Considerations for the statistical evaluation of cross-validation data following ICH M10 recommendations

One year into ICH M10 + EBF OS Barcelona 2023

Kamil Sklodowski, BA Manager (presenter)
Katja Heinig, BA Manager
Eginhard Schick, Scientific Expert
Alex Georgiou, Scientific Expert
Matthew Barfield, Chapter Lead

v1.02

Cross-validation

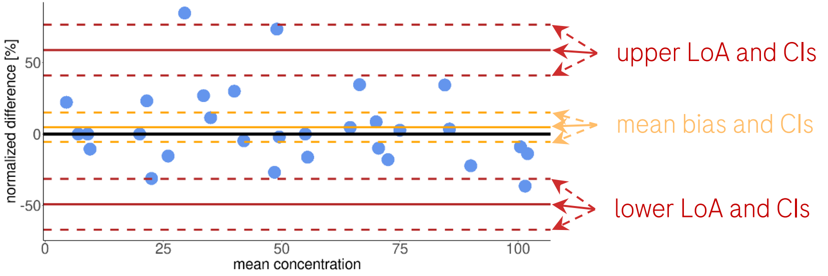
6.2. Cross validation - page 27/45

Bias can be assessed by **Bland-Altman** plots or **Deming regression**. Other methods appropriate for assessing **agreement** between two methods (e.g., **concordance correlation coefficient**) may be used too. Alternatively, the **concentration vs. time curves** for study samples could be plotted for samples analysed by each method to assess bias.

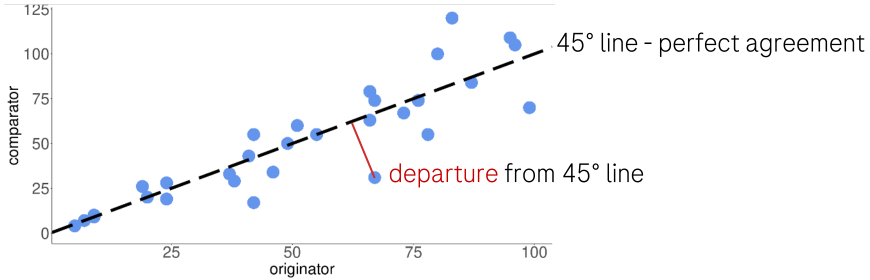
- Key takeaway is to assess:
  - agreement between results (no reference dataset)
- Using family of methods:
  - descriptive (e.g Bland-Altman)
  - regressions (e.g. Deming, Passing-Bablok)
  - indices (e.g. Lin's CCC)

Parameters and criteria relevant for BA and CCC

- Bland-Altman (BA):
  - Limits of agreement (LoA) and mean difference (bias):
    - Outliers effect on both estimates
    - Normality assumptions
    - Point estimate and interval calculations
    - Criteria focused on:
      - Predefined relevant limits (e.g. %diff)
      - Mean bias and the corresponding CIs



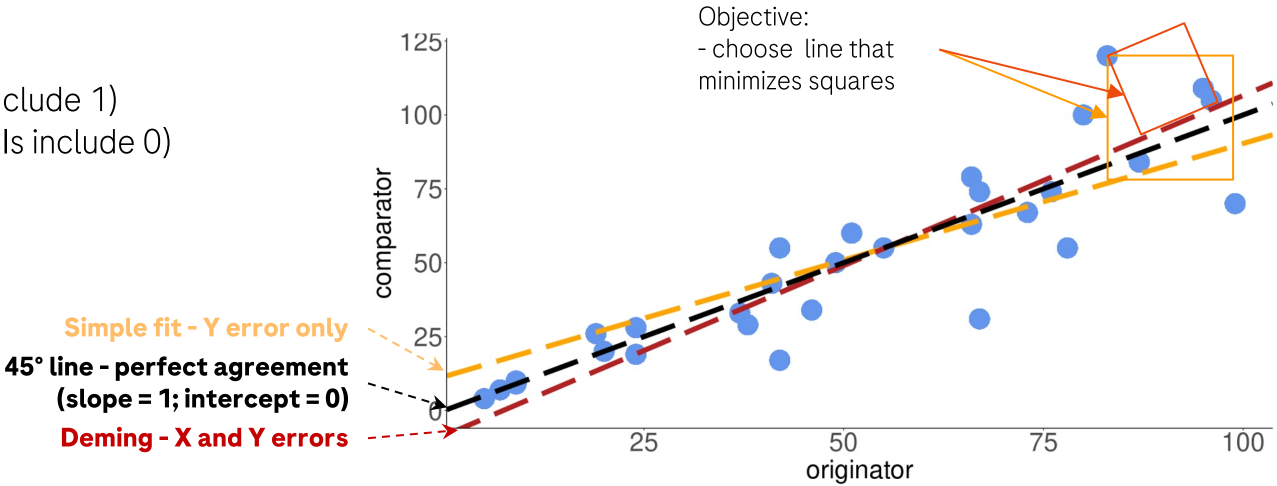
- Lin's Concordance Correlation Coefficient (CCC):
  - Effect of inter-subject variability
  - Criteria focused on:
    - CCC above a selected threshold (e.g. lower CI of the estimate ⩾ 0.9)
  - Alternative criteria for:
    - Precision (ρ)
    - Scale shift (σ)
    - Location shift (μ)



