

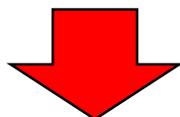
La concordanza

Concordanza

Ci sono situazioni in cui si vogliono confrontare due metodi o valutare la riproducibilità di un metodo confrontando due operatori diversi o il medesimo operatore in due tempi diversi.

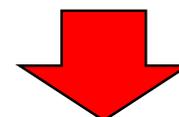
Gli strumenti statistici da utilizzare per la valutazione della concordanza dipendono dalla natura del fenomeno in oggetto:

misurazione
(variabile quantitativa)



Metodo di BLAND-ALTMAN

valutazione
(variabile qualitativa)



Indici K

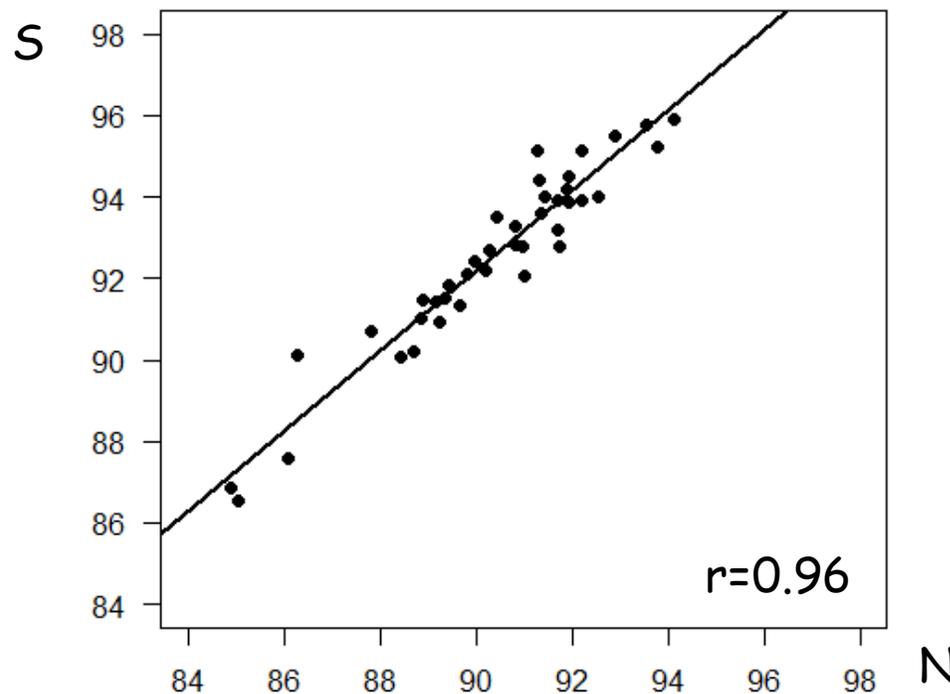
Concordanza tra misurazioni

ESEMPIO

Si vuole confrontare un nuovo metodo (N), che è veloce ed economico, con quello standard (S) nella misurazione di un certo parametro.

E' stato approntato un esperimento con 40 misurazioni effettuate utilizzando entrambi i metodi.

Sulla base dei risultati ottenuti, conviene sostituire S con N, o usare S ed N in maniera interscambiabile?



Concordanza tra misurazioni

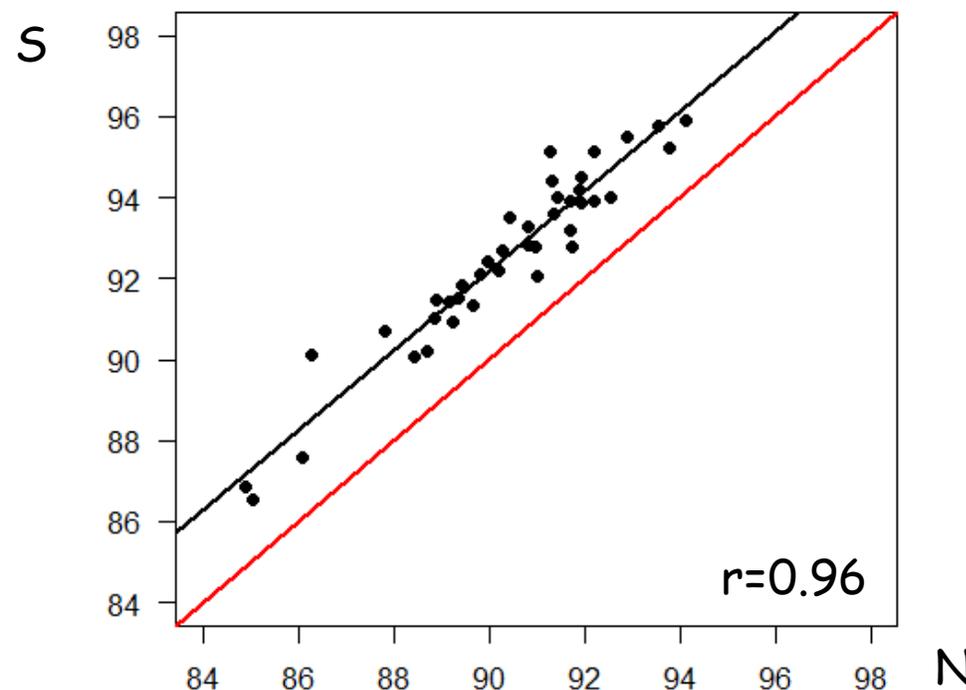
ESEMPIO

Si vuole confrontare un nuovo metodo (N), che è veloce ed economico, con quello standard (S) nella misurazione di un certo parametro.

E' stato approntato un esperimento con 40 misurazioni effettuate utilizzando entrambi i metodi.

Sulla base dei risultati ottenuti, conviene sostituire S con N, o usare S ed N in maniera interscambiabile?

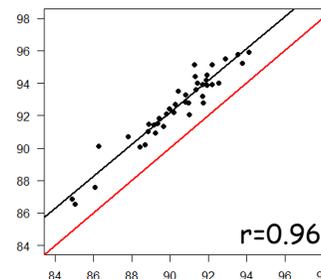
NO!!!



Concordanza tra misurazioni

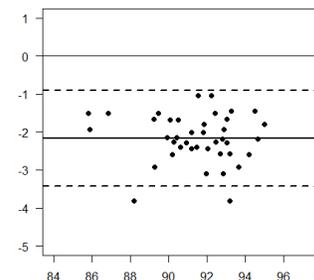
Vi è la consuetudine di valutare graficamente la concordanza tra due misure 1) utilizzando un grafico di dispersione e 2) quantificando la bontà della loro concordanza attraverso il coefficiente di correlazione.

