Bland–Altman Analyse

Burkhardt Seifert

Abteilung Biostatistik
Universität Zürich

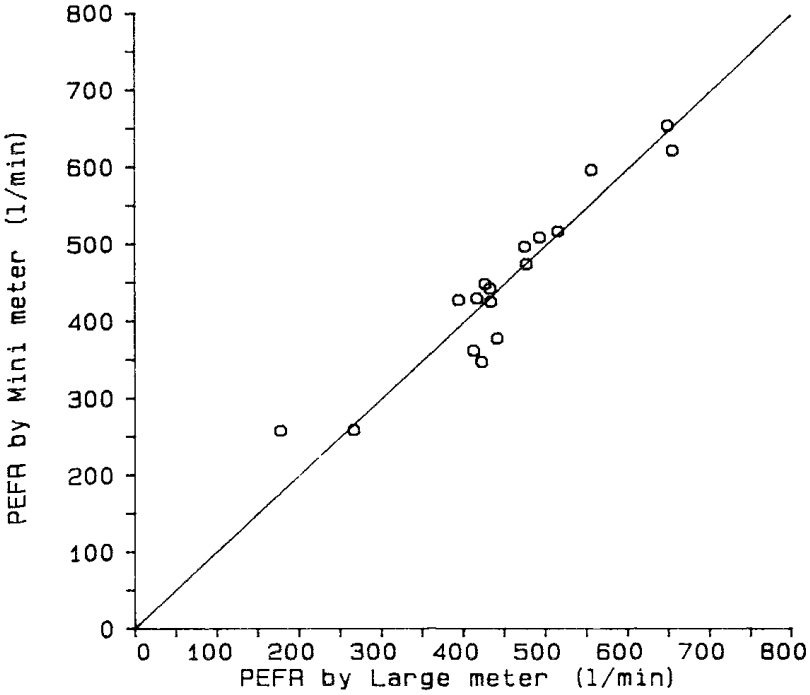
- Bland–Altman (1986). *Lancet*
- Bland–Altman Plot
- ergänzende Angaben

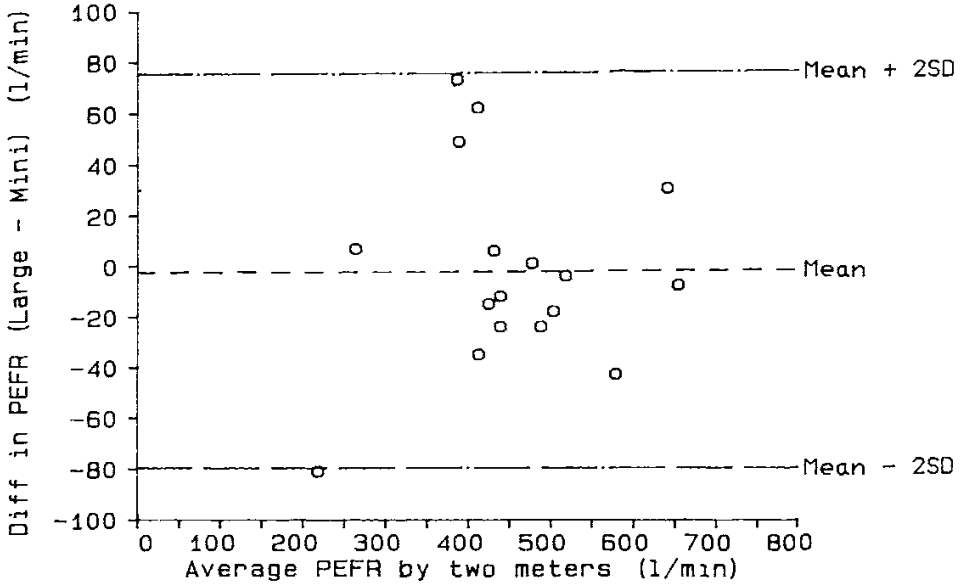
Wie alles begann . . .

