

XX Congresso della Società GITMO

RIUNIONE NAZIONALE GITMO

ROMA, ERGIFE PALACE HOTEL, 7-8 MAGGIO 2026

Valutazione della severità e della risposta nella cGVHD:
il ruolo delle piattaforme web

Giorgia Mancini

Clinica di Ematologia

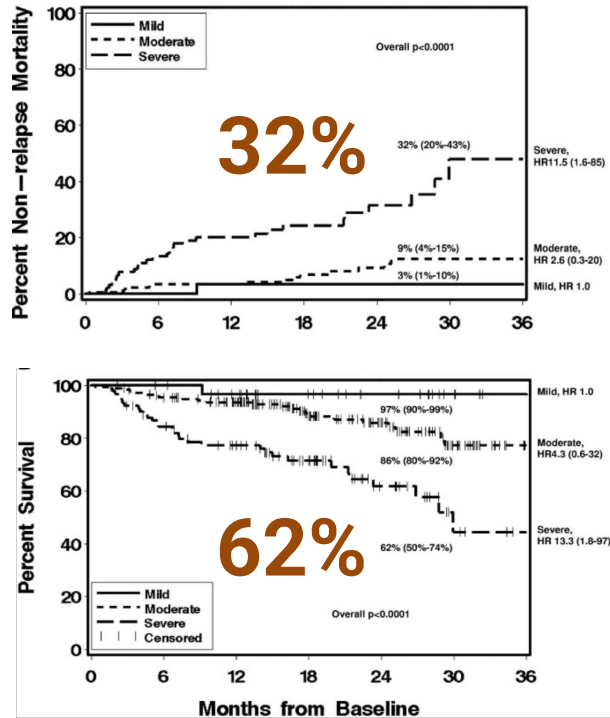
AOU delle Marche

Disclosures of Giorgia Mancini

Company name	Research support	Employee	Consultant	Stockholder	Speakers bureau	Advisory board	Other
Sanofi						X	

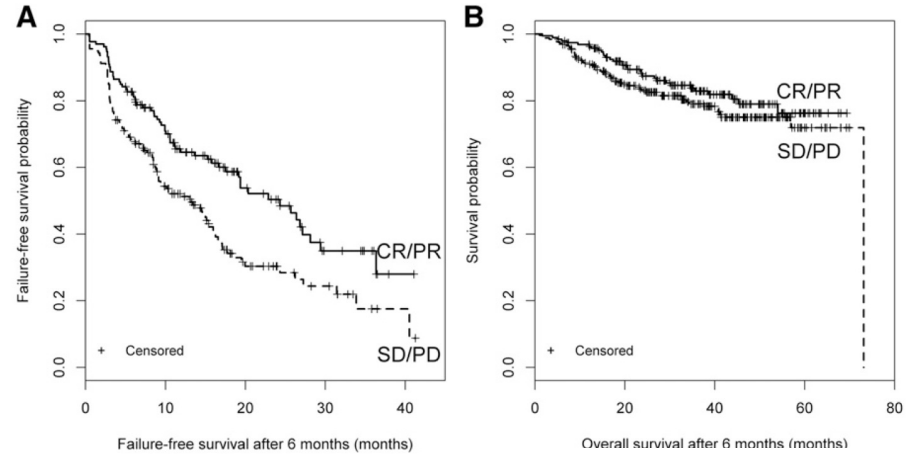
NIH Global Severity Score, Response Evaluation & Outcome

La severità al basale predice
NRM e OS

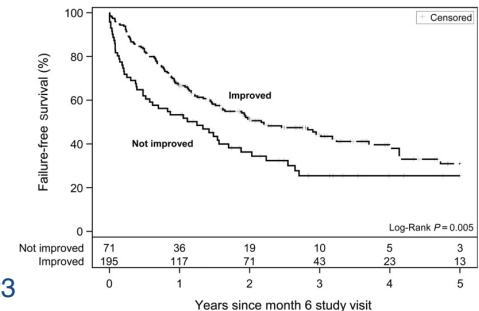


Arai et al., Blood, 2011

La risposta a 6 mesi predice FFS



Palmer et al. Blood, 2016



In et al. Cell Therapy & Immunotherapy 2023

EVOLUZIONE DEI CRITERI NIH: DALLA DESCRIZIONE ALLA MISURAZIONE

(Chronic Graft-versus-Host Disease)

NIH CRITERIA 2005

DESCRITTIVI (COSA È)

1 DEFINIZIONE CHIARE

- Categorie chiare (classica/overlap)
- Distinzione acuta/cronica

2 PUNTEGGIO DI SEVERITÀ GLOBALE (MILD, MODERATE, SEVERE)

- Schema d'organo semplice
- Risultato qualitativo (None/Final)

3 VALUTAZIONE D'ORGANO INIZIALE



Skin
Body Surface Area (BSA) descrittivo

Lung
Sintomi soggettivi

Eye
Sintomi soggettivi

Joint

2005

NIH CRITERIA 2014

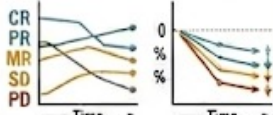
MISURABILI (QUANTO È)

1 AFFINAMENTO E SPECIFICITÀ

- Danno attivo/passato
- Criteri specifici (bocca, genitali, ecc.)