$$Y_N \text{ vs } Y_S$$



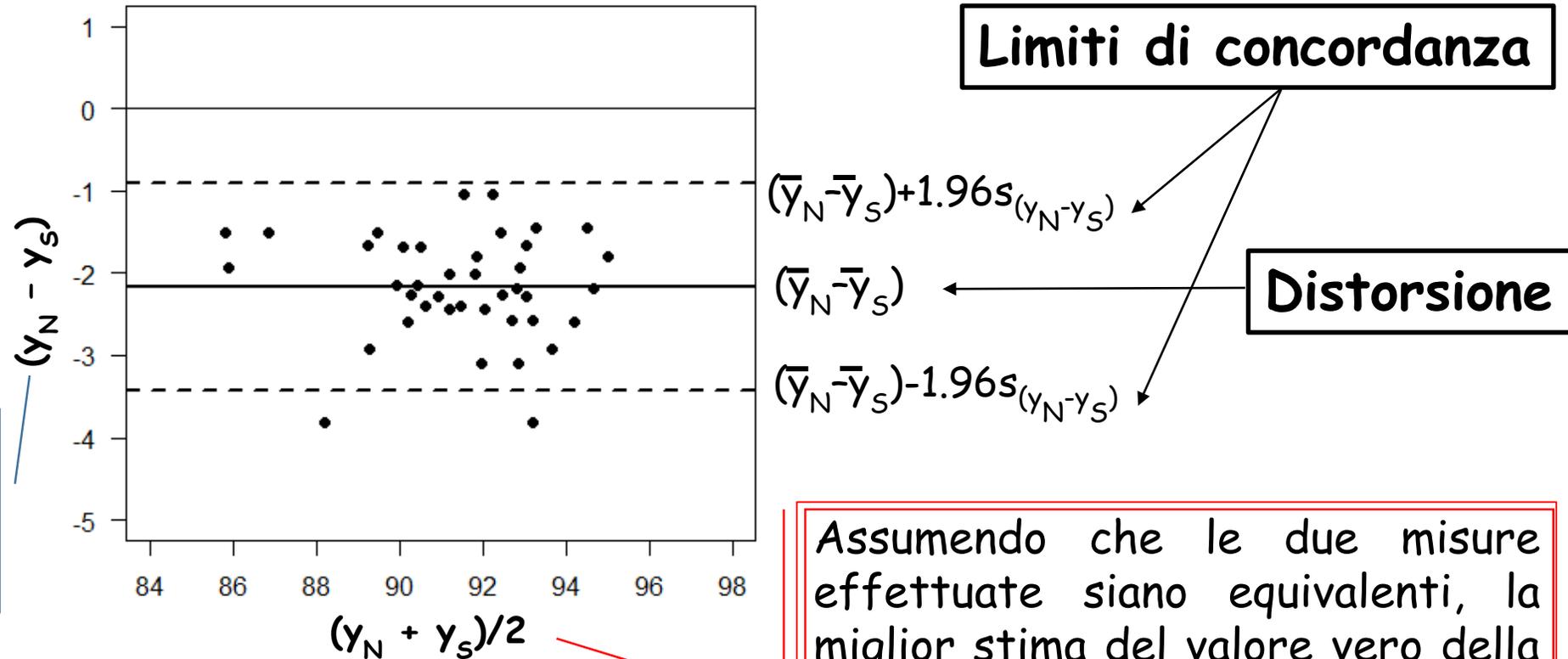
Bland e Altman (1999) hanno proposto una tecnica grafica più appropriata e di facile uso ed interpretazione che si basa sull'analisi delle differenze fra le due misurazioni:

$$(Y_N - Y_S) \text{ vs } (Y_N + Y_S)/2$$



Metodo di Bland & Altman

Il grafico di $(y_N - y_S)$ vs $(y_N + y_S)/2$ esplora la relazione tra l'errore di misura e il **valore vero della misura**

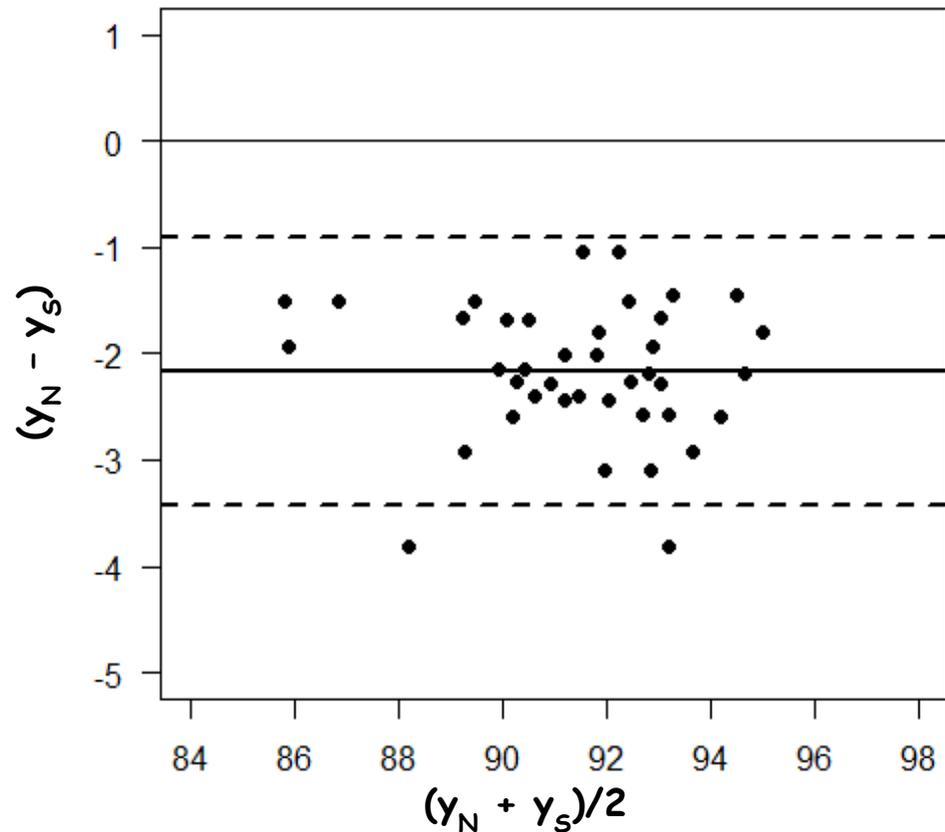


La differenza fra le due misurazioni rappresenta l'errore di misura.

Assumendo che le due misure effettuate siano equivalenti, la miglior stima del valore vero della misura è rappresentata dalla media aritmetica delle due misure.

Metodo di Bland & Altman

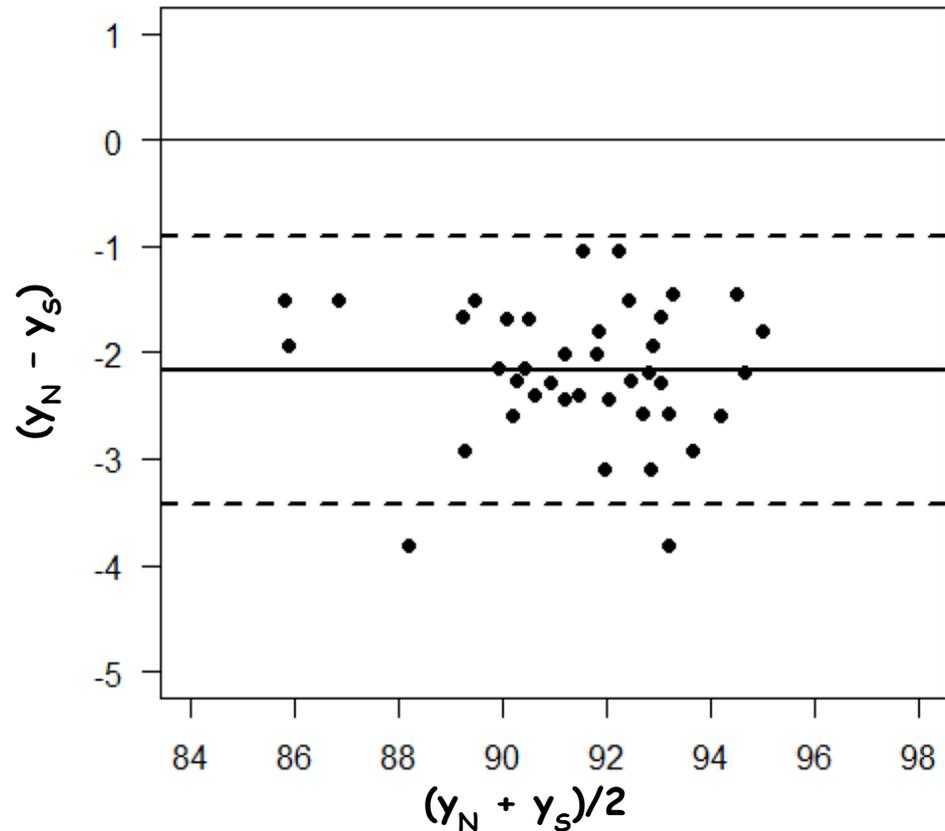
Interpretazione del grafico di B&A e dei relativi indicatori di concordanza:



- 1) La media delle differenze tra le misurazioni dei due metodi è una stima del livello di concordanza.
- 2) Ci aspettiamo che il 95% circa delle differenze sia compreso entro i limiti di concordanza, ovvero tra $(\bar{y}_N - \bar{y}_S) - 1.96s_{(y_N - y_S)}$ e $(\bar{y}_N - \bar{y}_S) + 1.96s_{(y_N - y_S)}$
- 3) Se l'ampiezza dell'intervallo di concordanza non è rilevante da un punto di vista clinico, allora le due misure possono essere utilizzate intercambiabilmente.

