COMPARABILITY OF METHODS AND ANALYSERS

Nora Nikolac

Zagreb, October 24-25, 2015

15th EFLM Continuing Postgraduate Course in Clinical Chemistry and Laboratory
When?

- Introducing new method or analyzer
- Multiple analytical systems in laboratory
- Using services of another laboratory
Why?

- To increase patient safety
  - To assure that method change is not going to influence laboratory result for the patient.
How?

- Experimental procedures following protocols
- CLSI EP09-A3: Measurement procedure comparison and bias estimation using patient samples

1. Number of samples
2. Measurement range
3. Time of analysis
4. Data analysis
5. Data interpretation
1. Number of samples

- Min: 40 samples
- Optimal: 100 samples
  - To identify unexpected errors from sample matrix or interferences
- Measurements in duplicate
2. Measurement range

- Cover 90% of the method measurement range

- Good agreement between methods

- Difference in higher concentration range
2. Measurement range

- Overlapping measurement range for both methods

- Method A determined using dilution protocol
- Method B reported as LOQ

Glucose concentration
3. Time of analysis

- Measurements done within 2 hours
  - Not for: glucose, lactate, ammonia, blood gasses testing...

- Measurements done over 5 days
  - Better over longer period of time

- Collecting samples over period of time (first method) and analyzing in batch using second method
4. Analyzing results

Several statistical approaches:

- Correlation
- Paired test for difference
- Linear regression
  - Deming regression
  - Passing-Bablok regression
- Bland-Altman analysis
4. Analyzing results

- Comparison of two methods for direct bilirubin concentration measurement

<table>
<thead>
<tr>
<th>Summary data</th>
<th>Method 1</th>
<th>Method 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Analyzer, method</td>
<td>Architect (Abbott) Diazo method</td>
<td>AU 680 (Beckman Coulter) DPD method</td>
</tr>
<tr>
<td>Min-Max</td>
<td>2.7-232.3</td>
<td>5.5-273.4</td>
</tr>
<tr>
<td>Mean ± SD</td>
<td>65.5 ± 67.9</td>
<td>82.4 ± 83.6</td>
</tr>
<tr>
<td>Median (IQR)</td>
<td>38.5 (7.9-127.8)</td>
<td>42.4 (11.1-158.2)</td>
</tr>
<tr>
<td>P (normality)</td>
<td>0.059</td>
<td>0.036</td>
</tr>
</tbody>
</table>
4.1 Correlation

- Spearman coefficient of correlation

\[ r (95\% \text{ CI}) = 0.97 (0.95-0.98) \]

Excellant correlation
What is the meaning of this result?

- Methods are significantly associated
- Linear relation between methods
- ↑ of Method A associated with ↑ of Method B
- Nothing about amount of increase!
4.2 Significance of difference

- Wilcoxon test (normality failed)

\[ P < 0.001 \]

Significant difference between methods
What is the meaning of this result?

- Calculating differences for each pair of measurement
- Comparing number of negative and positive differences
- If there is no difference between methods, number of differences is equal

More measurements were higher using Method 2
4.3 Linear regression

High correlation
Linear relationship

Equation to describe relationship between methods
Determine proportional and constant error

Deming regression
Passing and Bablok regression

Lessons in biostatistics

Comparison of methods: Passing and Bablok regression

Lidija Bilić-Zulle

Clinical Department of Laboratory Diagnostics, Clinical Hospital Centre and Department of Medical Informatics, Rijeka University School of Medicine, Rijeka, Croatia

Corresponding author: lidija.bilic-zulle@medri.hr

Linear regression

Regression equation:
\[ y = a + bx \]

Intercept:
\[ \text{Intercept} = a \]

g (α) = b

95% confidence intervals

Regression equation:
\[ y = a \text{ (95\% CI)} + b \text{ (95\% CI)} x \]
Constant and proportional error

Regression equation
\[ y = a \, (95\% \, \text{CI}) + b \, (95\% \, \text{CI}) \, x \]

Excluding 0
Constant error

Excluding 1
Proportional error

Intercept = \( a \)

\( \tan(\alpha) = b \)
Deming regression

- Includes analytical variability of both methods (CV)
- Assumes that errors are independent and normally distributed
- Both methods prone to errors

\[ y = 1.74 \text{ (-1.77 to 5.24)} + 1.23 \text{ (1.16 to 1.30)} x \]

- No constant error
- Proportional error
Passing-Bablok regression

- Non-parametric method
- No assumptions about distributions of samples
- No assumptions about distributions of errors
- Not sensitive to outliers
Why don’t we recalculate results?

Direct bilirubin (Method 2) = 1.23 x Direct bilirubin (Method 1)

Direct bilirubin (Method 1) = Direct bilirubin (Method 2) / 1.23
Residual analysis

- How well data fit to the regression model

\[ Y - F(x) \]

- Residuals: 
  - Positive residuals: residuals above the regression line, indicating underestimation.
  - Negative residuals: residuals below the regression line, indicating overestimation.

Graphs showing examples of residuals:
- Positive residuals: dots above the line, indicating consistent underestimation.
- Negative residuals: dots below the line, indicating consistent overestimation.

