POCT Method Validation

Step by step
Outline

• Background
  – Method validation
  – POCT glucose measurement

• The experiments
  – Linearity (reportable range)
  – Imprecision
  – Inaccuracy (bias)
  – Interference testing

• POCT glucose testing
  – Example: NOVA StatStrip®
    (POC Vol 13, Nr 4, Dec 2014)

• Conclusion
Background: Method Validation

• A standard method validation process provides a **systematic** way for evaluating the **performance** of a testing process.

• All about **error** assessment.

• **Aim:**
  – To determine how much error is present in a test result produced by the (POCT) method.

• **Want to know:**
  – Will the error introduced by the method influence the interpretation of the test (do the size of the errors exceed what is allowable for the test method)?
Background: Method Validation

1. Define analytical requirement
2. Identify candidate method
3. Plan validation experiments
   - Carry out experiments
   - Use data to assess fitness-for-purpose

Stages before validation

- Statement of validation
- Analytical requirements met

Yes
No
• Determine the linearity REPORTABLE RANGE EXPERIMENT

• Error estimation

  – Random error (imprecision)
  • Error that can be either (+) or (-)
  • Direction + magnitude cannot be predicted
  • Calculate the SD from the result of a set of replicate measurements
  • REPLICATION EXPERIMENT

  – Systematic error (inaccuracy)
  • Always in one direction, i.e. all results are either high or low
  Constant error: stays the same over a range of [ ]’s
  Proportional error: changes as the concentration changes
  • COMPARISON OF METHODS EXPERIMENT

  – Total error
  • Combined effect of random and systematic error
Background: Method Validation

High Accuracy
High Precision

Low Accuracy
High Precision

High Accuracy
Low Precision

Low Accuracy
Low Precision
Point-of-Care testing use in general is on the increase

With the increasing incidence of diabetes mellitus and metabolic syndrome, early testing has become more common

POCT glucose meter use widely used for screening and self-monitoring screening

Also used in hospitals and intensive care settings due to rapid TAT

Here patient factors such as hemodynamics and drug treatment becomes important, as POCT glucose meters are prone to interference

Concern: precision as well as accuracy in different clinical settings
In 2014, the US Food and Drug Administration (FDA) approved the NOVA StatStrip® Glucose Hospital Meter System for use in the critical care setting.

Evaluated by Rensburg et al.

2 aims:
- To assess the analytical performance of the StatStrip (Xpress and Connectivity) to 2 Accu-check Active meters (Roche Diagnostics), using the plasma glucose by the glucose oxidase method on the Siemens Advia 1800 as the reference method.
- To assess the glucose meters in the clinical setting.
• CLSI C30-A2: Point-of-Care Blood Glucose Testing in Acute and Chronic Care Facilities

  – “When rapid results are required...to make therapeutic decisions, and the time required to obtain results from the clinical laboratory would compromise patient care, POC blood glucose testing is appropriate”

  – Supplements rather than replaces testing in the clinical laboratory
Background: POCT Glucose Measurement

- Performance criteria
  - Precision
  - Accuracy
  - Linearity within assay range

- EP9 – Method Comparison and Bias Estimation Using Patient Samples

- EP10 – Preliminary Evaluation of Quantitative Clinical Laboratory Methods

- EP15 – User Demonstration of Performance for Precision and Accuracy
Background: POCT Glucose Measurement

- Patient specimens, commercially available material
- Determine if POC system is whole blood or plasma calibrated (NB when comparing POC results with laboratory result)
- Effect of hematocrit: most glucose monitoring systems provide accurate measurement only within a defined hematocrit range
- Potential interferences with glucose measurement, e.g. drugs (mannitol, acetaminophen, ascorbic acid, dopamine) – refer to manufacturers information
Method validation ➔ verify manufacturers claims regarding the analytical range, precision, accuracy (± interferences)