2 CRITERI DI RISPOSTA QUANTITATIVI

RISPOSTA AL TRATTAMENTO



Misurazione dei cambiamenti per valutare l'efficacia terapeutica

3 STRUMENTI DI MISURAZIONE STANDARDIZZATI



Updated NIH Skin Score (0-3)

% FEV₁ per Polmone

Modified Oral Mucosa Score (0-12)

NIH Joint and Fascia Score

P-ROM

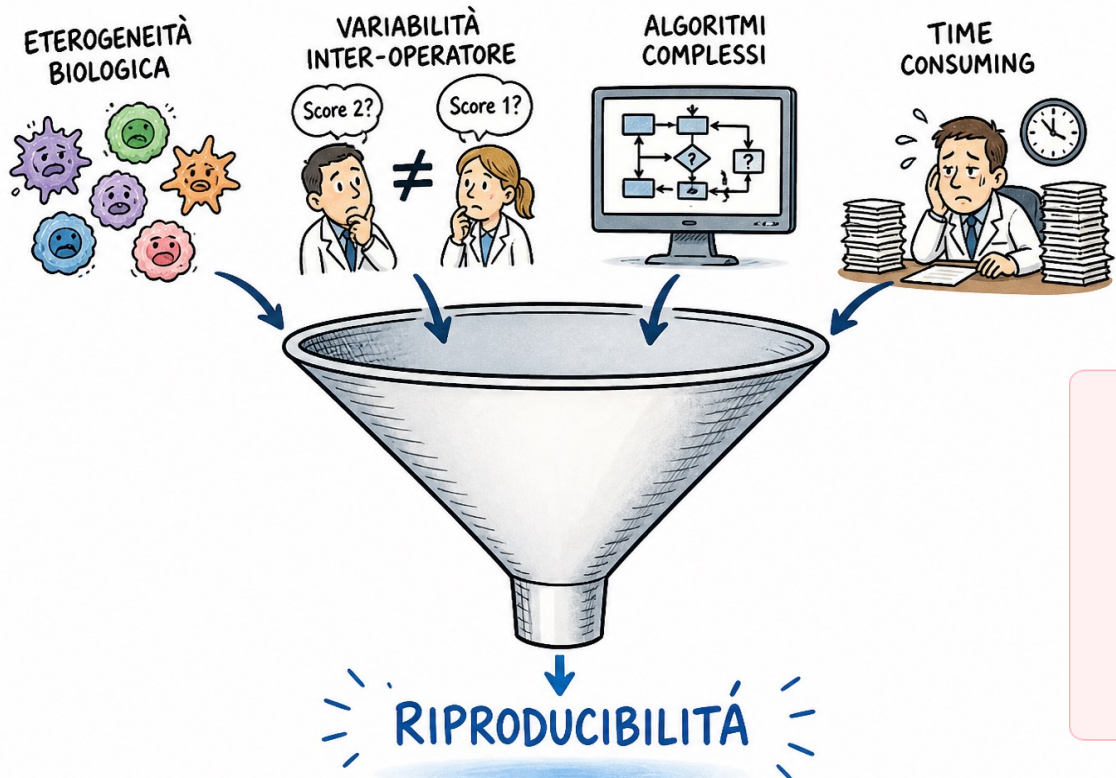
PARADIGMA SHIFT

MISURAZIONE DEI CAMBIAMENTI

Focus: da semplice diagnosi a misurazione del carico globale e risposta.

REAL WORLD

FROM COMPLEXITY TO REPRODUCIBILITY



Barriera Percepita	Gruppo Americano (%)	Gruppo EBMT (%)
Mancanza di tempo per lo scoring NIH	49%	31%
Mancanza di familiarità con i criteri	40%	23%
Accesso limitato agli specialisti	19%	37%
Mancanza di evidenze su outcome pazienti	38%	34%

Duarte et al., BMT, 2014



Barriere Diagnostiche

- ✓ 57% si affida solo alla propria conoscenza
- ✓ Vincoli di tempo stringenti in corsia
- ✓ Linee guida NIH complesse

38%

Tasso di Misclassificazione

4 paz/10: Errata Gravità o Stadiazione

“La misclassificazione del ~38% è un segnale d'allarme sulla sostenibilità clinica degli attuali criteri di ricerca applicati alla vita reale”.



- Scarsa comparabilità dei dati
- Debolezza end point clinici
- Scarsa riproducibilità degli outcome

POSSIBILI SOLUZIONI



Standardizzazione Digital

Implementazione di strumenti di supporto alle decisioni cliniche e app per automatizzare lo scoring NIH.



Formazione Mirata

Programmi di training intensivo per ridurre la variabilità inter-osservatore tra i diversi centri.



Biomarcatori

Integrazione di criteri oggettivi biologici per validare lo scoring clinico e ridurre la misclassificazione.



DIGITAL SOLUTION: eGVHD App www.uzleuven.be/egvhd

Applicazione mobile e web-based progettata per migliorare l'accuratezza e la standardizzazione della diagnosi di GvHD.