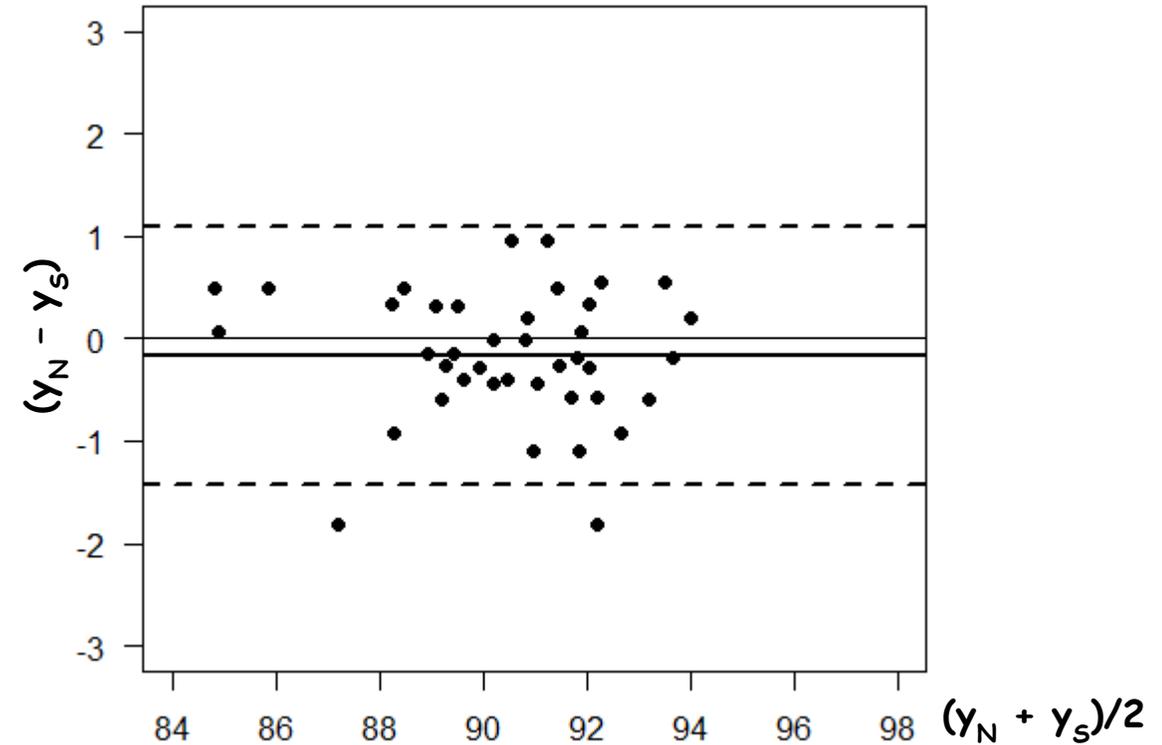
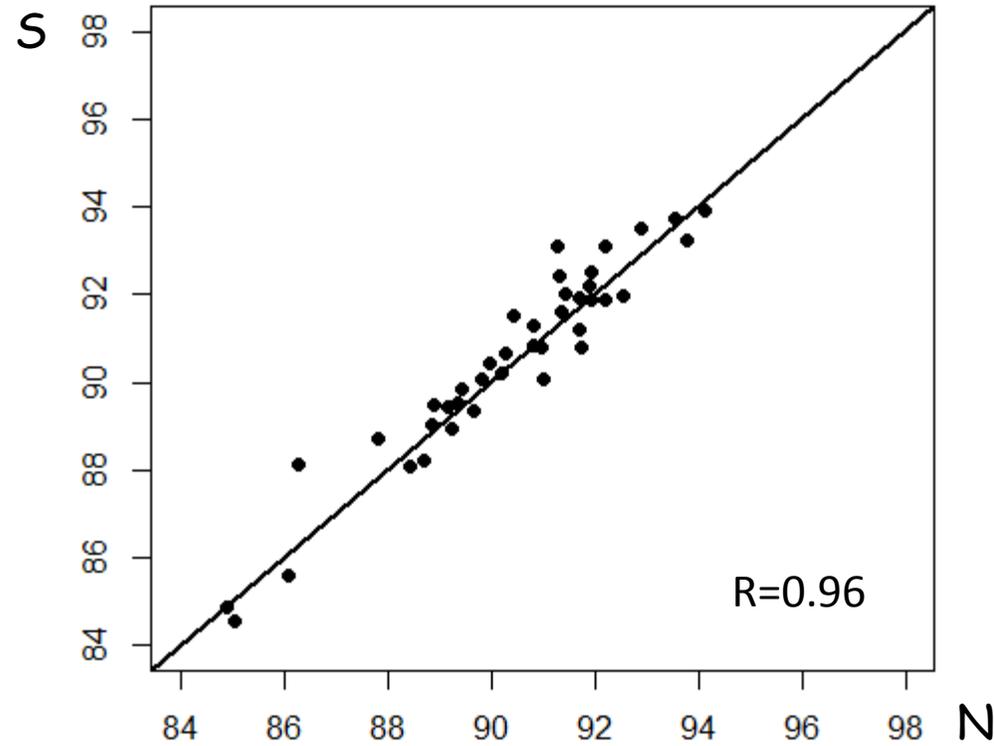
Metodo di Bland & Altman

Cosa posso dedurre da questo grafico?



- 1) Il nuovo metodo (N) tende sistematicamente a sottostimare le misure effettuate con il metodo standard (S) di 2.3 unità.
- 2) La differenza tra le misurazioni non è «quasi mai» inferiore a -1.1 o superiore a 3.4 unità (ci attendiamo solo il 5% delle differenze fuori da questo intervallo).
- 3) L'ampiezza dell'intervallo è clinicamente abbastanza stretta da giustificare un uso intercambiabile delle due misurazioni?

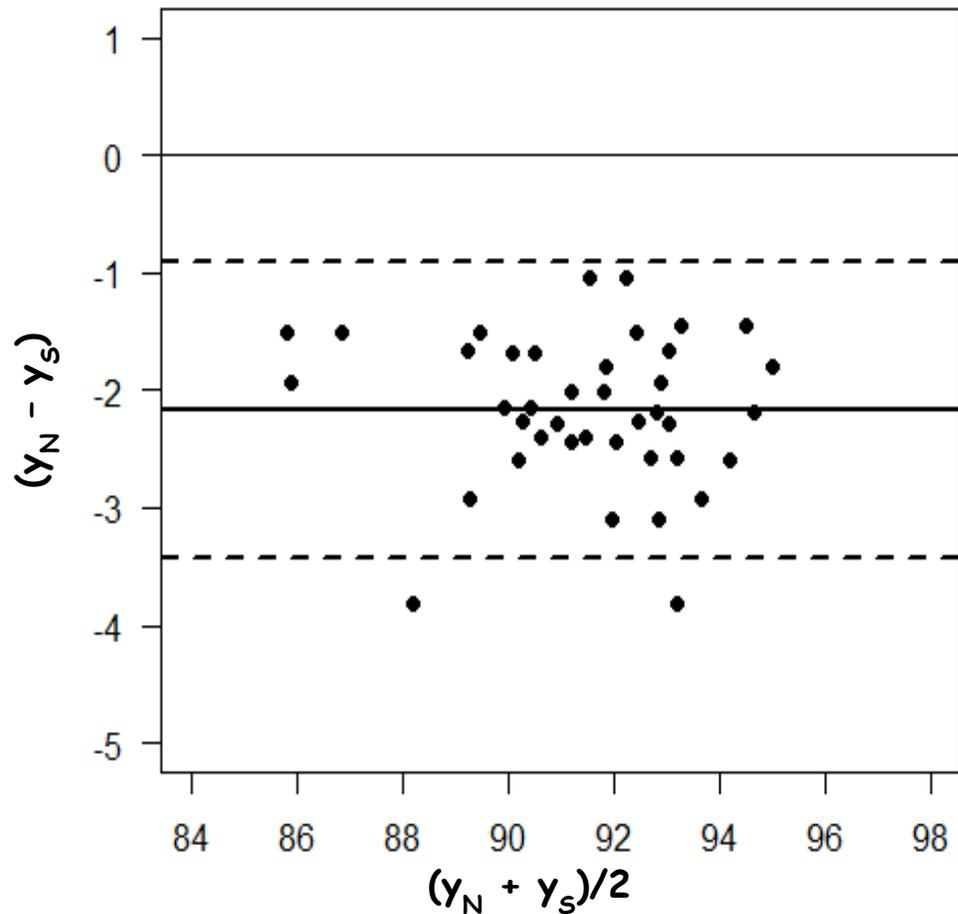
Metodo di Bland & Altman



Il diagramma di Bland-Altman mostra come non vi sia sostanziale differenza sistematica fra i due metodi e che le differenze tra i metodi oscillano nell'intervallo -1.3 e 1.1 unità indipendentemente dal valore della variabile misurata.