Correct model fit:

Incorrect model fit:
Residual analysis

Differences between measured and calculated values
4.3 Bland-Altman analysis

- Graphical method to compare two measurements technique
- Analyzing differences between measurement pairs

**Lessons in biostatistics**

**Understanding Bland Altman analysis**

Davide Giavarina

Clinical Chemistry and Hematology Laboratory, San Bortolo Hospital, Vicenza, Italy

Corresponding author: davide.giavarina@ulssvicenza.it

Mountain plot

More positive differences
4.3 Bland-Altman analysis

- Plotting differences against:
  - Mean of two methods (no reference method)
  - One method (reference method)

The diagram shows a scatter plot with differences calculated as

\[ \text{method 2} - \text{method 1} \]

against the mean of method 1 and method 2. The limits of agreement are indicated by the lines:

- Upper limit: +1.96 s
- Lower limit: -1.96 s

The 95% confidence interval is also shown.
LoA and mean difference

- **Absolute units**
  - **Constant bias**
  - Wide LoA = poor agreement
  - Including 0 = no constant bias
  - Excluding 0 = Constant bias
  - Narrow LoA = good agreement

- **Mean difference**
  - Mean of method 1 and method 2
  - Method 2 – method 1
  - 0

- **95% CI**
  - Limits of agreement
  - Limits of agreement
  - +1.96 s
  - -1.96 s

- **Mean of method 1 and method 2**

- **Absolute units**
  - Constant bias
  - Percentage (%)
  - Proportional bias

Including 0 = no constant bias
Excluding 0 = Constant bias
Narrow LoA = good agreement

Wide LoA = poor agreement
Bland-Altman analysis

Plotting against mean difference

No constant bias

Plotting against % difference

Proportional bias
5. Data interpretation

Statistical significance ≠ Clinical significance

Comparing values with predefined acceptance criteria

Sunday, 25 Oct, 2015, morning

9:00 – 9:45 Six sigma metrics
Sten Westgard

9:45 – 10:30 Performance criteria
Gunnar Nordin

10:30 – 11:15 Biological variation
Sten Westgard

11:15 Poster award & Closing

"How to assess the quality of your method?"
October 24-25, 2015, Zagreb, Croatia
Method comparison

- Important laboratory procedure for verification
- Included into validation protocols for new reagents
  - Comparison with the reference method
  - Comparison with different manufacturers
  - Comparison with same manufacturer
- Results are presented in manufacturers declarations
Can we rely on manufacturers declarations?

- **Comparing 7 insert sheets for glucose concentration measurement**

<table>
<thead>
<tr>
<th>Manufacturer</th>
<th>N</th>
<th>Unit</th>
<th>r</th>
<th>Intercept (95% CI)</th>
<th>Slope (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>102</td>
<td>mg/dL</td>
<td>0.9993</td>
<td>-4.54 (?,?)</td>
<td>1.06 (?,?)</td>
</tr>
<tr>
<td>B</td>
<td>117</td>
<td>mmol/L</td>
<td>0.998</td>
<td>-0.081 (?,?)</td>
<td></td>
</tr>
<tr>
<td>C</td>
<td>75</td>
<td>mmol/L</td>
<td>1.000</td>
<td>0.179 (?,?)</td>
<td>0.996 (?,?)</td>
</tr>
<tr>
<td>D</td>
<td>43</td>
<td>mg/dL</td>
<td>0.9977</td>
<td>-2.6 (?,?)</td>
<td>1.084 (?,?)</td>
</tr>
<tr>
<td>E</td>
<td>?</td>
<td>mg/dL</td>
<td>0.999</td>
<td>0.68 (?,?)</td>
<td>0.99 (?,?)</td>
</tr>
<tr>
<td>F</td>
<td>40</td>
<td>mg/dL</td>
<td>0.98</td>
<td>-3.14 (?,?)</td>
<td>0.98 (?,?)</td>
</tr>
<tr>
<td>G</td>
<td>60</td>
<td>mmol/L</td>
<td>0.998</td>
<td>0.000 (?,?)</td>
<td>1.008 (?,?)</td>
</tr>
</tbody>
</table>

- **Correlation for determination of agreement**

- **No 95% CI for evaluation of bias**

- **No BA analysis**
To conclude

- Linear regression analysis
- Bland-Altman plot
- Interpretation of results
- Data analysis
- Verification procedure
- Laboratory methods
- Clinically relevant criteria
- Number of samples
- Measurement range
- Time of analysis
- Data analysis
- Data interpretation
Comparability of methods and analyzers

- Coefficient of correlation doesn’t allow conclusions about comparability of methods, but only about linear association between them, even when it is very high (close to 1).

- Regression equation: $Y = 0.67 (-0.15-1.32) + 1.09 (1.03-1.22) x$ is an example of proportional bias between methods (95% CI for slope not including 1) without constant bias between methods (95% CI for intercept including 0).

- Regression equation for glucose concentration: $Y = 0.07 (0.01-0.13) + 1.15 (0.85-1.23) x \text{ (mmol/L)}$ is an example of statistically significant, but clinically non-significant constant bias. Value of 0.07 (0.01-0.13) mmol/L glucose is lower than conventional analytical performance of the test.