Bland, Altman (1986). Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet*

- Zitationen: 12 696 (15. 4.), 12 711 (15.5.)
- **Vergleich neuer mit etablierter Methode:** Korrelationskoeffizient ist irreführend:
  - r misst linearen Zusammenhang, nicht Übereinstimmung
  - Übereinstimmung ist skalenabhängig, aber nicht r
  - r hängt vom Bereich ab: Wenn Stichprobe bzl. Wertebereich in zwei Gruppen geteilt wird, ist Korrelation für gesamte Stichprobe grösser als für beide Gruppen. Wäre für Übereinstimmung absurd.
  - Test auf Signifikanz ist irrelevant ($p < 0.001$).
  - Messungen mit hoher Korrelation ($r = 0.94$) können schlechte Übereinstimmung verbergen.

Wie alles begann . . . (cont.)



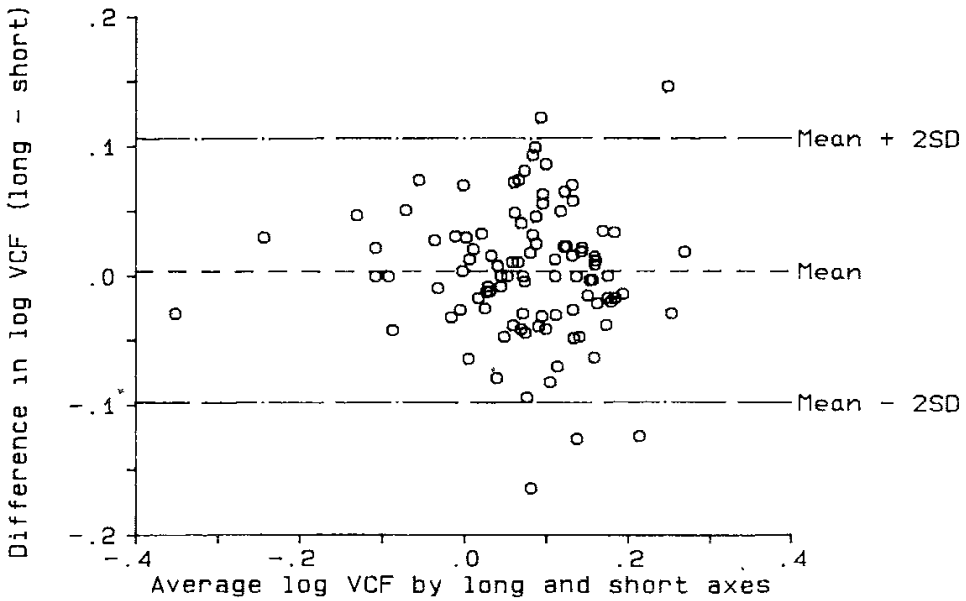


- $d_j = Y_{1j} - Y_{2j}$, $s = \mathrm{SD}(d)$
- $\bar{d}$: "bias"
- $\bar{d} \pm 2s$: "limits of agreement"
- $\mathrm{SE}(\bar{d}) = s/\sqrt{n}$, $\mathrm{SE}(s) = s/\sqrt{2n}$ $\longrightarrow$ $\mathrm{SE}(\bar{d} \pm 2s) \approx s\sqrt{3}/\sqrt{n}$

Wie alles begann . . . (cont.)

Bland & Altman diskutieren auch:

- logarithmische Transformation (einzig sinnvolle)



- Wiederholbarkeit (nehmen E$d = 0$ an)
- Formel für limits of agreement für 2 Methoden bei je 2 Messungen (zeigen aber keinen Plot)

Wie macht man einen Bland–Altman Plot?

1. Wir sind keine Graphiker!

Michel, Rake, Götzmann, Seifert, Ferrazzini, Chaoui, Treiber, Keller, Marincek, Kubik–Huch (2002). Pelvimetry and patient acceptability compared between open 0.5–T and closed 1.5–T MR systems. *European Radiology*.