Silveira et al., 2023; Frey et al., 2020; Grossmann, 2020; Shieh, 2018; Jan & Shieh, 2018; Giavarina, 2015; Watson & Petrie, 2010; Myles & Culi, 2007; Barnhart et., 2006
R references: Potapov et al., 2023; Signorell et al., R Core Team 2022; Wickham 2016

Parameters and criteria relevant for selected regression models

- Deming and Passing-Bablok regressions:
  - Ratio of errors
  - Weights
  - Effect of outliers and normality (not applicable for Passing Bablok)
  - Parametrization method for slope and intercept estimates
  - CI calculation methods (e.g. bootstrap, jacknife, analytical)
  - Linear relationship
  - Criteria focused on:
    - slope (e.g. CIs include 1)
    - intercept (e.g. CIs include 0)

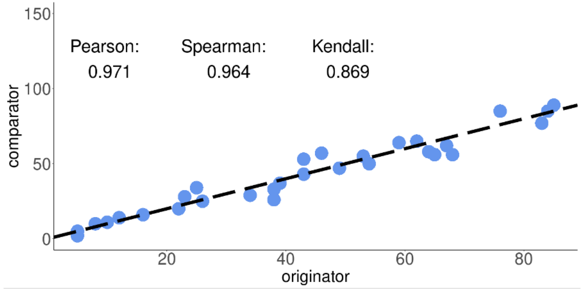


Baumdicker & Hölker 2023; Geistanger, 2020; Therneau, 2018; Payne, 1997; Linnet 1990; Passing & Bablok, 1983; Deming, 1943
R references: Potapov et al., 2023; Signorell et al., R Core Team 2022; Wickham 2016

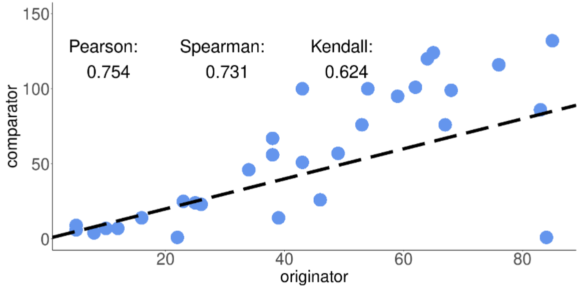
in silico cases

- Random sample set (n=30); 5 [LLOQ] to 100 [ULOQ] (x-axis) = **originator**
- Generation of additional datasets with caveats (y-axis) = **comparator** (**A-D**)
- Direction of change plays no role;negative is used as an example

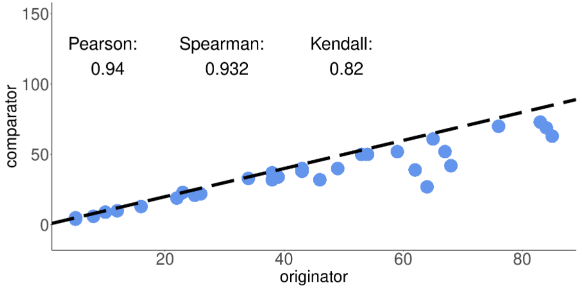
(**A**) no bias, var ~ 20% of the originator



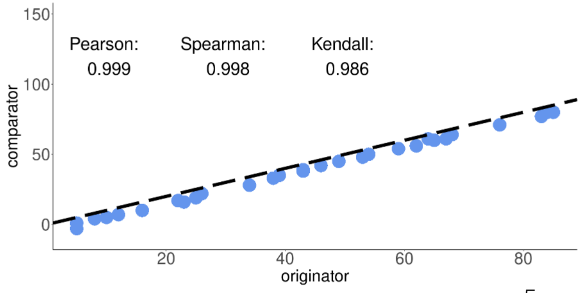
(**B**) no bias, var ~ 60% of the originator