Metodo di Bland & Altman

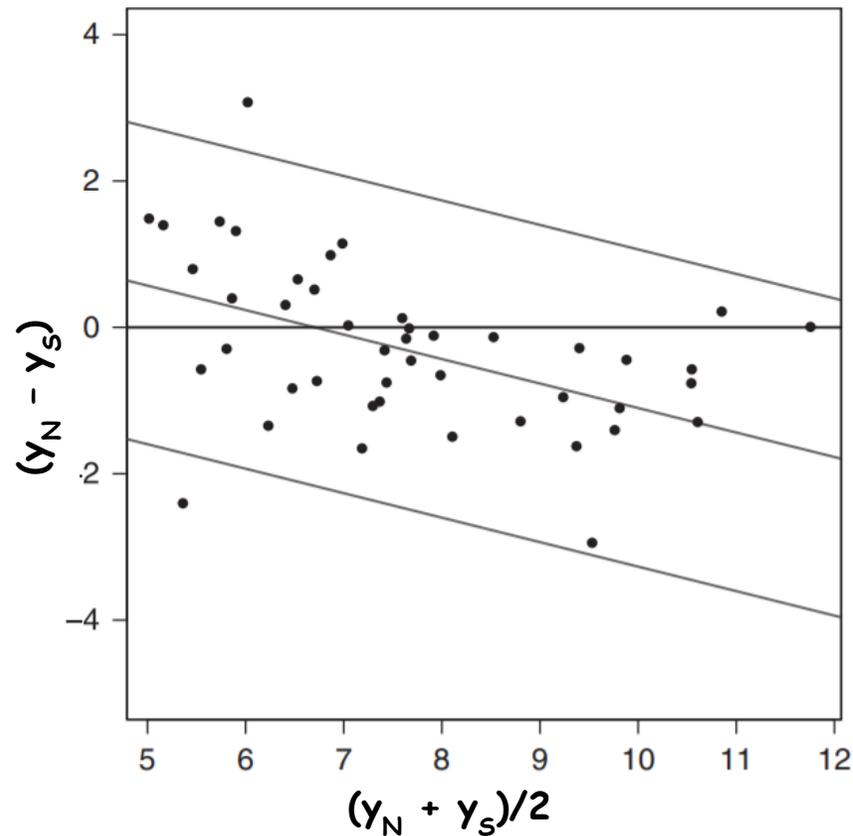
Assunzioni del metodo di B&A:



- 1) «**Differenza costante**»: la differenza tra due misurazioni non dipende dall'ordine di grandezza (ovvero dalla media).
- 2) «**Varianza costante**»: La variabilità delle differenze non dipende dall'ordine di grandezza (ovvero dalla media).

Metodo di Bland & Altman

Assunzioni del metodo di B&A e **come verificarle**:



1) «Differenza costante»: la differenza tra due misurazioni non dipende dall'ordine di grandezza (ovvero dalla media).

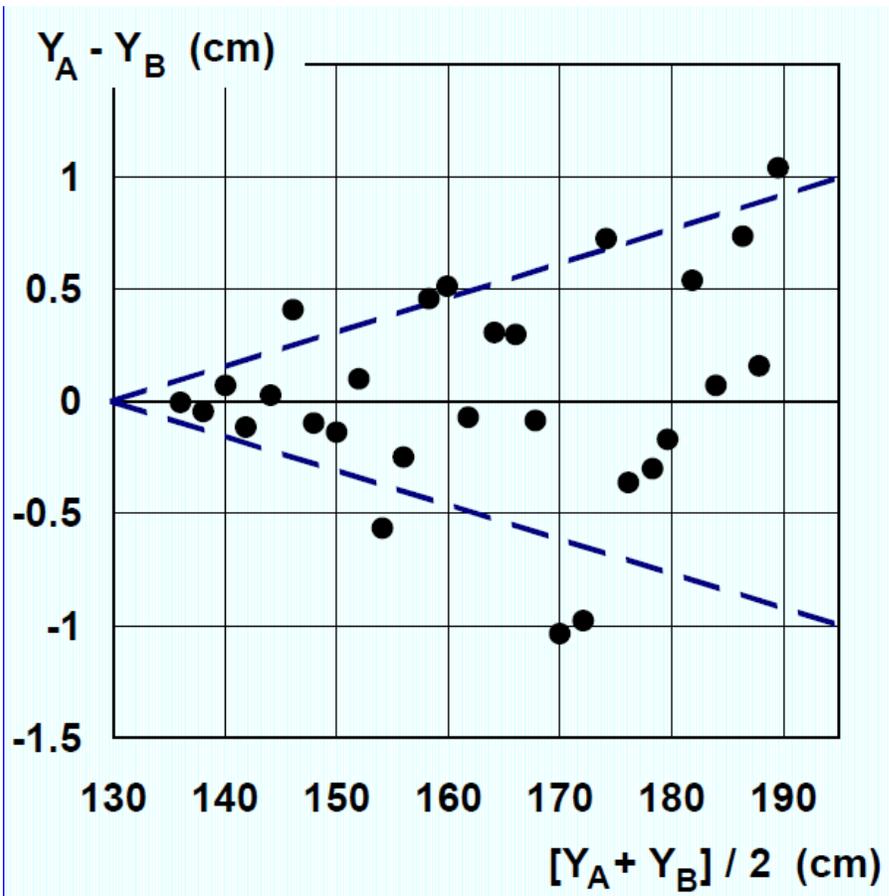
Stimare un modello di regressione lineare con Y =«differenze» e X =«medie» e verificare che il coefficiente angolare sia ≈ 0 .

$$(y_N - y_S) = 2.24 - 0.33 (y_N + y_S)/2$$

Per valori bassi, il metodo N tende a sovrastimare S, ma al crescere dei valori la tendenza si inverte. I limiti di concordanza vanno calcolati tenendo conto di questo aspetto.

Metodo di Bland & Altman

Assunzioni del metodo di B&A e **come verificarle**:



2) «**Varianza costante**»: La variabilità delle differenze non dipende dall'ordine di grandezza (ovvero dalla media).

Stimare un modello di regressione lineare con Y =«residui del modello differenze vs medie» e X =«medie» e verificare che il coefficiente angolare sia ≈ 0 .

Anche se non sembra che vi sia una differenza sistematica tra i due metodi, l'imprecisione aumenta all'aumentare delle medie. I limiti di concordanza vanno quindi opportunamente calcolati tenendo conto di questo aspetto.

Concordanza tra valutazioni

ESEMPIO

Due clinici (A e B) esprimono un parere sulla presenza o assenza di una certa condizione patologica in un gruppo di 100 pazienti.

		Osservatore A		Totale
		Presente	Assente	
Osservatore B	Presente	29	14	43
	Assente	9	48	57
Totale		38	62	100

Come si può quantificare il grado di accordo tra i due osservatori?

Concordanza tra valutazioni

		Osservatore A		Totale
		Presente	Assente	
Osservatore B	Presente	a=29	b=14	43
	Assente	c=9	d=48	57
Totale		38	62	n=100

L'accordo osservato (P_o) è la proporzione di pazienti sui quali i due osservatori esprimono lo stesso giudizio:

$$P_o = \frac{a+d}{n} = (29+48) / 100 = 77\%$$

Concordanza tra valutazioni

Vi è tuttavia una quota attesa di giudizi concordanti che si verificano per il solo effetto del caso.