Table 2 Open vs closed system pelvimetry (cm). *ISD* interspinous; *ITD* intertuberous; *OC* obstetric conjugate; *SO* sagittal outlet; *TD* transverse diameter

Measure		Mean±SD	Minimum	Maximum	Δ±SD
OC	Open	11.7±0.8	10.0	12.9	–0.0±0.2
	Closed	11.8±0.8	10.5	13.0	
SO	Open	11.4±0.9	9.8	13.3	–0.1±0.2
	Closed	11.5±0.9	10.0	13.4	
TD	Open	13.0±0.9	11.1	14.7	–0.1±0.2
	Closed	13.1±0.9	11.0	14.7	
ISD	Open	10.9±0.9	8.6	12.1	–0.1±0.3*
	Closed	11.0±0.8	8.9	12.8	
ITD	Open	12.2±1.0	10.1	13.8	–0.2±0.3*
	Closed	12.4±1.0	10.1	13.9	

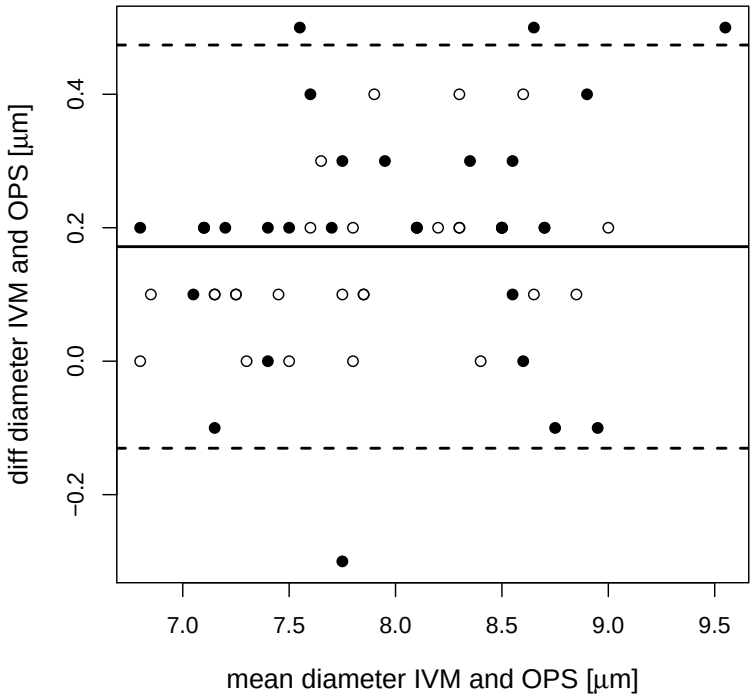
*Significant difference between the two methods p=0.01

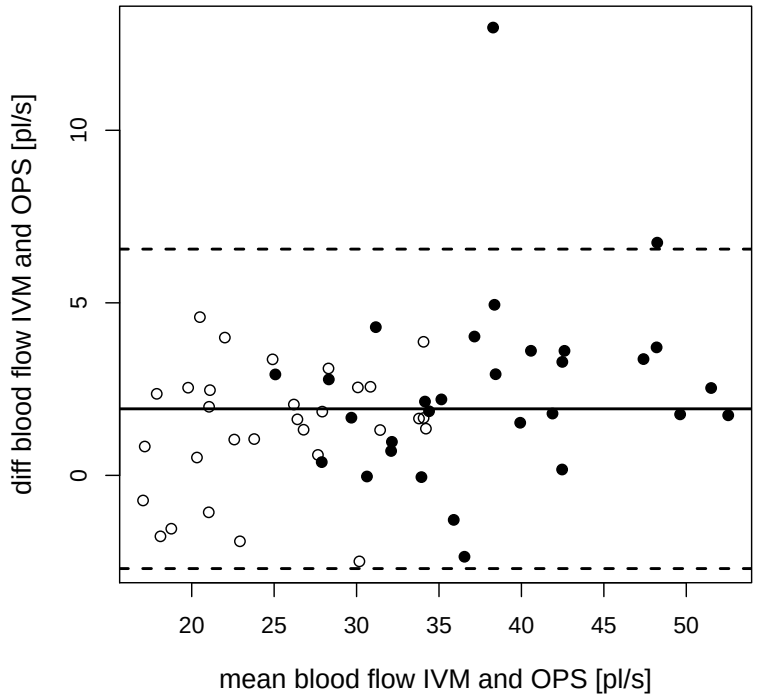
Wie macht man einen Bland–Altman Plot? (cont.)