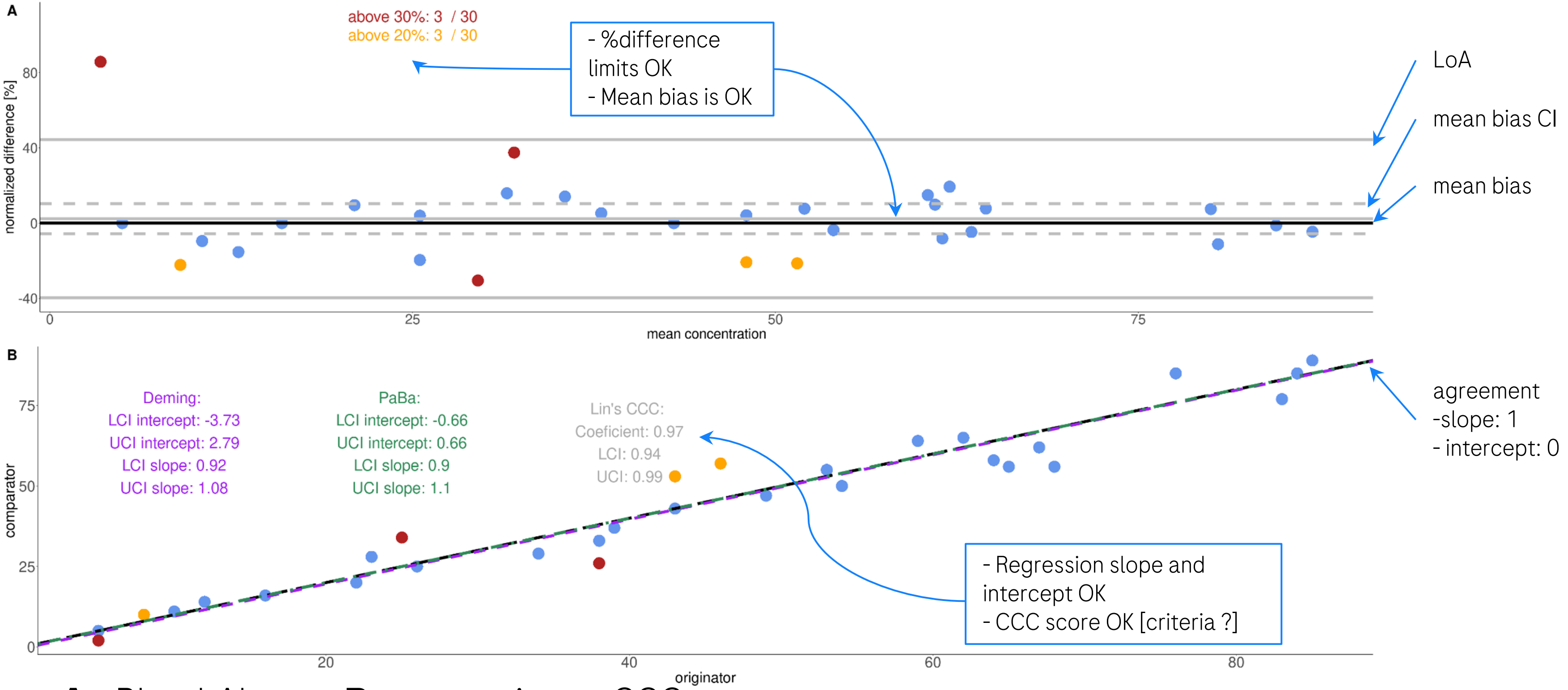
(**C**) negative proportional bias (~20%)



(**D**) negative non-proportional bias (~LLOQ)