Osservatore A

		Osservatore A		Totale
		Presente	Assente	
Osservatore B	Presente	a=29	b=14	43
	Assente	c=9	d=48	57
Totale		38	62	n=100

Tale accordo casuale atteso (P_a) si ottiene nel modo seguente:

$$P_a = \frac{(a+c)(a+b)+(b+d)(c+d)}{n^2} = (38*43 + 62*57)/100^2 = 51.7\%$$

Indice K di Cohen

Un indice di «accordo oltre il caso» è quindi ottenibile scontando l'accordo casuale dall'accordo osservato e rapportando questo valore al massimo accordo non casuale osservabile.

$$K = \frac{P_o - P_a}{1 - P_a} = (0.77 - 0.517) / (1 - 0.517) = 52.4\%$$

L'indice K vale 1 in caso di completo accordo e vale 0 (o <0) in caso di accordo nullo (o addirittura accordo osservato inferiore a quello casuale).

Il calcolo di K è immediatamente estendibile al caso di giudizi con >2 categorie. Inoltre, ne esiste una versione pesata (**weighted K**) applicabile al caso di **categorie ordinali**.

Scandinavian Journal of Gastroenterology. 2015; 50: 429–438

informa
healthcare

ORIGINAL ARTICLE

Simultaneous liver iron and fat measures by magnetic resonance imaging in patients with hyperferritinemia

STEFANIA GALIMBERTI^{*1}, PAOLA TROMBINI^{*2}, DAVIDE PAOLO BERNASCONI¹,
IRENE REDAELLI³, SARA PELUCCHI⁴, GIORGIO BOVO⁵, FILIBERTO DI GENNARO⁶,
NICOLA ZUCCHINI⁵, NICOLETTA PARUCCINI³ & ALBERTO PIPERNO^{4,7}

¹*Department of Health Sciences, Centre of Biostatistics for Clinical Epidemiology, University of Milano-Bicocca, Monza, Italy,* ²*Internal Medicine Unit, ICP Sesto S. Giovanni Hospital, Milan, Italy,* ³*Medical Physics Unit, S. Gerardo Hospital, Monza, Italy,* ⁴*Department of Health Sciences, University of Milano-Bicocca, Monza, Italy,* ⁵*Clinical Pathology Unit, S. Gerardo Hospital, Monza, Italy,* ⁶*Clinical Radiology Unit, S. Gerardo Hospital, Monza, Italy,* and ⁷*Centre for Hemochromatosis and Iron Metabolism Disorders, S. Gerardo Hospital, Monza, Italy*

Risultati

Sample description

Table I. Characteristics of the 67 patients with hyperferritinemia who underwent MRI and liver biopsy.

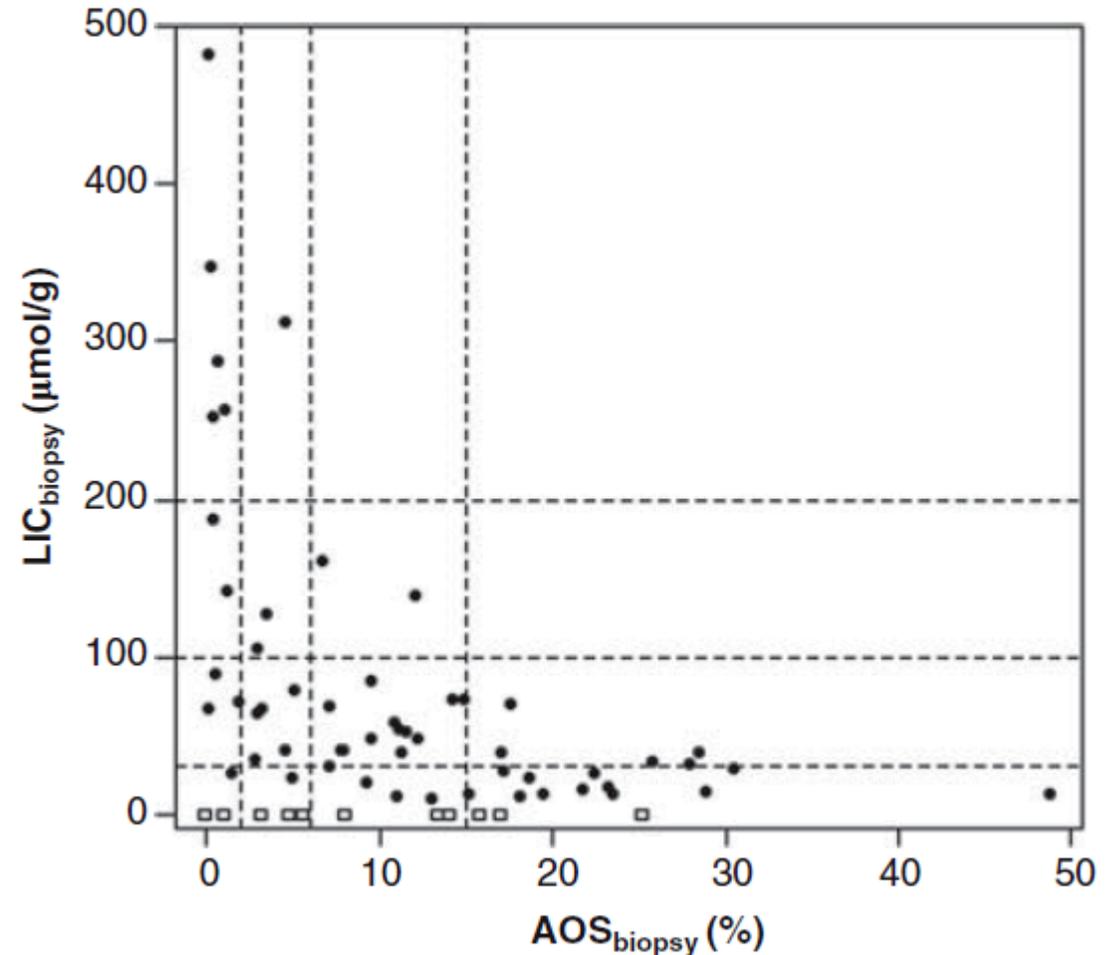
Parameter	Results ^a
Age at liver biopsy (years)	50.5 (45.2–60)
Males	59 (88)
Fibrosis (grade)	
0–1	4 (6)
2–4	50 (75)
5–6	13 (19)
Hepatic iron	
Total iron score	15 (12–25)
Liver iron concentration ^b ($\mu\text{mol/g}$)	44.6 (26.4–81.6)
Hepatic fat	
Grading of fatty hepatocytes (%)	70.0 (25.0–90.0)
Area of steatosis (%)	9.5 (3.2–17.1)

^aResults are reported as *n* (%) or median (I–III quartile).

^bLIC_{biopsy} was missing in 11 subjects.