Sviluppo EBMT

E-tool free open-source web application

Accesso: website or a mobile application

Focus: Stadiazione/Scoring aGVHD e cGVHD
(MAGIC/NIH 2014)

Uso: bed-side

Validazione: Studio di validazione
Marchio CE

Esportazione dei report in PDF



Accuratezza
diagnostica (OR 6.14)
e di scoring (6.29)

The eGVHD App has the potential to improve the accuracy of graft-versus-host disease assessment: a multicenter randomized controlled trial

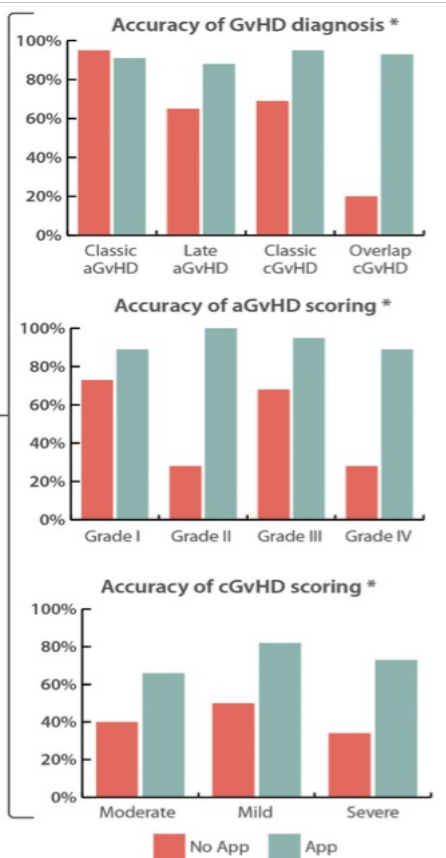
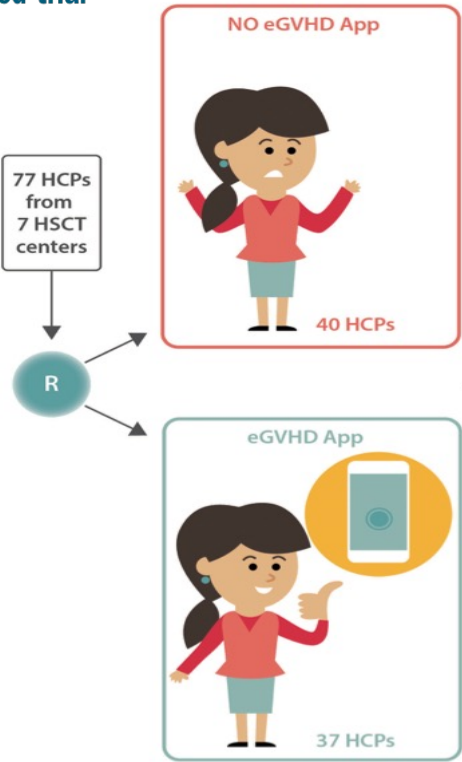
Ferrata Storti Foundation



Strumenti No App

- Proprie conoscenze **62%**
- Scheda di valutazione GvHD 2014 **23%**
- Scheda di valutazione NIH 2005 **15%**
- Documento di punteggio auto-progettato **15%**

Strumenti App



* Compared to the Expert Gold Standard

Accuratezza Diagnosi
10 casi/10 vs 6.5/10

Trade-off di tempo
48 vs 25 min

Accuratezza Scoring
9 casi/10 vs 4.5/10

Concordanza inter-operatore
(0.73 vs 0.56)

DIGITAL SOLUTION: BleedHD

Applicazione web focalizzata sul miglioramento della qualità dei dati raccolti

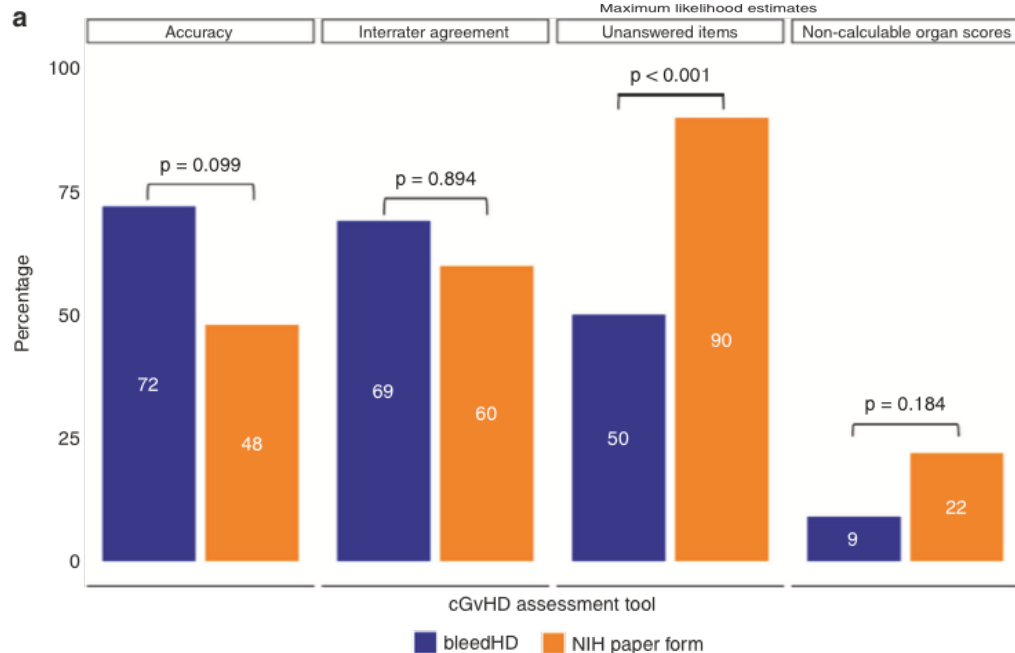
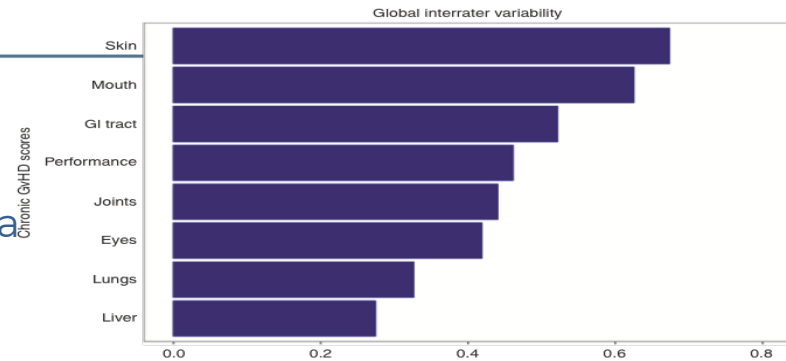
Sviluppo Basel