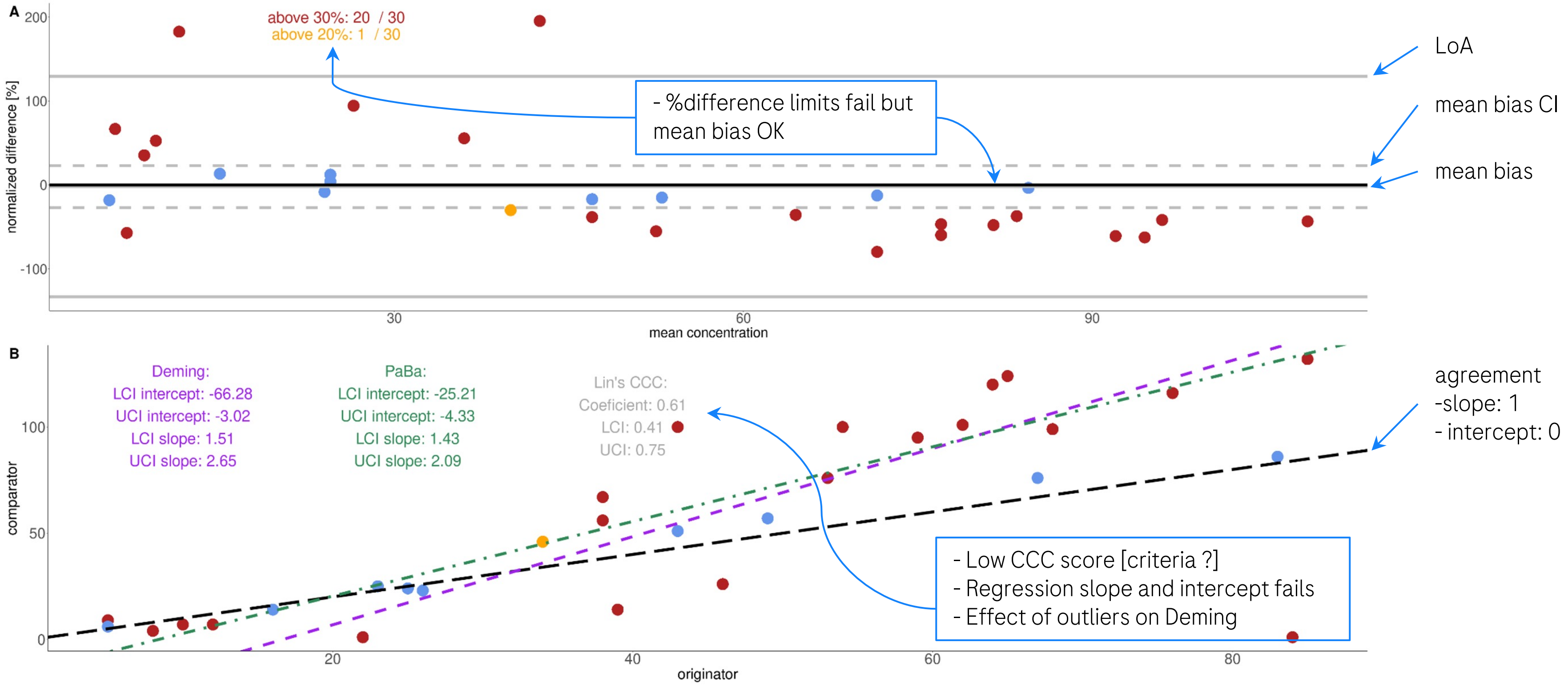


Results of *in silico* case A (no bias, var ~20%)



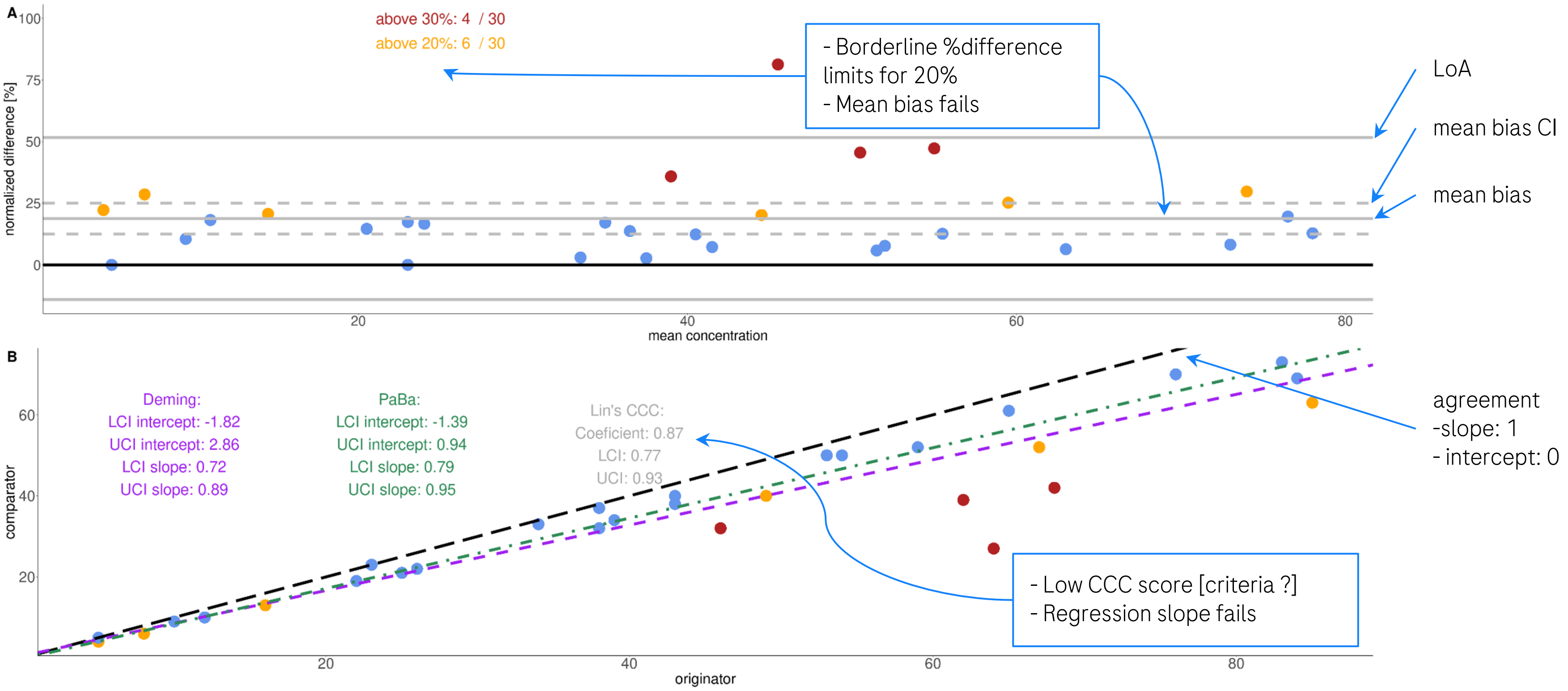
A - Bland-Altman; **B** - regressions + CCC

