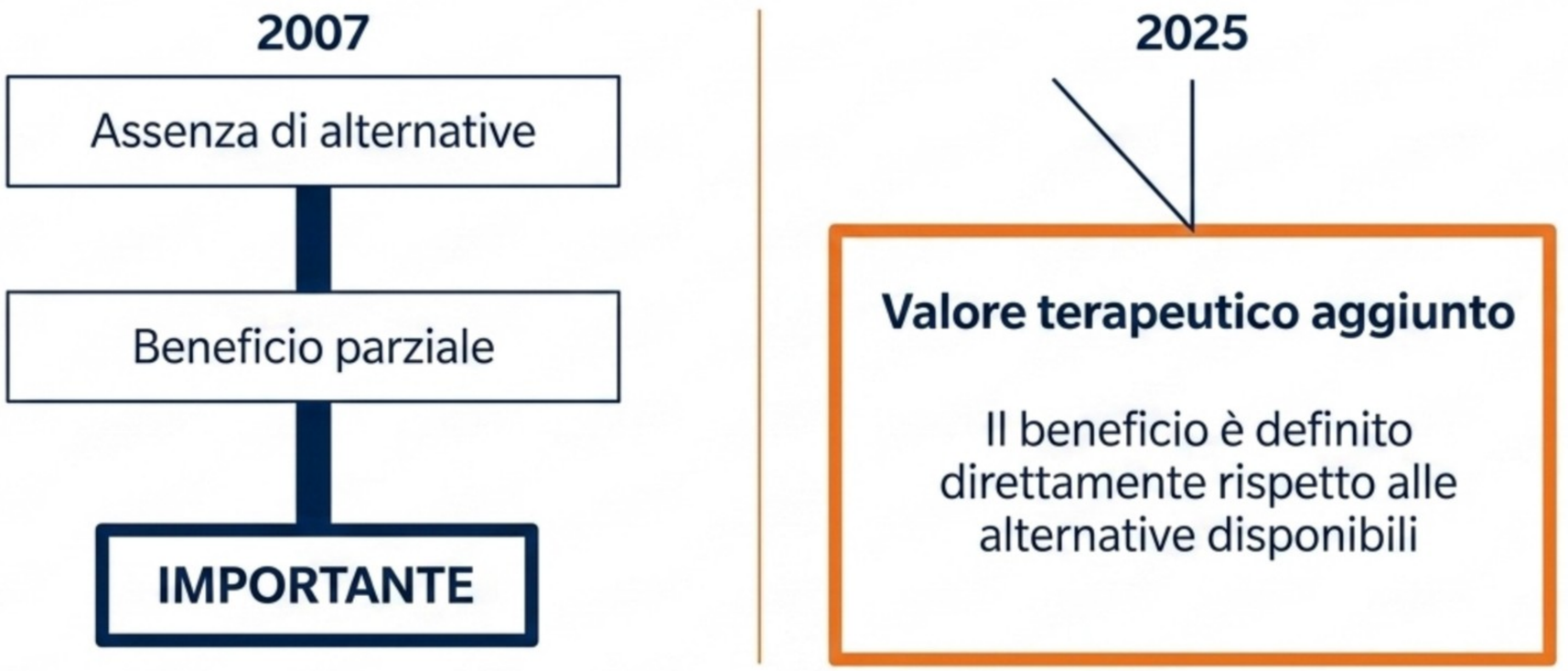
Ma guardando quali farmaci vengono selezionati, la storia cambia.

7 casi su 20 sono completamente discordanti. La concordanza complessiva è solo del 65%.

2007 \ 2025	Innovativo	Non Innovativo
Innovativo	5 Entrambi	5 Solo 2007
Non Innovativo	2 Solo 2025	8 Entrambi no

La differenza nasce prima del verdetto.