Accesso: website

Focus: Scoring cGvHD

Validazione: Validata su pazienti reali in contesto multiprofessionale

RIDUCE MISSING DATA



DIGITAL SOLUTION: GvHDRtrackR

Una suite di applicazioni per aGVHD e cGVHD sviluppata per il monitoraggio a lungo termine

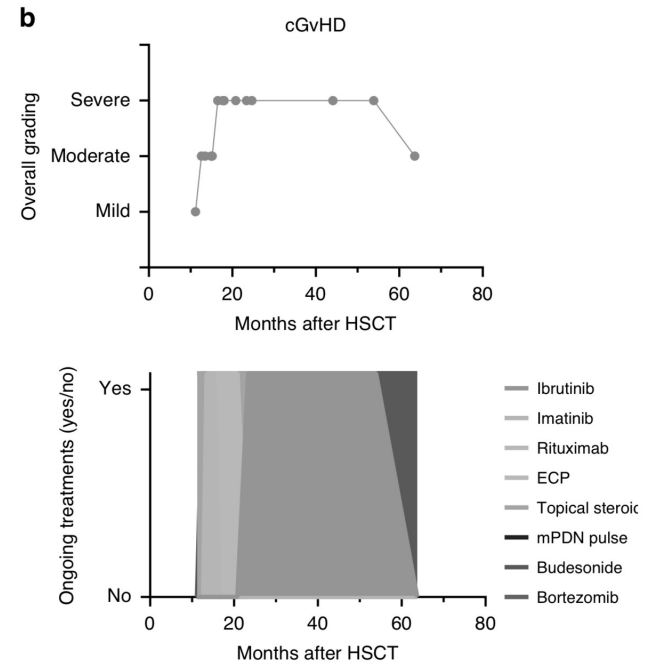
Sviluppo gruppo italiano (*Milano, Monza*)

Accesso: Desktop (Software Shiny)

Focus: Stadiazione/Scoring aGVHD e cGVHD

Validazione: Testata da team HSCT pediatrici (in fase di validazione multicentrica)

Follow up Longitudinale
Integrazione con Farmaci
Esportazione dati



DIGITAL SOLUTION: Crosy Software

ON

"Early software predicts survival in a p
diagnosed cG



TMO

predicts
58 newly

Grazie a GITMO e a
tutti i centri
partecipanti!

DIGITAL SOLUTION: Crosy Software

Sviluppo gruppo GITMO

Accesso: Software Bimind

Focus: Staging & Risposta cGVHD (&PROs) con
algoritmo automatico

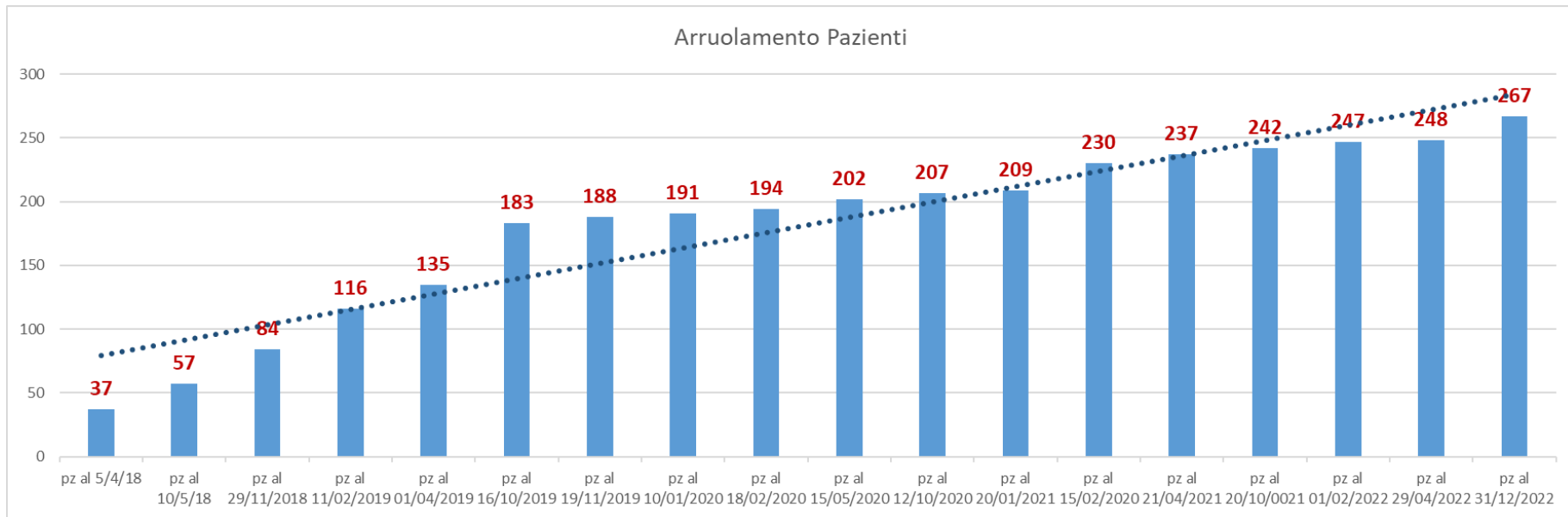
Validazione: Validata retrospettivamente
Database real-time