POC glucose measurement ➔ additional accuracy requirements from American Diabetes Association (ADA) and International Standardization Organization (ISO)
  – ADA: Analytical error <5% across all levels
  – ISO:

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Target blood glucose level from which to base mmol/L (mg/dL) bias or % bias</td>
<td>4,2 mmol/L (75 mg/dL)</td>
<td>5,6 mmol/L (100 mg/dL)</td>
</tr>
<tr>
<td>Acceptable bias from reference value for lower target glucose levels</td>
<td>± 0,83 mmol/L (± 15 mg/dL)</td>
<td>± 0,83 mmol/L (± 15 mg/dL)</td>
</tr>
<tr>
<td>Acceptable bias from reference value for higher target glucose levels</td>
<td>± 20%</td>
<td>± 15%</td>
</tr>
<tr>
<td>Acceptable % of all results within bias limits</td>
<td>95%</td>
<td>95%</td>
</tr>
<tr>
<td>Parkes Error Grid</td>
<td>Not required</td>
<td>99% of results within Zones A and B</td>
</tr>
</tbody>
</table>
POCT Method Validation
Get to know your method (s)

- Background information on method to be evaluated

<table>
<thead>
<tr>
<th>Information</th>
<th>StatStrip (Xpress and Connectivity)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Waived test</td>
<td>Can perform outside the laboratory setting, no method validation required</td>
</tr>
<tr>
<td>Intended use</td>
<td>Quantitative glucose measurement in whole blood (can be used in ICU); not for diagnosis or screening of diabetes</td>
</tr>
<tr>
<td>Methodology</td>
<td>Modified glucose oxidase-based amperometric technology with hematocrit correction</td>
</tr>
<tr>
<td>Units of measurement</td>
<td>mmol/L (can set to mg/dL)</td>
</tr>
<tr>
<td>Calibration</td>
<td>Plasma calibrated</td>
</tr>
<tr>
<td>Manufacturer claims</td>
<td>0.6 – 33.3 mmol/L (10 mg/dL – 600 mg/dL)</td>
</tr>
<tr>
<td>- - Measurement range</td>
<td>3.6% @ 2.2 mmol/L; 2.8% @ 7.2 mmol/L; 2.4% @ 17.5 mmol/L; 2.9% @ 31.7 mmol/L</td>
</tr>
<tr>
<td>- - Precision</td>
<td>5% @ 1.6 mmol/L; 2.3% @ 4.1 mmol/L; 3.3% @ 11.2 mmol/L; 1.7% @ 33.0 mmol/L</td>
</tr>
<tr>
<td>- - Accuracy</td>
<td>Capillary: Slope = 1.004; y-intercept = -4.16; r² = -0.9601</td>
</tr>
<tr>
<td></td>
<td>Venous: Slope = 0.991; y-intercept = -2.22; r² = 0.9858</td>
</tr>
<tr>
<td>Interferences</td>
<td>No interference from acetaminophen, ascorbic acid, dopamine, ephedra, D-galactose, ibuprofen, L-Dopa, Methyl-Dopa, Salicylate, Tetracycline, Tolazamide, Tolbutamide</td>
</tr>
<tr>
<td></td>
<td>No interference with hematocrits from 20% to 65%</td>
</tr>
</tbody>
</table>

- Reference assay: Glucose oxidase; Siemens Advia 1800 (laboratory)
- Other glucometers: Accu-check Active (Roche); glucose dehydrogenase pyrroloquinolinquinone electrochemical method
- Training NB
POCT Method Validation

Reportable Range Experiment (Linearity)
Reportable range experiment

- NB to determine linearity/reportable range
- Manufacturers make claims that should be checked
  - StatStrip: 0.6 – 33.3 mmol/L (10 mg/dL – 600 mg/dL)
- Series samples of known concentrations/series of known dilutions of highly elevated specimen or patient pool
  - StatStrip: 5 levels “linearity solution” 1.6/4.1/11.2/16.7/33.0 mmol/L
- CLSI:
  - Need 4 different levels, preferably 5
  - 4 measurement per level (Westgard 3 replicates); calculate mean
  - Plot data
    - Westgard Linear Data Plotter; Analyse-it® Method Validation software
    - mean of measurement on y-axis and assigned value on x-axis
- Determine linear range with visual inspection
### Reportable range experiment