Results of *in silico* case B (no bias, var~60%)



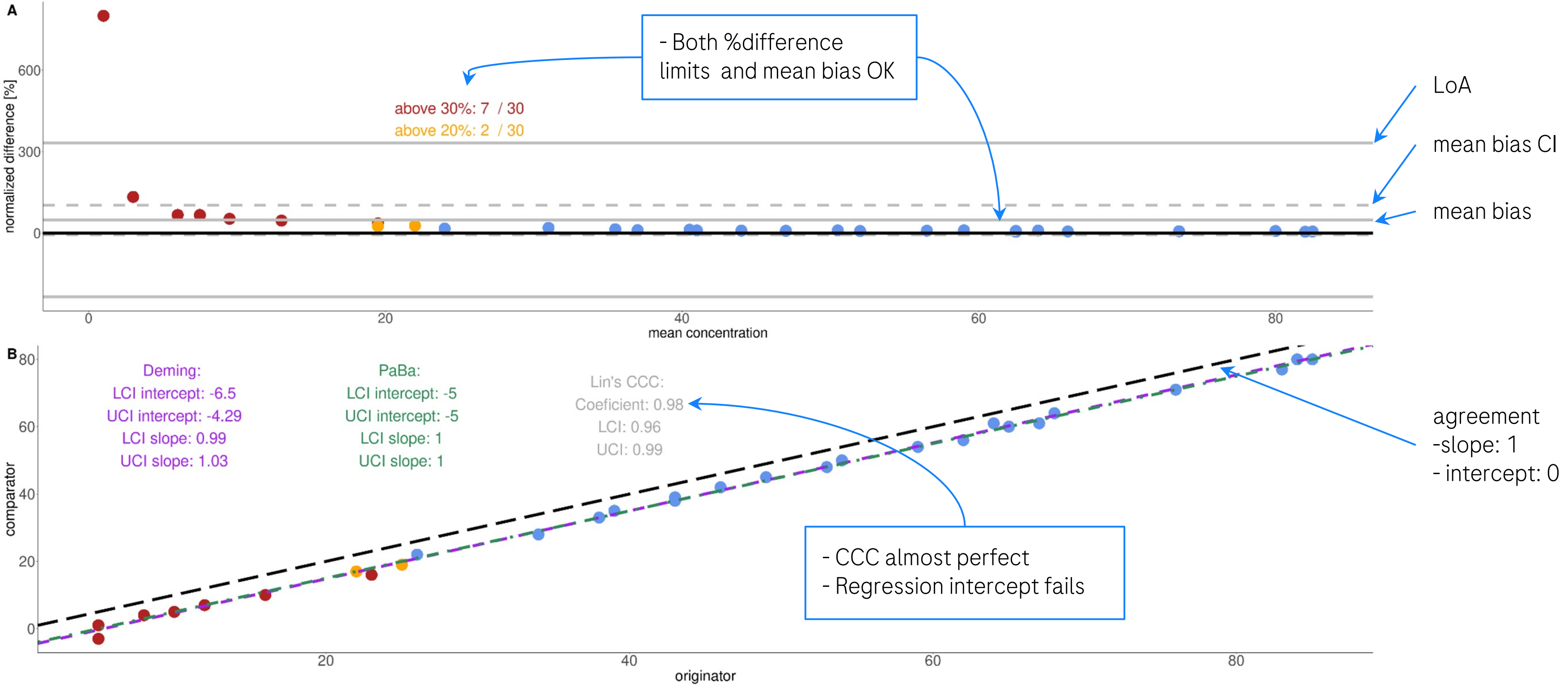
A - Bland-Altman; **B** - regressions + CCC