CENTRI TRAPIANTO IN ITALIA



Arruolamento: 23 centri



Primo Paziente inserito 31/10/2017 dal CIC 141 (Brescia)

Ultimo Paziente inserito 02/03/2022 dal CIC 788 (Ancona)

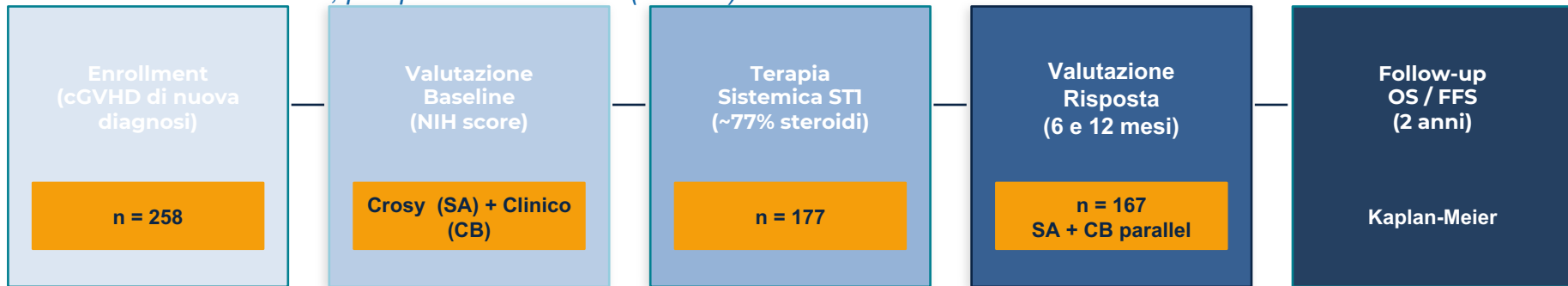


Barriere

- ✓ Tecniche (installazione e uso, approvazione CE)
- ✓ Tempo

CrOSY: Disegno dello Studio — NCT02991846

Studio non-interventistico, prospettico multicenter (GITMO)



Endpoint

PRIMARY

Failure-Free Survival (FFS) — sopravvivenza senza progressione, senza nuova IST, senza recidiva o tossicità grave

SECONDARY

ORR a 6 e 12 mesi (SA + CB); OS; % pazienti che sospendono IST

KEY ANALYSIS

Confronto SA (Crosy) vs CB response in termini di concordanza e predittività di OS — landmark analysis a 6 mesi

CrOSY: ORR a 6 mesi

I tassi a 6 mesi mostrano una concordanza moderata

PARAMETRO DI RISPOSTA	SA – CROSY (N, %)	CB – CLINICO (N, %)	DELTA / NOTE
CR (Remissione Completa)	33 (19%)	41 (23%)	Lieve discrepanza
PR (Risposta Parziale)	76 (43%)	74 (42%)	Sostanziale accordo
ORR (CR + PR)	109 (62%)	115 (65%)	Differenza marginale
Risposta Mista (MR)	40 (23%)	12 (7%)	Divergenza critica
Concordanza SA vs CB	65.9% (kappa = 0.53)		Accordo moderato