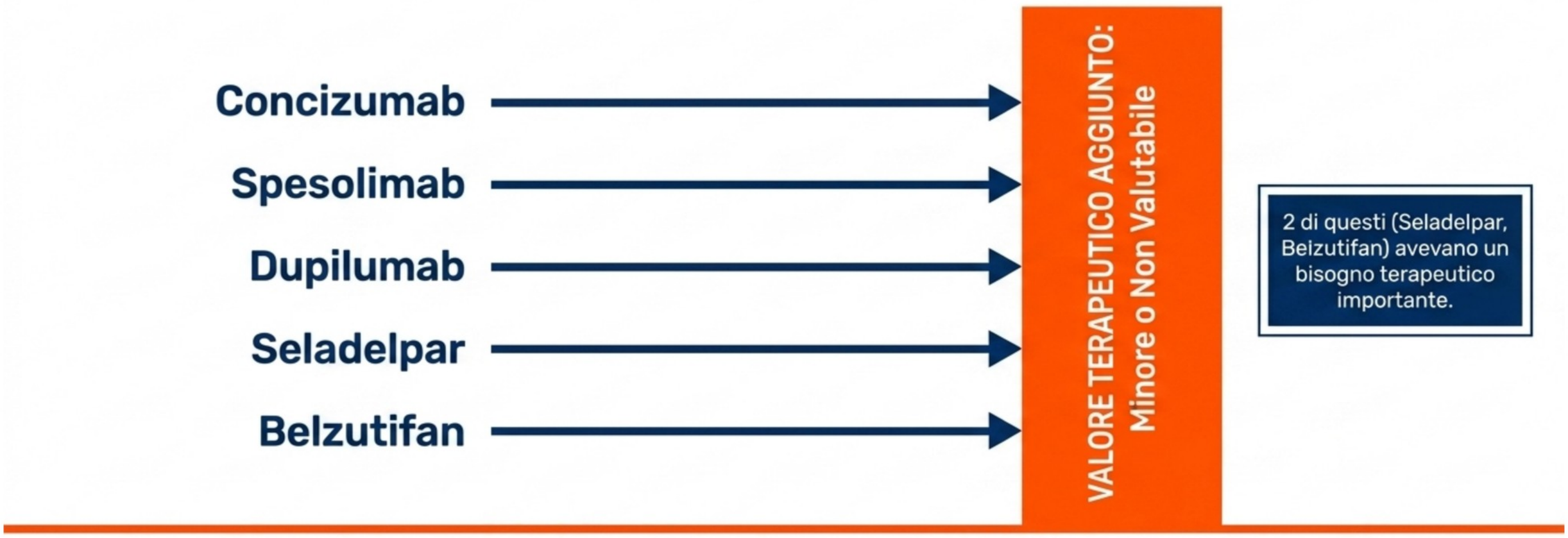
I due framework collocano il confronto con le alternative in punti diversi del ragionamento.
Quindi non cambia soltanto la soglia, cambia cosa decidiamo di far pesare di più.



Nel 2007 l'assenza di alternative può pesare più dell'entità del beneficio.

La discordanza lascia una traccia regolare.

Nei 5 farmaci premiati solo dal 2007, il valore terapeutico aggiunto nel 2025 è sempre giudicato minore o non valutabile.



Spesolimab: la differenza diventa concreta.

Stesso caso clinico, ma il **confronto** viene costruito in modo radicalmente diverso a causa del framework.

Caso: Spesolimab (prevenzione riacutizzazioni di GPP)

Logica 2007

Nessun farmaco autorizzato

+

Beneficio parziale

↓

IMPORTANTE

Logica 2025

Esistono opzioni off label (bisogno moderato)

+

Manca un confronto quantificabile

↓

VALORE AGGIUNTO NON VALUTABILE

↓

NON INNOVATIVO

La divergenza non dipende dai dati clinici, ma da come il criterio costruisce il comparatore.

Punti di forza e limiti dell'esercizio

L'obiettivo non è sostituire una commissione, ma isolare l'effetto del criterio.

PUNTI DI FORZA	LIMITI
✓ Stesse valutazioni AIFA, criterio diverso	! Analisi esplorativa su un corpus limitato
✓ Stesso valutatore AI già istruito sugli esercizi AIFA	! L'AI non riproduce il confronto collegiale della CSE
✓ Ogni assegnazione tracciabile a frase e pagina	! Tre schede non classificabili con il 2007
✓ Regole fissate prima di osservare i risultati	! Kappa incerto: l'IC 95% comprende lo zero (κ = 0,30)

La domanda che resta.

È un'analisi esplorativa, su pochi casi. Non dice quale criterio sia migliore.
Ma una cosa emerge chiaramente: i numeri complessivi sono vicini (10 vs 8).
Le innovazioni selezionate, molto meno.

Quando cambia il metodo, cambia soltanto la soglia dell'innovazione o anche ciò che scegliamo di premiare?

10 vs 8	7 / 20	5
totali vicini	esiti discordanti	innovativi con entrambi

DiSFarM Dipartimento di Scienze Farmaceutiche