Results of *in silico* case C (proportional bias, ~20%)



A - Bland-Altman; **B** - regressions + CCC

Results of *in silico* case D (non-proportional bias, ~LLOQ)



A - Bland-Altman; **B** - regressions + CCC

Real-life examples with variable approaches

	Original criteria applied	Molecule type	Data used for evaluation	Conclusions
Case 1	**ISR like (with 30%)**	LM	**Combined QC and study samples**	Passed (67%)
Case 2	ISR like (with 30%)	LM	Combined QC and study samples	Failed
Case 3	**ISR like (with 20%) & nominal for QCs**	SM	**Separate QC and study samples**	Samples fail; QCs pass
Case 4	ISR like & nominal for QCs	SM	Separate QC and study samples	Samples pass; QCs pass
Case 5	**ISR like (with 20%) & nominal for QCs**	SM	**Only QC samples** (low n = 24)	Passed
Case 6	ISR like (with 20%) & nominal for QCs	SM	Separate QC and study samples	Samples fail; QCs pass
Case 7	ISR like & nominal for QCs	SM	Separate QC and study samples	Samples pass; QCs pass
Case 8	ISR like & Nominal for QCs	SM	Separate QC and study samples	Samples pass; QCs pass
Case 9	ISR like (with 30%)	LM	Combined QC and study samples	Failed
Case 10	ISR like (with 20%) & nominal for QCs	SM	Separate QC and study samples	Samples fail; QCs pass

Samples and criteria applied for evaluation of selected methods using real-life examples

	Samples used in analysis	Criteria for BA	Criteria for Deming, Passing-Bablok and Lin's CCC
Case 1	**Combined** QC and study samples	**Pre-specified %difference limits criteria following ISR to replace LoA:** - Must contain 67% of observations below 30% for LM and below 20% for SM **Mean %difference/bias:** - CIs for mean bias must include "0"	**Slope:** - CIs must include "1" **Intercept:** - CIs must include "0" **CCC:** - Lower CI ⩾ 0.9 (arbitrary)
Case 2			
Case 3			
Case 4			
Case 5	QC samples		
Case 6	Study samples		
Case 7	**Combined** QC and study samples		
Case 8			
Case 9			
Case 10	Study samples		