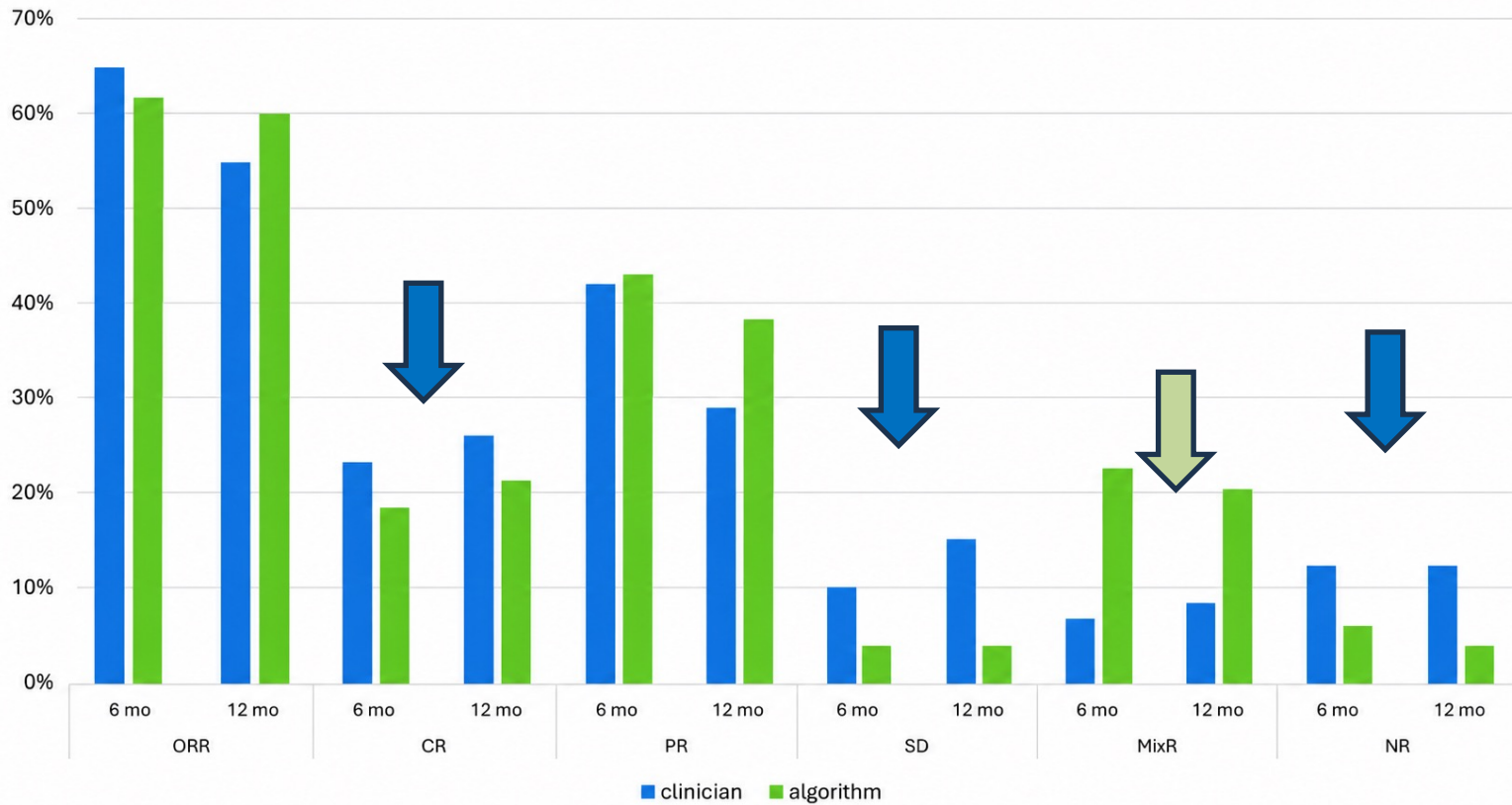
- Il software SA cattura le Risposte Miste (MR) molto meglio della valutazione clinica (23% vs 7%)
- I clinici tendono a classificare le MR come PR o PD in base all'impressione globale
- L'algoritmo NIH applicato uniformemente mantiene la distinzione prognostica della MR
- Questo differenzia la capacità predittiva dei due metodi



CrOSY: Discordanze

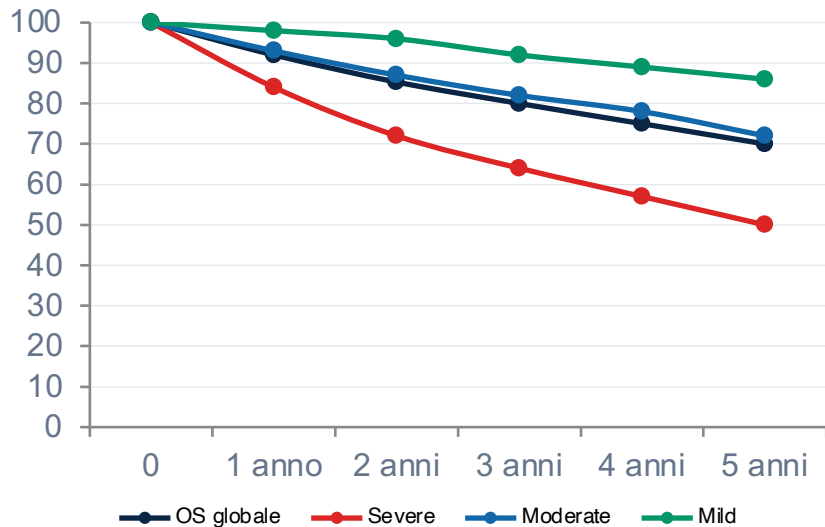
Il clinico tende a sovrastimare la stabilità (SD) e la progressione (PD)

Categoria Clinica (CB)	Tasso Discordanza	Destinazione Software (SA)
Unchanged (SD)	83%	39% verso PR 33% verso MR
Progressione (PD)	59%	41% verso MR
Risposta Mista (MR)	70%	30% verso PR 30% verso PD (CB)



CrOSY: OS e FFS

OS — Intera Coorte (n=258)



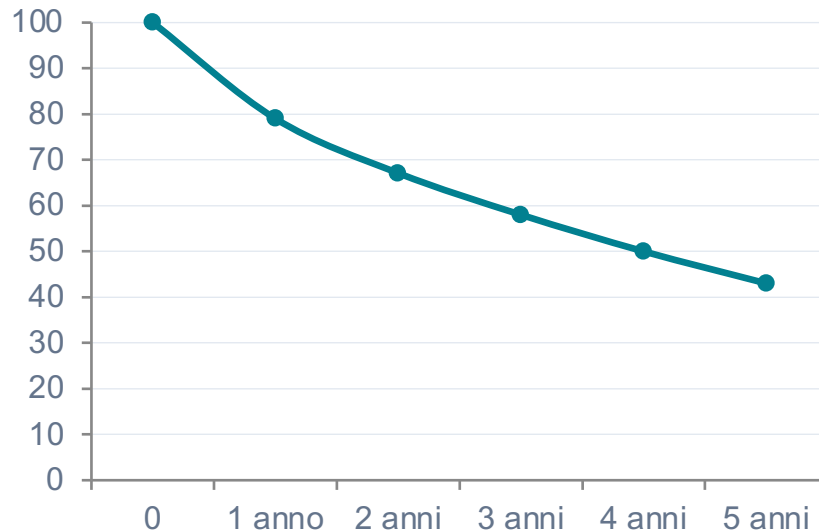
85.3%

OS a 2 anni
(intera coorte)