#### Westgard QC

**POCT Method Validation**

<table>
<thead>
<tr>
<th>Analyst</th>
<th>M Hoffmann</th>
</tr>
</thead>
<tbody>
<tr>
<td>Test</td>
<td>POCT glucose</td>
</tr>
<tr>
<td>Method</td>
<td></td>
</tr>
<tr>
<td>X Label</td>
<td></td>
</tr>
<tr>
<td>Y Label</td>
<td></td>
</tr>
</tbody>
</table>

**Linear-data Plotter**

- Change the number of pools: 5
- Change the number of samples: 3

Enter values and click the 'Calculate' button below.

<table>
<thead>
<tr>
<th>Value</th>
<th>Assigned Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td></td>
</tr>
</tbody>
</table>

**Pool 1**

- Average: -

**Pool 2**

- Average: -

**Pool 3**

- Average: -

**Pool 4**

- Average: -

**Pool 5**

- Average: -

**Reportable range data not published in article by Rensburg et al**
Replication Experiment (Random Error/Imprecision)
A replication experiment is performed to estimate the imprecision or random error of the analytical method. Typically performed by obtaining test results on 20 samples of the same material and then calculating the mean ($\bar{x}$), standard deviation (SD), and coefficient of variation (CV)

$$s = \sqrt{\frac{\sum(x_i - \bar{x})^2}{(n - 1)}}$$

$$CV = \frac{SD}{\bar{x}} \times 100$$

Close to medical decision limits

Method validation usually requires “within-day” and between-” day runs to calculate “total imprecision”

Can use standard solutions, control solutions, or pools of fresh patient samples (not possible for glucose testing)

2 – 3 control levels (near medical decision limits)

2 alternatives – Standard protocol EP15 protocol
Replication Experiment

Standard protocol

- Analyze 20 samples of each material within a run or within a day to obtain short-term imprecision (calculate short term imprecision)

\[
\bar{x} = \frac{\sum x_i}{n}
\]

\[
s = \sqrt{\frac{\sum (x_i - \bar{x})^2}{n - 1}}
\]

\[
CV = \frac{SD}{\bar{x}} \times 100
\]

- Determine if short term imprecision is acceptable BEFORE proceeding (reject method if not acceptable = “best case scenario”)

- Analyze 1 sample of each of the 2 materials on 20 different days to estimate long-term imprecision. Determine whether long-term imprecision is acceptable

- Acceptability: CLIA or Manufacturers Claims
Replication Experiment

CLIA
- within-day SD should be $\frac{1}{4}$ or less of the defined allowable total error
- Total standard deviation should be $\frac{1}{3}$ or less of the defined total error allowable

Manufacturers claims:
- Obtain SD and nr of measurements used in the experiment from manufacturers claims
- Obtain SD and nr of replications from your replication experiment
- Calculate the F-value (larger SD squared divided by smaller SD squared)
- Look up critical F-value for 20 degrees of freedom ($dF = n-1$) in the numerator (your nr of replicates) and “X” degrees of freedom as denominator
- If calculated-F is less than critical-F, the manufacturers claim has been verified

POCT Method Validation
Replication Experiment

- F-test tells whether the difference in the variances is statistically significant
- SD of test method and comparative methods are squared, and the larger variance is divided by the smaller variance
- \( F = \frac{(s_1)^2}{(s_2)^2} \)
- Calculated F-value compared to critical F-value, and if calculated F-value less than critical, it means that there is no difference between the random errors of the two methods

F Table: Critical values of F for \( p=0.05 \) (probability)
Replication Experiment