2.

- kürzliche Beratung: "Vergleich der orthogon. Spektrographie mit der intravitalen Fluoreszenzmikroskopie zur Erfasssung des Reperfusionsschadens am Periost". Kunden hatten ein Jahr verzweifelt nach Graphiker gesucht.

- 2 Messmethoden (OPS, IVM), 2 Messpunkte (Baseline "o", 1 h Reperfusion "•"), 4 Parameter (FCD, Diameter, RBC-Velocity, Blood flow), 6 Tiere, je 5 Messungen





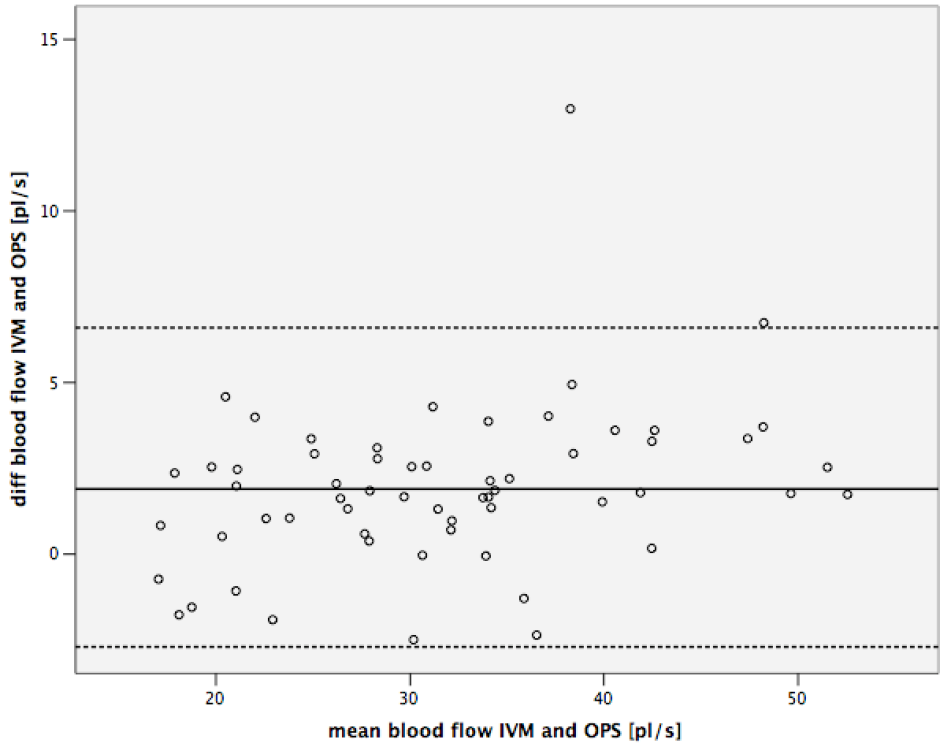
Wie macht man einen Bland–Altman Plot? (cont.)

3. **SPSS Geht nicht !**

Wie macht man einen Bland–Altman Plot? (cont.)

3. **SPSS Geht doch!**

- SPSS 13, nur Gallery, "Add a reference line to the Y axis"



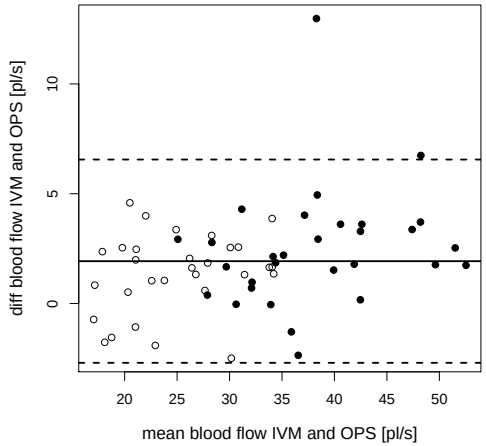
- auch bei overlay scatter

Formel für Bias und Limits of Agreement

- Darf ich die Clusterung ignorieren?

$$Y_{ijk\ell} = \mu_i + a_j + b_k + c_{ij} + e_{ik} + f_{jk} + g_{ijk} + \varepsilon_{ijk\ell}$$

