Comparing ABL and numerical eGFR

Christoph Jäggli, Alexander Leichtle

October 2019

1 Introduction

TODO: *outwrite GFR: (Wikipedia) Glomerular filtration rate (GFR) describes the flow rate of filtered fluid through the kidney.*

2 Method

2.1 Data

Data from 780 patients (16-97 years, 340 female and 440 male) was used for the analysis. All patients were consecutively referred for determination of GFR by two different machines, henceforth called ABL and COBAS, respectively. After determining the serum creatinine level, they both used the estimation given by the CKD-EPI formula of Levey et al. [2009]

According to TODO: *put ref* we group the GFR values into six groups G1, G2, G3a, G3b, G4, and G5 as shown in table 1. Table 2 shows some basic statistics of the estimated GFR values from the ABS and the COBAS method, respectively.

Table 1: GFR-categorization of the chronically kidney diseases.

Category	**Description**	**Range** ($\mathrm{ml}/\min/1.73\mathrm{m}^2$)
G1	Normal	90+
G2	Mildly decreased	60-89
G3a	Mildly to moderately decreased	45-59
G3b	Moderately to severely decreased	30-44
G4	Severely decreased	15-29
G5	Kidney failure	<15

Table 2: Basic statistics for the estimated GFR values by ABL and COBAS

	ABL	**COBAS**
Min	5	5
Max	171	165
Mean	89.93	89.13
Std	28.29	28.11
CV	0.3146	0.3154
p_{G1}	0.5603	0.5474
p_{G2}	0.2949	0.3051
p_{G3a}	0.0564	0.059
p_{G3b}	0.059	0.059
p_{G4}	0.0167	0.0179
p_{G5}	0.0128	0.0115

2.2 Evaluation

The GFR estimation of Levey et al. [2009] is given by

$$\mathrm{GFR} = \beta \min(S_{cr}/k, 1)^{\alpha} \max(S_{cr}/k, 1)^{-1.209} 0.993^{age} * 1.018(\text{if female})$$

where $\alpha = -0.329, k = 0.7$ for females and $\alpha = -0.411, k = 0.9$ for males. The constant β is defined based on the origin and the sex of the patient as in table 3.

Table 3: Definition of β based on the origin and the sex of the patient.

sex	**black**	**white or other**
female	$\beta = 166$	$\beta = 144$
male	$\beta = 163$	$\beta = 141$

3 Results and Discussion

3.1 Correlation

The correlation between two variables X and Y (also called Pearson's Correlation Coefficient (PCC) TODO: *put ref here*) is defined as

$$\rho_{X,Y} = \frac{\mathrm{cov}(X,Y)}{\sigma_X \sigma_Y}$$

where $\text{cov}(X, Y)$ is the covariance between X and Y and σ_X (resp. σ_Y) denotes the standard deviation of X (resp. Y).

Assuming that the two samples X and Y are drawn from independent normal distributions, the probability density function of the sample correlation coefficient ρ reads [Student, 1908]

$$f(\rho) = \frac{(1-\rho^2)(n/2-2)}{B(1/2, n/2-1)} \tag{1}$$

where $B(\cdot, \cdot)$ is the beta function and n the number of samples. This function can be used to approximately compute a confidence interval as well as probability value (p-value) that uncorrelated and normally distributed samples give a correlation values at least as extreme as $\rho_{X,Y}$.

The GFR values described above produced correlation coefficient of $\rho_{X,Y} = 0.9908$ with p-value $< 10^{-10}$ and confidence interval between 0.9894 and 0.992.

3.2 Regression

Deming Regression (DR) is commonly used for fitting a linear spline to two-dimensional samples where both variables, X and Y, are measured with error TODO: *put ref here*. DR assumes that the error ratio (denoted by λ) is constant. For the following analysis we assumed the error ratio to be equal to 1 in which case DR gives the same result as orthogonal regression. The left figure of 1 shows the Deming regression (dashed red) and its uncertainty region (magenta) between the GFR values of ABL and COBAS, while the right part shows the corresponding Bland-Altman-Plot. TODO: *Describe the plot's in more detail*

3.3 Prediction

The Deming regression line from figure 1 can be used for predicting the G-category (according to table 1) for eGFR (CKD-EPI) from the values of GFR (Opus ABL). Figure 2 shows the confusion matrix that was obtained by a 10-fold cross-validation technique.. The x-axis shows the predicted and the y-axis the true categories. In total, there are 725 ($p = 0.9295$) subjects that are classified correctly what leaves 55 ($p = 0.0705$) misassignmenets. Among the defective classifications, there are 35 objects that are assessed 1 class too severe and 25 objects that are assessed 1 class too mild. We observe that both, the true and the predicted distribution, highly emphasize the categories G1 and G2. This becomes even more clear in Table 4, where assignment frequencies for both distributions are presented.