Rensburg et al

- CP 30-A2 does not specify how precision should be tested StatStrip (2) and Accu-Check (2)

- Precision Studies:
  - Calculated within-run imprecision only
  - Quality control material: 3 levels run repeatedly for 20 replicates
    - $\bar{x} = 3.5$ mmol/L (range = 3.3 – 3.7; SD = 0.1; CV = 2.8%)
    - $\bar{x} = 6.2$ mmol/L (range = 5.7 – 6.6; SD = 0.23; CV = 3.7%)
    - $\bar{x} = 16.6$ mmol/L (range = 16.2 – 17.0; SD = 0.2; CV = 1.2%)
  - Spiked heparinized whole blood: 3 levels run repeatedly for 20 replicates
    - $\bar{x} = 5.2$ mmol/L (5.0 – 5.4)
    - $\bar{x} = 15.7$ mmol/L (14.9 – 16.5)
    - $\bar{x} = 21.5$ mmol/L (20.2 – 21.9)

- All glucose meters achieved precision of less than 5% at all levels (NB: hypoglycaemic range not covered)
Replication Experiment

Preparing whole blood specimens (manufacturers protocol)
Replication Experiment

- StatStrip (Rensburg et al): $\bar{x} = 3.5$ mmol/L; $n=20$; $SD = 0.1$
- Manufacturer claim: $\bar{x} = 4.1$ mmol/L; $SD = 0.09$, CV 2.3%
- $F = (s_1)^2/(s_2)^2 = (0.1)^2/(0.09)^2 = 1.23$ (Calculated F-value)
- $F = \text{table} = 1.748$ (Critical F-value)
- Calculated $F < \text{Critical } F$: NO DIFFERENCE in random error/variance
- Claim verified
Replication Experiment

- EP 15 – Precision protocol intended for the laboratory
- Shorter and simpler
- 2 Control levels; 3 replicates; period of 5 days; total of 15 measurements
- Estimate total imprecision
- Mathematics complicated (use a template)
# Replication Experiment

## POCT Method Validation

**Analyte:** paracetamol  
**Units:** umol/l  
**Method:** Enzymatic

### LEVEL 1

<table>
<thead>
<tr>
<th>Run</th>
<th>Date</th>
<th>Replicate 1</th>
<th>Replicate 2</th>
<th>Replicate 3</th>
<th>Run Mean</th>
<th>(x1 - Mean)^2</th>
<th>(x2 - Mean)^2</th>
<th>(x3 - Mean)^2</th>
<th>Sum (Run Mean - Mean)^2</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td></td>
<td>91,500</td>
<td>91,800</td>
<td>92,200</td>
<td>91,833</td>
<td>0,11111</td>
<td>0,00111</td>
<td>0,13444</td>
<td>0,24667</td>
</tr>
<tr>
<td>2</td>
<td></td>
<td>90,600</td>
<td>91,800</td>
<td>92,900</td>
<td>91,767</td>
<td>0,00111</td>
<td>0,44444</td>
<td>0,40111</td>
<td>0,84667</td>
</tr>
<tr>
<td>3</td>
<td></td>
<td>91,100</td>
<td>90,400</td>
<td>91,700</td>
<td>91,067</td>
<td>0,02778</td>
<td>0,01778</td>
<td>0,00111</td>
<td>0,04667</td>
</tr>
<tr>
<td>4</td>
<td></td>
<td>90,300</td>
<td>90,600</td>
<td>90,500</td>
<td>90,400</td>
<td>0,00000</td>
<td>0,04000</td>
<td>0,04000</td>
<td>0,08000</td>
</tr>
<tr>
<td>5</td>
<td></td>
<td>90,400</td>
<td>90,600</td>
<td>90,200</td>
<td>90,400</td>
<td>0,00000</td>
<td>0,04000</td>
<td>0,04000</td>
<td>0,08000</td>
</tr>
</tbody>
</table>