85.6%

OS a 2 anni
(pazienti in ST1)

FFS — Pazienti in ST1 (n=177)



67.1%

FFS a 2 anni
(ST1)

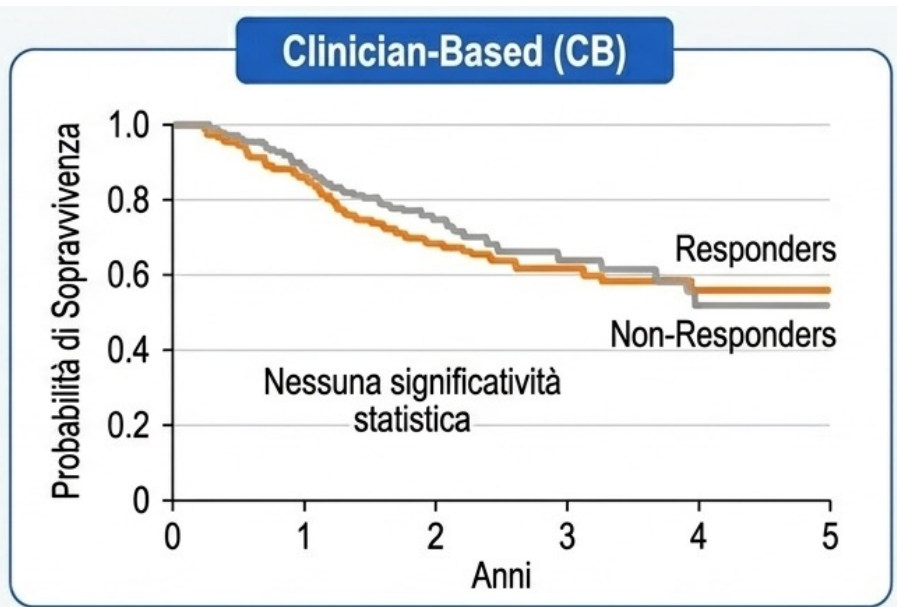
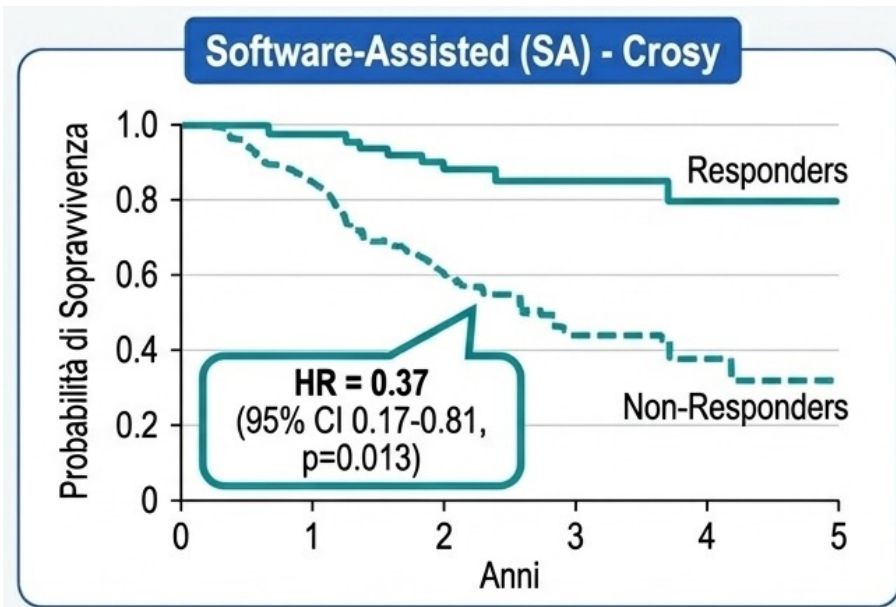
23%

off IST a 12 mesi
(in ST1 invariata: 39%)

CrOSY: OS

La risposta a 6 mesi calcolata tramite il Software è un **predittore indipendente** di OS.

OS a 2 anni 86%



L'Accuratezza Predice la Sopravvivenza

Multivariate Analysis for Overall Survival (OS)

Variabile Landmark (6 mesi)	Hazard Ratio (HR)	95% CI	P-Value
Risposta SA (Software-Assisted)	0.37	0.17 - 0.81	0.013
Risposta CB (Clinician-Based)	0.52	0.24 - 1.11	0.090 (NS)
Baseline ECOG PS	1.72	1.11 - 2.66	0.014
Baseline Skin NIH score	1.65	1.13 - 2.41	0.010

Conclusioni Crosy



Primo studio che confronta prospetticamente la valutazione automatica con quella convenzionale a intervalli di 6–12 mesi.

La valutazione della risposta SA a 6 mesi dopo STI influenza la OS → utile per studi futuri su nuovi farmaci.