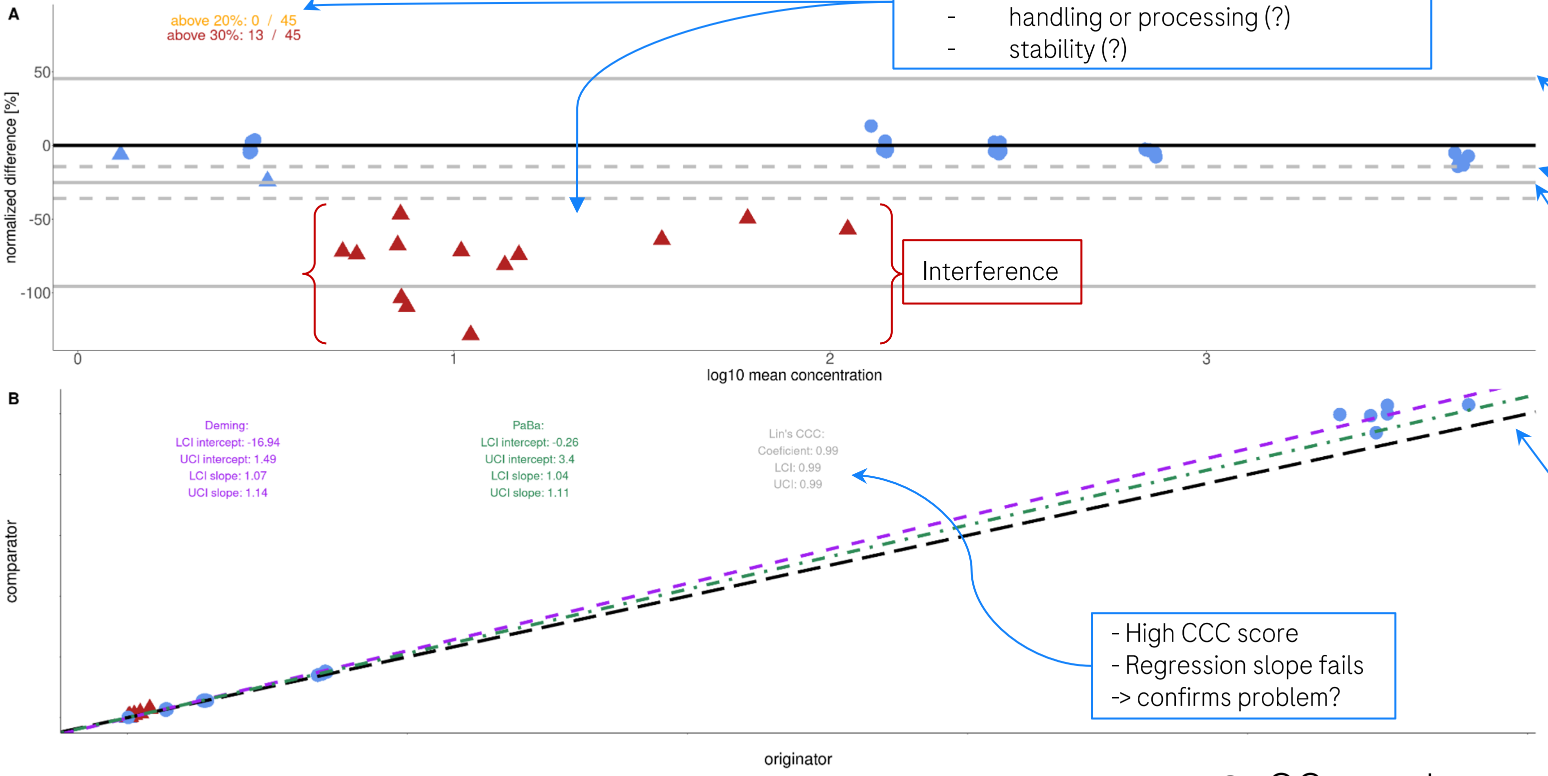
Performance summary - failing rate

	Original decision	Bland-Altman	Lin's CCC	Deming regression	Passing- Bablok regression	Simple majority vote (> 50% failed parameters)
Case 1	Passed	%difference: Passed Mean bias: **Failed**	Passed	Slope: **Failed** Intercept: Passed	Slope: **Failed** Intercept: Passed	Passed (2/5)
Case 2	**Failed**	%difference: Passed Mean bias: **Failed**	**Failed**	Slope: Passed Intercept: **Failed**	Slope: **Failed** Intercept: Passed	**Failed (3/5)**
Case 3	**Failed**	%difference: Passed Mean bias: **Failed**	Passed	Slope: **Failed** Intercept: Passed	Slope: **Failed** Intercept: Passed	**Passed (2/5)**
Case 4	Passed	%difference: Passed Mean bias: **Failed**	Passed	Slope: **Failed** Intercept: Passed	Slope: **Failed** Intercept: Passed	Passed (2/5)
Case 5	Passed	%difference: Passed Mean bias: **Failed**	Passed	Slope: **Failed** Intercept: Passed	Slope: **Failed** Intercept: Passed	Passed (2/5)
Case 6	Passed	%difference: Passed Mean bias: Passed	Passed	Slope: Passed Intercept: Passed	Slope: Passed Intercept: **Failed**	Passed (0/5 or 1/5)
Case 7	Passed	%difference: Passed Mean bias: **Failed**	Passed	Slope: **Failed** Intercept: **Failed**	Slope: Passed Intercept: Passed	**Failed or Passed (3/5 or 1/5)**
Case 8	Passed	%difference: Passed Mean bias: **Failed**	Passed	Slope: Passed Intercept: Passed	Slope: Passed Intercept: Passed	Passed (1/5)
Case 9	**Failed**	%difference: **Failed** Mean bias: **Failed**	Passed	Slope: **Failed** Intercept: **Failed**	Slope: **Failed** Intercept: **Failed**	**Failed (4/5)**
Case 10	**Failed**	%difference: Passed Mean bias: **Failed**	**Failed**	Slope: **Failed** Intercept: **Failed**	Slope: **Failed** Intercept: Passed	**Failed (4/5 or 3/5)**
Summary of failed cases	**4**	**9**	**2**	**8**	**7**	**3**

Why?

Example 1: case study 3

- Mean bias fails but %difference limits are OK
- Why only study samples fail?
    - handling or processing (?)
    - stability (?)