**Mean:** 91,107 umol/l  
**Number of days:** 5  
**Replicates per day:** 3  
**Within-Run Imprecision:** 0,622  
**Total Imprecision:** 0,852  
**Variance for daily means (B):** 0,4685556  
**Mean:** 981,30 umol/l  
**Number of days:** 5  
**Replicates per day:** 3  
**Within-Run Imprecision:** 2,9439  
**Total Imprecision:** 10,74  
**Variance for daily means (B):** 104,74111

### LEVEL 3

<table>
<thead>
<tr>
<th>Run</th>
<th>Date</th>
<th>Replicate 1</th>
<th>Replicate 2</th>
<th>Replicate 3</th>
<th>Run Mean</th>
<th>(x1 - Mean)^2</th>
<th>(x2 - Mean)^2</th>
<th>(x3 - Mean)^2</th>
<th>Sum (Run Mean - Mean)^2</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td></td>
<td>984,50</td>
<td>984,50</td>
<td>995,40</td>
<td>988,133</td>
<td>13,20111</td>
<td>13,20111</td>
<td>52,80444</td>
<td>79,20667</td>
</tr>
<tr>
<td>2</td>
<td></td>
<td>974,00</td>
<td>968,50</td>
<td>969,50</td>
<td>970,667</td>
<td>11,11111</td>
<td>4,69444</td>
<td>1,36111</td>
<td>17,16667</td>
</tr>
<tr>
<td>3</td>
<td></td>
<td>987,50</td>
<td>995,40</td>
<td>989,80</td>
<td>990,900</td>
<td>11,56000</td>
<td>20,25000</td>
<td>1,21000</td>
<td>33,02000</td>
</tr>
<tr>
<td>4</td>
<td></td>
<td>989,70</td>
<td>988,20</td>
<td>985,30</td>
<td>987,067</td>
<td>6,93444</td>
<td>0,75111</td>
<td>3,12111</td>
<td>10,80667</td>
</tr>
<tr>
<td>5</td>
<td></td>
<td>973,30</td>
<td>968,20</td>
<td>967,70</td>
<td>969,733</td>
<td>12,72111</td>
<td>2,35111</td>
<td>4,13444</td>
<td>19,20667</td>
</tr>
</tbody>
</table>

**Mean:** 981,30 umol/l  
**Number of days:** 5  
**Replicates per day:** 3  
**Within-Run Imprecision:** 3,1887333  
**Total Imprecision:** 3,4654944  
**Variance for daily means (B):** 104,74111  
**Degrees of freedom:** 10  
**Percent Point (C):** 20,483177  
**Within-Run Verification value:** 14,444444  
**Total precision degrees of freedom (T):** 4,8329623  

---

**Verification Limit**

**Claimed Within-Run Imprecision:** 0,54664  
**Claimed Total Imprecision:** 3,1887333  
**Within claim**

---

**Within-Run Imprecision:** 0,622  
**Total Imprecision:** 10,74  
**Verification Limit**

**Levels:** 3  
**Claimed Within-Run Imprecision:** 0,54664  
**Claimed Total Imprecision:** 3,1887333  
**Within claim**

---

**Verification Limit**

**Degrees of freedom:** 10  
**Percent Point (C):** 20,483177  
**Within-Run Verification value:** 4,2132935  
**Total precision degrees of freedom (T):** 4,8329623  

---

**Variance for daily means (B):** 0,4685556  
**Degrees of freedom:** 10  
**Percent Point (C):** 20,483177  
**Within-Run Verification value:** 0,7823482  
**Total precision degrees of freedom (T):** 8,5738892
POCT Method Validation

Comparison of Methods (Systematic Error/bias)
Comparison of methods

• Performed to estimate inaccuracy or systematic error

• Experiment performed by analyzing patient samples on the new method (test method) and comparative method