$i = 1, 2$ Methode
$j = 1, 2$ Messpunkt
$k = 1, 6$ Tier
$\ell = 1, 5$ Nr. innerhalb (ijk)



($d_{jk\ell}$ unklar: Frage, ob Reihenfolge ℓ wichtig; sonst 5 statt einer Differenz?)

$$d_{jk\ell} = \text{bias} + (c_{1j} - c_{2j}) + (e_{1k} - e_{2k}) + (g_{1jk} - g_{2jk}) + (\varepsilon_{1jk\ell} - \varepsilon_{2jk\ell})$$

$\mathrm{E}(\bar{d}) = \mathrm{E}(d_{jk\ell}) = \text{bias}$ **Clusterung kann ignoriert werden!**

$$\mathrm{V}(d_{jk\ell}) = 2\left(\sigma_c^2 + \sigma_e^2 + \sigma_g^2 + \sigma_\varepsilon^2\right)$$

$\mathrm{E}(s^2) = 1.0\sigma_c^2 + 1.7\sigma_e^2 + 1.9\sigma_g^2 + 2\sigma_\varepsilon^2$ **zu eng!**

Halten Messpunkte und Tiere fest
$\longrightarrow$ s^2 ist richtiges Mass der Variabilität um Bias

Konfidenzintervalle für Bias und Limits of Agreement

1. Ignorieren!

2. SPSS und lmer{lme4} liefern asymptotische Kovarianzmatrix der Varianzkomponentenschätzer (REML).

$\widehat{\sigma} \sim N(\sigma, \Sigma)$

V = V($d_{jk\ell}$) = 2 * Summe der Varianzkomponenten, die „Methode“ enthalten

$V = a^T \sigma$

limits of agreement:

bias $\pm$ $2\sqrt{V}$

Var(V) = 4 * Summe aller Kovarianzen, die zweimal „Methode“ enthalten

$\text{Var}(a^T \widehat{\sigma}) = a^T \Sigma a$

$\text{Var}(f(x)) = (f'(\mu))^2 * \text{Var}(x)$

$\text{Var}(\sqrt{V}) = 1/(4 * V) * \text{Var}(V)$

$\left(x^{1/2}\right)' = \frac{1}{2} x^{-1/2}$

Grenzen des CI für limit of agreement:

$\text{SE(limit)} = \sqrt{\text{Var(bias)} + 4\text{Var}(\sqrt{V})}$

Warum erzähle ich das hier alles?

Meier, Schramm, Holstein, Seifert, Trentz, Menger (2006). Measurement of compartment pressure of the rectus sheath during intra–abdominal hypertension **in rats**. *Intensive Care Medicine*

$\longrightarrow$

Benninger, Laschke, Cardell, Holstein, Seifert, Keel, Trentz, Menger, Meier (2008). Compartment pressure of the rectus sheath accurately reflects intra-abdominal pressure **in a porcine model**. *submitted*

$\longrightarrow$ Referenz:

Malbrain et al. (2006) Results from the International Conference of Experts on Intra-abdominal Hypertension and Abdominal Compartment Syndrome. I. Definitions. *Intensive Care Medicine*

$\longrightarrow$ Web-page: `http://www.wsacs.org/research_recommendations.php`

WSACS (World Society of the Abdominal Compartment Syndrome) Recommendations for Research

• Statistical considerations

Descriptive statistics of the IAP values in the study population should be provided. The measurements should be compared using **Bland and Altman statistics, and bias with 95% confidence interval, including the lower and upper limits of agreement** should be reported. A **graphical representation** of the data is needed to evaluate performance across the range of IAP in the study group. **Correlation coefficients and scatterplots** may be added. A new method can considered to be acceptable only if **at least 20 relevant patients** have been studied, if **the bias** (the difference between two measurements) **does not exceed 1mmHg** (either positive or negative) and the **precision** (the standard deviation of the bias) **is not higher than 2mmHg**. The **percentage difference** (defined as the precision divided by mean IAP) should also be provided, and be **no higher that 25%** (. . .). The **number of measurements per patient** should be reported, and if this is not the same in all patients, appropriate statistical correction is needed.