LoA

mean bias CI

mean bias

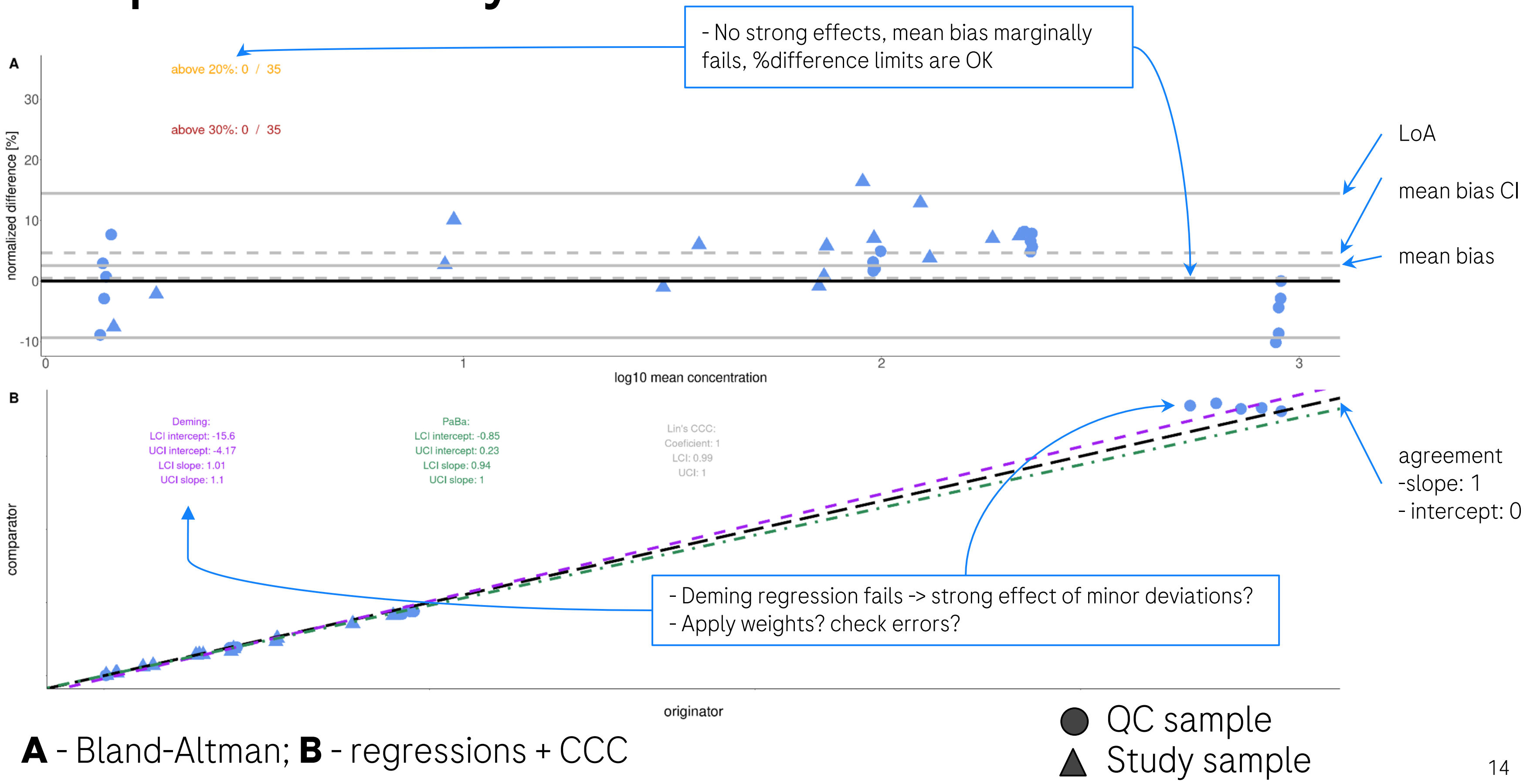
agreement
- slope: 1
- intercept: 0

- High CCC score
- Regression slope fails
-> confirms problem?

A - Bland-Altman; **B** - regressions + CCC

● QC sample
▲ Study sample

Example 2: case study 7



A - Bland-Altman; **B** - regressions + CCC

Summary and Conclusions

- Understand your data - visualize and investigate
- Know your tools - limitations (e.g. outlier effect), corrections (e.g. weights)
- No approach is perfect - there are no one-fits-all methods but a combination of the right ones can get the job done!

Poster:
Heinig et al., 2023

Take home message

"Which method is best?
(...) Essentially, if the data is good, all the methods will agree on that fact. If there are assay issues, outliers in particular, then the actual source of the problem needs to be investigated rather than just using a "better" regression tool. ***Understanding data requires more than pushing a button.****"*
~ Terry M. Therneau

Acknowledgements

- Eginhard Schick, Katja Heinig, Alex Georgiou
- Matthew Barfield
- Entire Roche RBBO and BioAM Chapters
- EBF Organizing Committee
  and the entire EBF Community - yes it's You!

- All past, present and future Mentors
  - no journey should be walked alone!

Doing now what patients need next