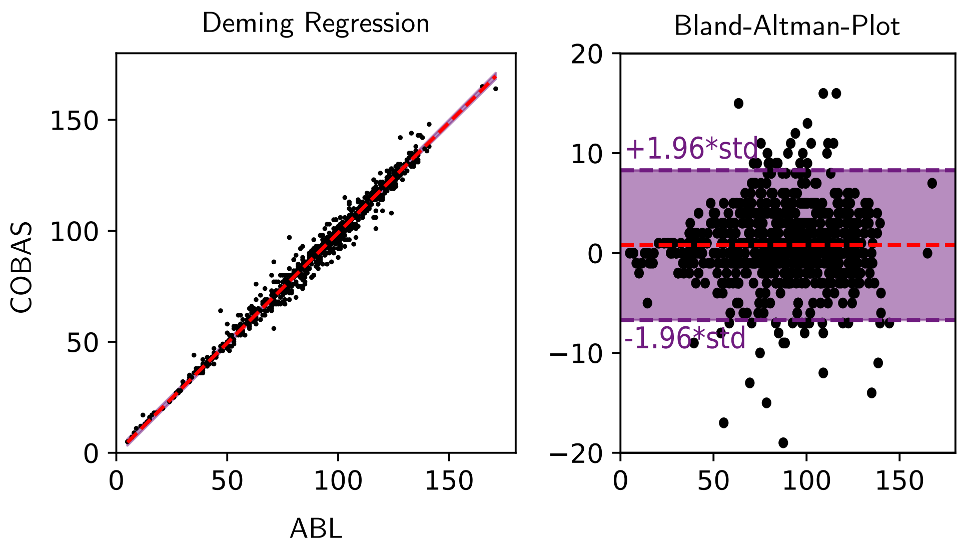

Figure 1: Deming regression line and the corresponding Bland-Altman-Plot of the two data sets.

3.4 Agreement measures

Cohen's and Fleiss' kappa are two similar measures that can be used to express the inter-rater agreement for categorical indicators. They are thought to be more robust reliability measures than simple percent agreement calculation, because they take into account that some agreements might occur purely by chance (c.f. baseline agreement) TODO: *put ref here; explain the measures in more detail*. It is commonly agreed that measures larger than 0.8 - 1 express very to almost perfect agreement. Table 5 shows the agreement measure values for the predicted (DR from ABS) and the true

Table 4: Assignment frequencies for the true and the predicted categories.

Category	**True**		**Predicted**	
G1	427	($p = 0.5474$)	428	($p = 0.5487$)
G2	238	($p = 0.3051$)	231	($p = 0.2961$)
G3a	46	($p = 0.059$)	50	($p = 0.0641$)
G3b	46	($p = 0.059$)	43	($p = 0.0551$)
G4	14	($p = 0.0179$)	16	($p = 0.0205$)
G5	9	($p = 0.0115$)	12	($p = 0.0154$)

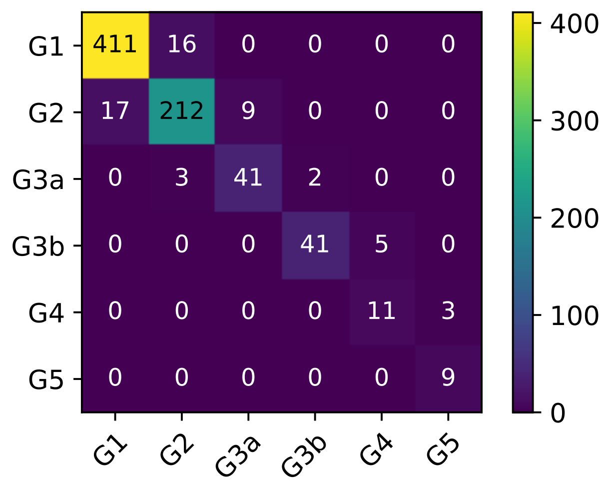

Figure 2: 10-fold cross-validation confusion matrix for the G-category prediction by a Deming regression line.