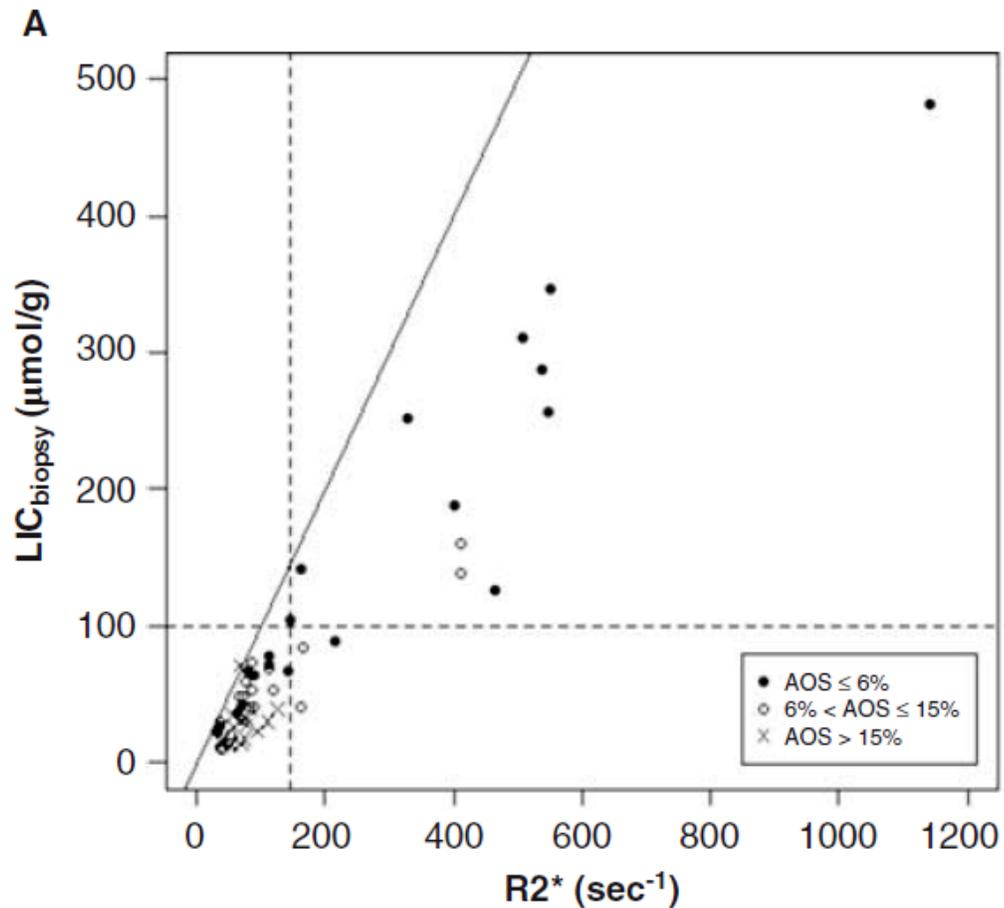
Abbreviation: LIC = Liver iron concentration.

Liver biopsy: iron vs fat



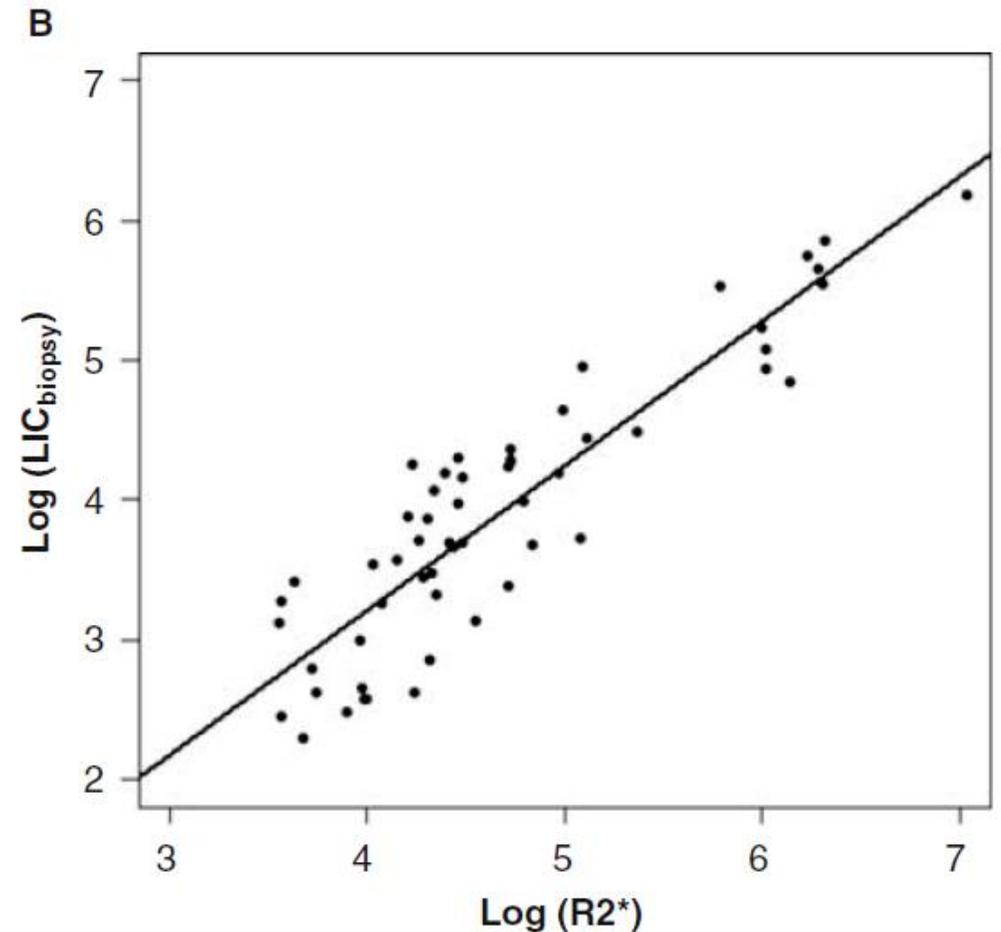
Biopsia vs MRI: ferro

Misurazioni originali

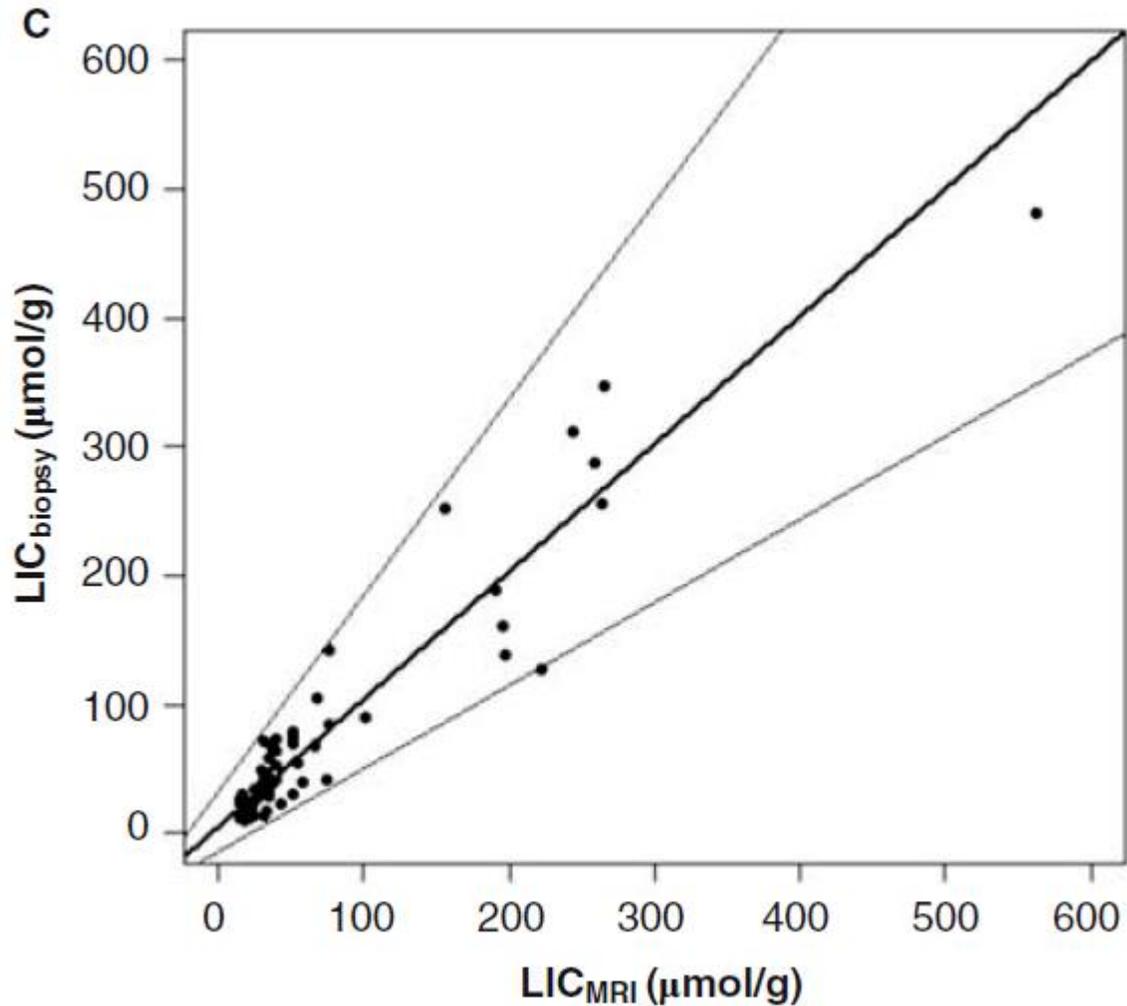


Retta di calibrazione

$$\log(\text{LIC}_{\text{MRI}}) = -0.92 + 1.03\log(\text{R2}^*) \quad (R^2 = 0.81)$$