- 7 Tiere, Druck stufenweise in 14 Stufen erhöht und wieder reduziert (2 Perioden)

Fig. 1. Bland-Altman plot for P_{Insuff} and corresponding CPRS for all IAP between physiological pressure and 30mmHg. Mean P = (P_{Insuff} +CPRS)/2; difference P_{Insuff}-CPRS. Between 25 and 28 measurements per animal were performed, depending on the initial physiological IAP. This resulted in a total of 188 measurements. *Lines* Lower and upper limit of agreement (bias±1.96*standard deviations)

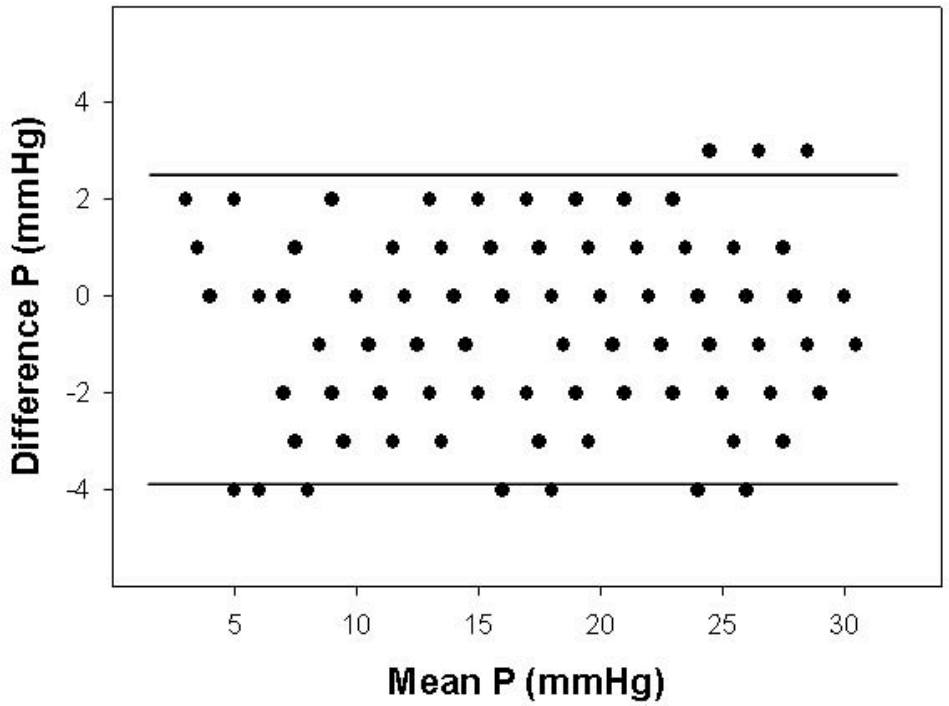


Table T2. Analysis of direct IAP measurement (P_{Insuff}) and different indirect IAP measurement techniques for all IAP between physiological pressure and 10mmHg. Between 5 and 8 measurements per animal were performed, resulting in a total of 48 measurements.

	CPRS vs. P_{Insuff}	FVP vs. P_{Insuff}	IVP vs. P_{Insuff}	Reference [11]
Mean IAP (mmHg)	8.1	7.9	7.0	
Bias (mmHg)	-1.3	-1.0	0.7	-1 to +1
Precision (SD, mmHg)	1.7	1.5	1.1	2
Lower LA (mmHg)	-4.7	-3.8	-1.6	Bias -4
Upper LA (mmHg)	2.1	1.9	2.9	Bias +4
Percentage error [8]	42%	36%	32%	<25%
COVA	0.21	0.18	0.16	
R^2 within animals	0.759	0.846	0.897	

IAP intra-abdominal pressure, P_{Insuff} automatically controlled insufflator pressure, *CPRS* compartment pressure of the rectus sheath, *FVP* femoral vein pressure, *IVP* intra-vesical pressure, *SD* standard deviation=precision, *LA* limit of agreement (Bias± (1.96*SD)), *COVA* coefficient of variation (*SD*/Mean IAP).