Le discrepanze tra valutazioni cliniche e tra centri → L standardizzati

La valutazione della risposta multicentrici se support



...take your time

abilità dell'esperienza dei applicando criteri NIH

nche in contesti
...E di facile "accesso")

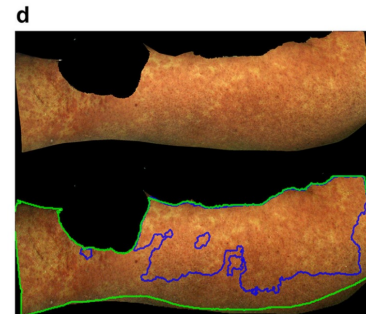
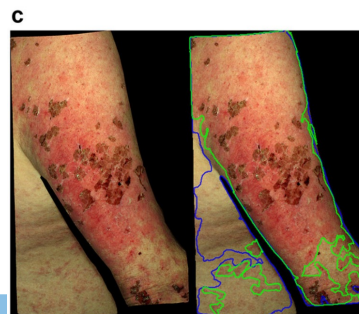
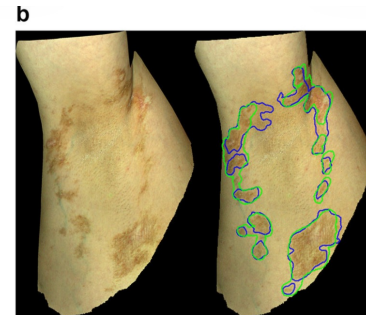
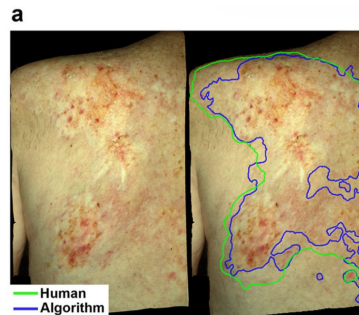
FUTURO: Dall'Algoritmo all'AI & MACHINE LEARNING



Modelli di Machine Learning per l'elaborazione di dati temporali e stratificazione dinamica del rischio.

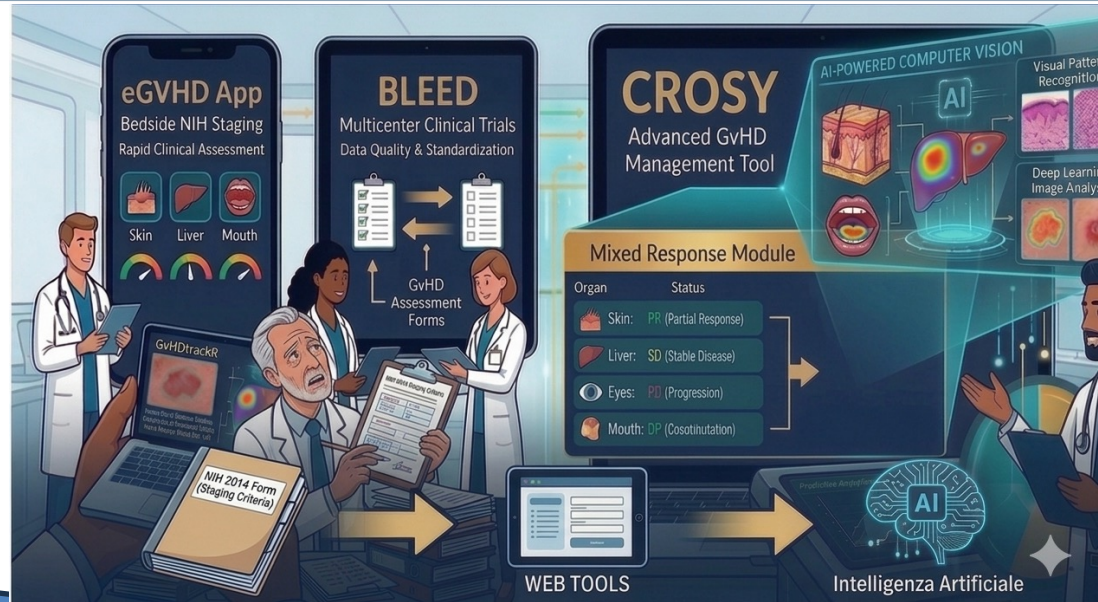
✓ **Visione Artificiale:** Segmentazione 3D delle lesioni cutanee con performance esperte (Errore di Superficie AI 9% vs 20% clinico).

✓ **Risk Assessment Dinamico & Adaptive Strategies:** Personalizzazione delle terapie in tempo reale (ATS)



L'ACCURATEZZA NON E' SOLO MIGLIORAMENTO TECNICO

Sviluppo e l'implementazione di strumenti digitali **semplici, intuitivi, applicabili** nella pratica clinica, **affidabili** e in grado di integrare tutte le informazioni rilevanti.



Corretta
analisi dei dati

Miglior
predittività
degli outcome

Più precisa
definizione
delle strategie
terapeutiche

Reale
miglioramento
degli outcome

Grazie!

- **Ai Centri GITMO partecipanti,**
- **Al GITMO (past & present) in particolare Dr.ssa Sonia Mammoliti, Prof. Andrea Bacigalupo, Prof.ssa Francesca Bonifazi**

- **Team Trapianti** della Clinica di Ematologia di Ancona
Dr.ssa Irene Federici
Dr.ssa Selene Guerzoni
- **Dr Jacopo Olivieri** per l'elaborazione statistica
- **BIMIND** per l'assistenza software ed estrazione dati

Special Thanks

Prof Attilio Olivieri

Dr. Francesco Saraceni

....e a tutti voi per l'attenzione