(COBAS) GFR values.

Table 5: Correlation measures when predicting the G-category of eGFR (CKD-EPI) from the value of GFR (Opus ABL) using a 10-fold cross-validation scheme.

Measure	Agreement
Baseline	0.3983
Cohen's κ	0.8828
Fleiss' κ	0.881

3.5 Variation of COBAS values

In the technical datasheet of the COBAS measurement device they report an empirical test for assessing the variation coefficient of the observed serum creatinine level. For this they used fast repeated measurements ($n = 21$) as well as daily measurements (21 days) at two different levels. The results are reported in Table 6 These variation coefficients can be used to construct a similar categorization as reported above by taking the error of the creatinine measurement into account. For this, we perturbed the 780 creatinine measurements from the COBAS tool by a Gaussian error with standard deviation corresponding tho the variation coefficient of the daily measures at

Table 6: Empirical test for measuring the variation coefficient of the COBAS method for evaluationg the serum creatinine level

Repeated Measures			**Daily Measures**		
μ (μmol/L)	σ (μmol/L)	VC (%)	μ (μmol/L)	σ (μmol/L)	VC (%)
86.2	0.7	0.8	94.9	1.4	1.4
353	2	0.6	338	4	1.1

the lower level (i.e. 1.4%). After using again the GFR estimation from the CKD-EPI equation, we compared the obtained kidney disease categories to the original ones. This results in the confusion matrix in figure 3 and their corresponding probability measures in table 7

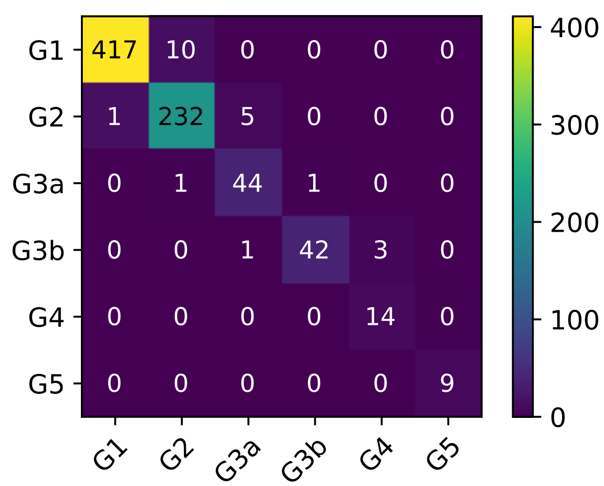

Figure 3: Confusion matrix of the estimated GFR values that are computed from the perturbed creatinine levels measured by the COBAS device.

Again, the x-axis shows the predicted and the y-axis the true categories. In total, there are 758 ($p = 0.9718$) subjects that are classified correctly what leaves 22 ($p = 0.0282$) misassignmenets. Comparing the two confusion matrices in figures 2 and 3 we clearly observe that the misassignmenets that are due to measurement uncertainties are significantly less frequent than the ones from the prediction method TODO: *does this mean that the prediction method should not be used; does this mean that the ABL method has a large measurement error?*

Table 7: Assignment frequencies for the true and the predicted categories for the perturbed COBAS measurements.

Category	**True**		**Predicted**	
G1	427	($p = 0.5474$)	418	($p = 0.5359$)
G2	238	($p = 0.3051$)	243	($p = 0.3115$)
G3a	46	($p = 0.059$)	50	($p = 0.0641$)
G3b	46	($p = 0.059$)	43	($p = 0.0551$)
G4	14	($p = 0.0179$)	17	($p = 0.0218$)
G5	9	($p = 0.0115$)	9	($p = 0.0115$)

4 Conclusion

References

Andrew S. Levey, Lesley A. Stevens, Christopher H. Schmid, Yaping (Lucy) Zhang, III Castro, Alejandro F., Harold I. Feldman, John W. Kusek, Paul Eggers, Frederick Van Lente, Tom Greene, Josef Coresh, and for the CKD-EPI (Chronic Kidney Disease Epidemiology Collaboration). A New Equation to Estimate Glomerular Filtration Rate. *Annals of Internal Medicine*, 150(9):604–612, 05 2009. ISSN 0003-4819. doi: 10.7326/0003-4819-150-9-200905050-00006. URL `https://doi.org/10.7326/0003-4819-150-9-200905050-00006`.

Student. Probable error of a correlation coefficient. *Biometrika*, pages 302–310, 1908.