Biopsia vs MRI: ferro



Concordanza dopo calibrazione

Distorsione = 2.8 micromol/g

Assunzioni:

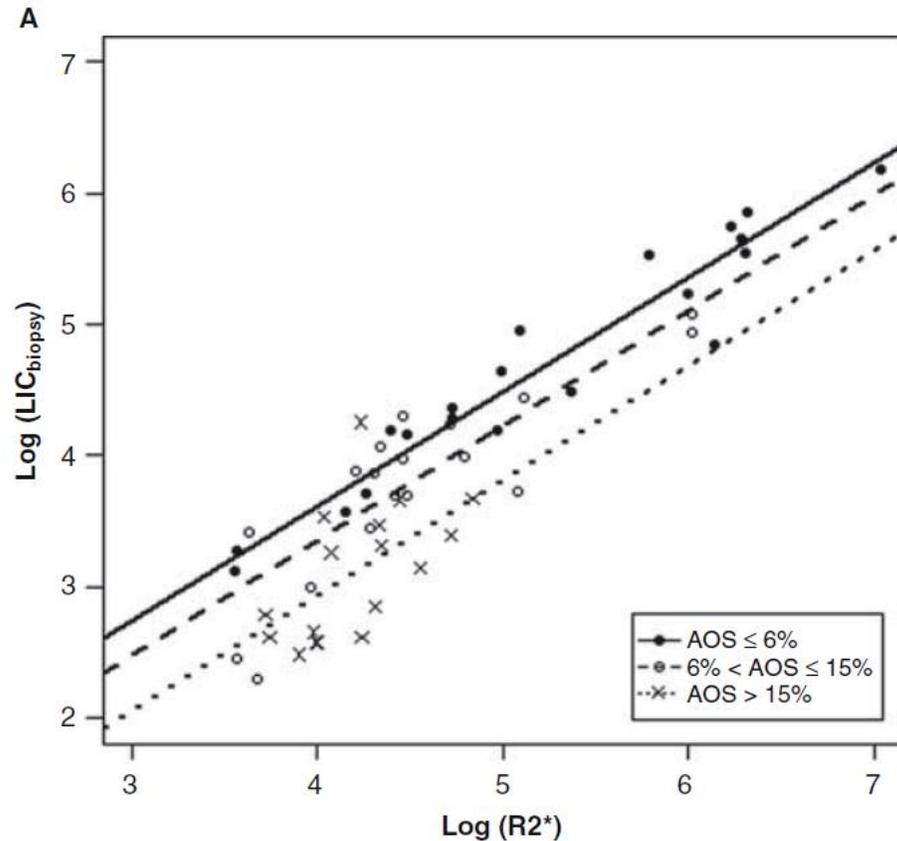
- Differenza costante
- Varianza delle differenze cresce all'aumentare delle medie.

I limiti di concordanza sono quindi calcolati in funzione delle medie delle misurazioni

Biopsia vs MRI: ferro by grasso

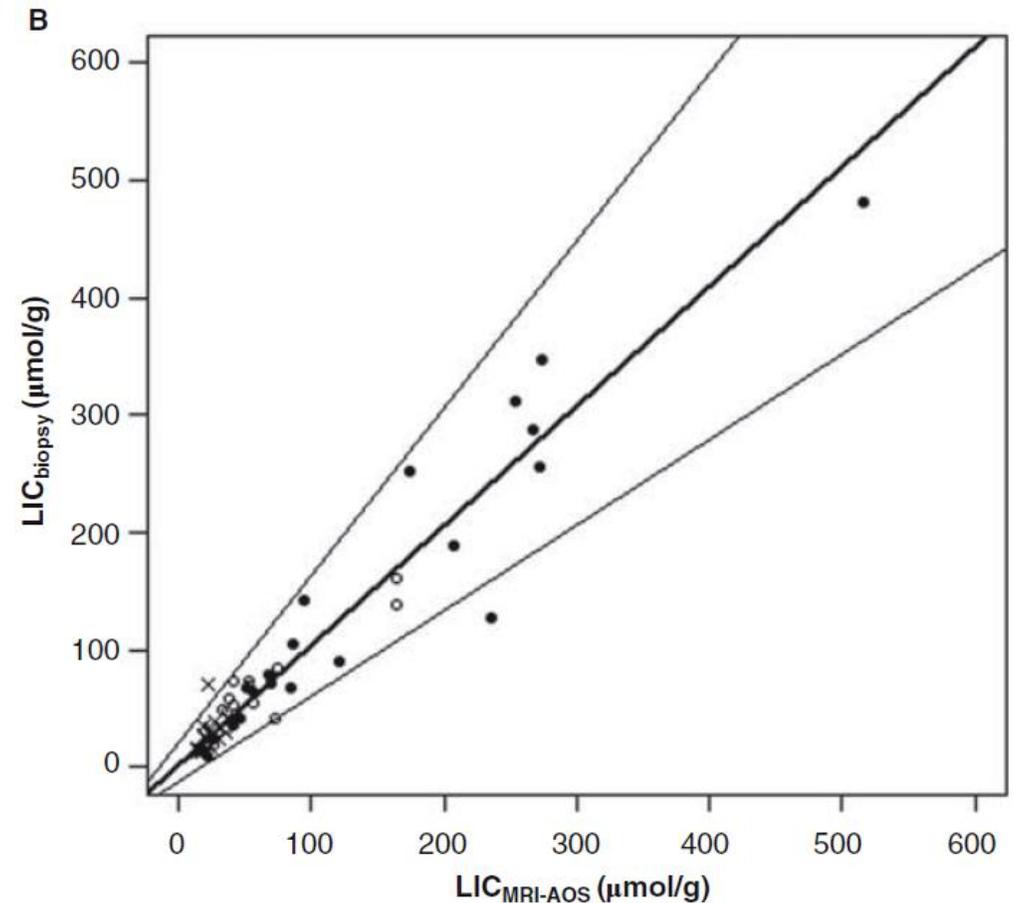
Rette di calibrazione ($R^2 = 0.88$)

Mild AOS $\log(\text{LIC}_{\text{MRI}}) = 0.12 + 0.87\log(\text{R2}^*)$
Moderate AOS $\log(\text{LIC}_{\text{MRI}}) = -0.14 + 0.87\log(\text{R2}^*)$
Severe AOS $\log(\text{LIC}_{\text{MRI}}) = -0.56 + 0.87\log(\text{R2}^*)$



Concordanza dopo calibrazione

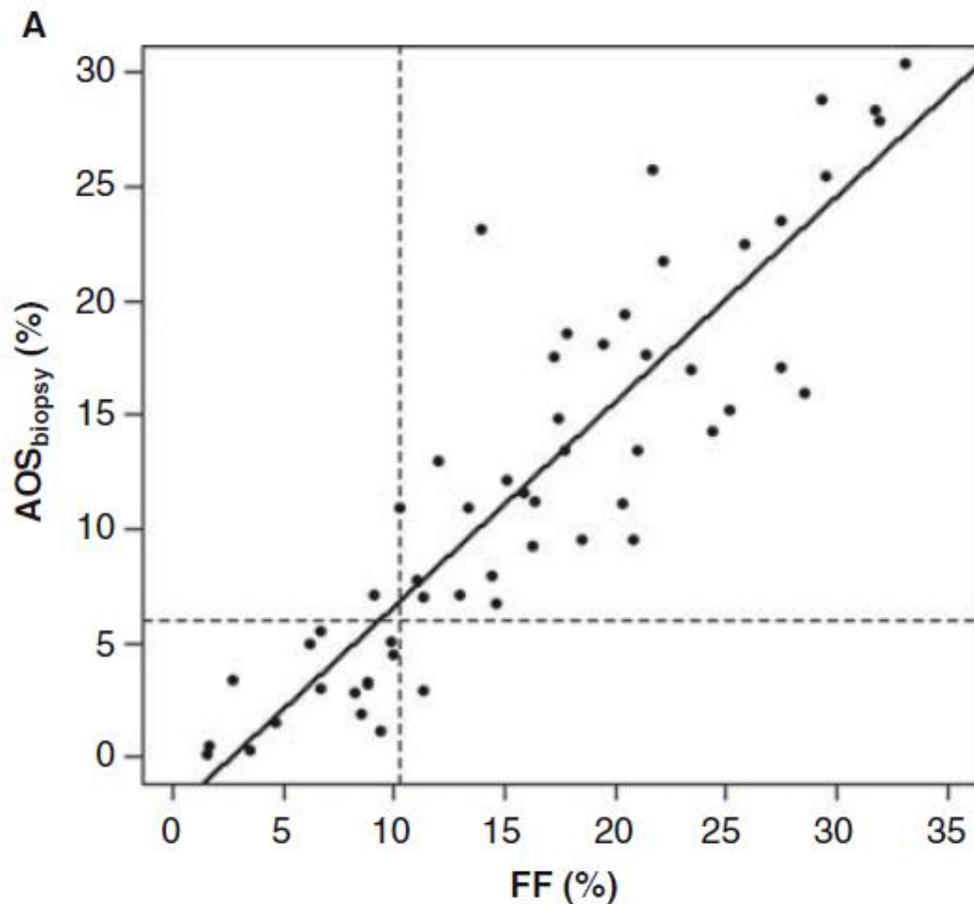
Distorsione ≈ 0



Biopsia vs MRI: grasso

Retta di calibrazione

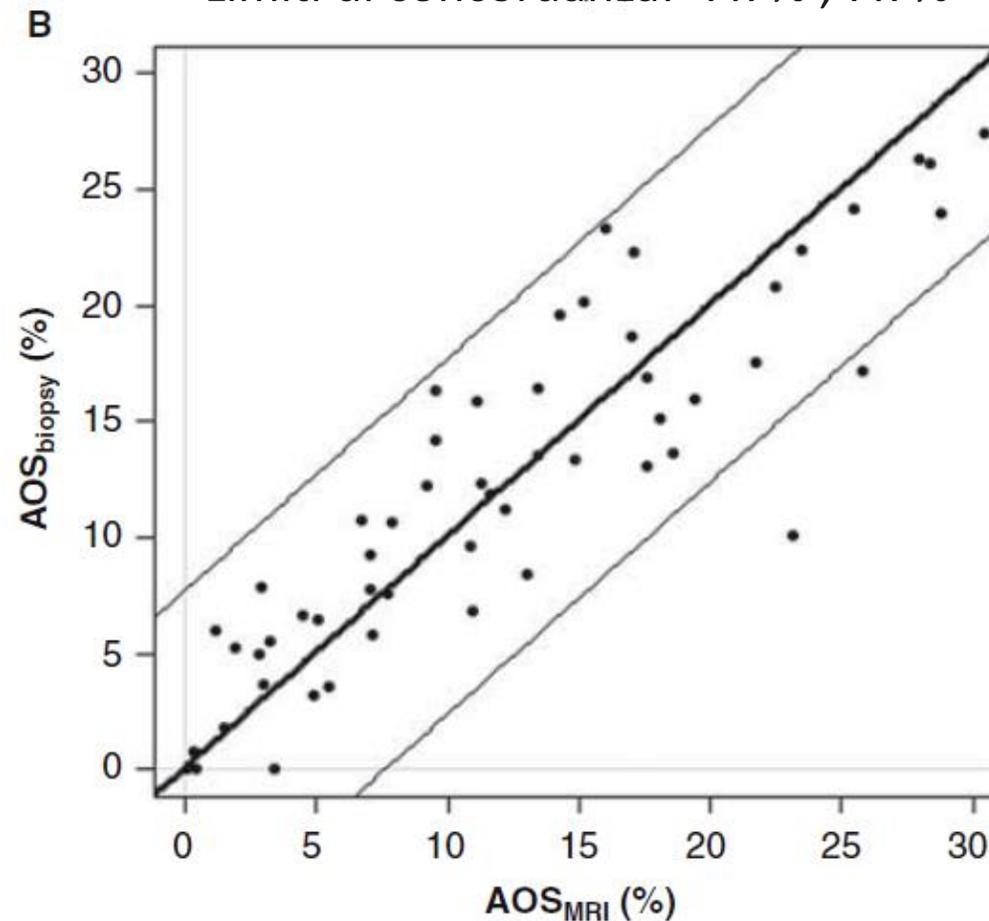
$$\text{AOS}_{\text{MRI}} = 2.40 + 0.90\text{FF} \quad (R^2 = 0.79)$$



Concordanza dopo calibrazione

Distorsione = -0.04%

Limiti di concordanza: -7.7% ; 7.7%



Biopsia vs MRI: concordanza qualitativa

Table II. Agreement between observed and predicted values of hepatic iron (a)–(b) and fat (c) by MRI.

(a)

LIC _{biopsy}	LIC _{MRI}				Total
	Absent	Mild	Moderate	Severe	
Absent	14	3	0	0	17
Mild	2	26	0	0	28
Moderate	0	1	3	1	5
Severe	0	0	1	5	6
Total	16	30	4	6	56 ^a

(b)

LIC _{biopsy}	LIC _{MRI-AOS}				Total
	Absent	Mild	Moderate	Severe	
Absent	17	0	0	0	17
Mild	4	23	1	0	28
Moderate	0	1	2	2	5
Severe	0	0	1	5	6
Total	21	24	4	7	56 ^a

(c)

AOS _{biopsy}	AOS _{MRI}				Total
	Absent	Mild	Moderate	Severe	
Absent	4	1	0	0	5
Mild	1	9	1	0	11
Moderate	0	0	17	3	20
Severe	0	0	3	17	20
Total	5	8	22	21	56 ^b

$$P_o = 48/56 = 85.7\%$$

$$wK = 0.84, 95\% \text{ CI } 0.74\text{--}0.95$$

$$P_o = 47/56 = 83.9\%$$

$$wK = 0.83, 95\% \text{ CI } 0.74\text{--}0.93$$

$$P_o = 47/56 = 83.9\%$$

$$wK = 0.84, 95\% \text{ CI } 0.75\text{--}0.94$$

Biopsia vs MRI: validazione interna

Dati di 10 «nuovi» pazienti arruolati consecutivamente

Table III. Hepatic iron and fat values obtained by biopsy and predicted by MRI in the validation sample. Classes of severity of steatosis and iron overload are also reported.

ID	FF (%)	AOS _{biopsy} (%)	AOS _{MRI} (%)	R2* (sec ⁻¹)	LIC _{biopsy} (μmol/g)	LIC _{MRI} (μmol/g)	LIC _{MRI-AOS} (μmol/g)
7	-	0.11	<2	369.46	118.0	175.81	193.15
3	3.53	0.18	0.78	161.39	87.6	74.91	93.97
6	-	1.20	<2	611.59	93.3	295.46	299.46
5	-	1.73	<2	761.32	256.4	370.22	362.31
8	6.32	3.06	3.28	68.25	27.4	30.87	44.44
1	11.07	6.22	7.56	143.36	66.9	66.31	65.36
10	20.98	13.93	16.48	101.17	36.9	46.31	31.71
9	25.26	21.34	20.33	65.74	15.9	29.70	21.79
2	21.90	22.49	17.31	48.36	22.2	21.65	16.69
4	24.36	34.53	19.52	65.62	27.8	29.65	21.76

Abbreviations: A = Absent; AOS = Area of steatosis; FF = Fat fraction; LIC = Liver iron concentration; M = Mild; Mo = Moderate; S = Severe.