• Estimate the errors on basis of the differences observed between the methods

• Error at medical decision levels NB

• Important (general):
  – Need a minimum of 40 different patient specimens that cover the reportable range
  – Need to be performed in different analytical runs on different days (minimum 5 days)
  – Specimens should be analyzed within 2 hrs on each of the test and comparative method
Comparison of methods

- Graph the comparison data and visually inspect the data.
- Discrepant results between test- and comparative methods should be re-analyzed to confirm that difference is real (remove outliers).
- Difference plot (Bland-Altman plot).
Comparison of methods

- Graph the comparison data and visually inspect the data
- Discrepant results between test- and comparative methods should be re-analyzed to confirm that difference is real (remove outliers)
- “Comparison Plot”: 

![Comparison Plot](image)
Comparison of methods

• Numerical estimates of the errors between test and comparative method can be obtained from statistical calculations.

• Statistics should provide:
  – Error at medical decision limits
  – Know constant and proportional nature of error

• Regression statistics: Ideal to have three statistical parameters that can each estimated a different type of error
  – Proportional error can be estimated from the slope
  – Constant error by the Y-intercept
  – Systematic error can be estimated at ANY concentration using the regression equation ($Y_c = bX_c + c$)
Comparison of methods

- Regression line: \( y = 1,171x + 0,1308 \); \( R^2 = 0,92 \)
- Slope (proportional error) = 1,17 (17%)
- Y-intercept (constant error) = 0,13 mmol/L
- \( R^2 \): how close the data are to the fitted regression line (goodness of fit)

Systematic error at medical decision level of 3 mmol/L:

\[
Y = 1,1711x + 0,1308 \\
Y = 1,1711 (3 \text{ mmol/L}) + 0,1308 = 3,64 \text{ mmol/L}
\]

Difference \( Y \) and \( X \) = 0,64 mmol/L = systematic error at a level of 3 mmol/L = 21%
Comparison of methods

CLSI C30-A2: POC Blood Glucose Testing

- Using altered blood samples
  - Allow anticoagulated blood to undergo glycolysis
  - Prepare \textit{40} specimens that span the measurement range
  - Target of \(1.4 + 2.8 + 4.4 + 8.3 + 13.9 + 22.2 + 33.3\) mmol/L
  - Spike samples only to obtain sufficient samples for glucose \(<2.8\text{mmol/L}\) and \(>22.2\text{mmol/L}\)
  - Measure samples in \textit{duplicate} on test and comparative method
  - Centrifuge and separate sample within \textit{5 min} of analysis on POC
  - Separated plasma should be tested within \textit{60 min}
Comparison of methods


• Comparing results
  – Duplicate results should match within 4% or 0.22 mmol/L
  – 95% of individual results test method agree
    • within ± 0.83 mmol/L of comparative method below 4.2 mmol/L (2003) and 5.6 mmol/L (2013)
    • within ± 20% of the comparative method above 4.2 mmol/L (2003)/5.6 mmol/L (2013)
  – 99% of results within Zones A and B of Parkes Error Grid
Comparison of methods

Parks error grid

- Developed to analyze the clinical significance of the bias between blood glucose system results and the laboratory reference results

- Data points assigned to one of the 5 zones (A – E) on the error grid

- Zones A and B clinically acceptable, bias from lab results would not lead to treatment decisions that may put a patient at risk

- As the bias increases (Zones C, D and E), there is greater risk of under- or overtreating a patient based on the glucose system result
Comparison of methods

AACB Guidelines for the Evaluation of PoCT instruments

- **Linearity**
  - Minimum of 2 replicates at 5–7 concentrations over range
- **Precision testing:**
  - Include both quality control material and patient samples
  - Both within- and between run imprecision should be determined
  - Minimum of 20 replicates
  - CLSI document EP-15A2 can be used

- **Method comparison**
  - 40 samples covering clinically meaningful range
  - Time between analysis on 2 methods should not exceed 2 hrs
    - Bland-Altman plot illustrating bias and allowable limits
    - Regression analysis
    - Error Grid (glucose)
Comparison of methods

StatStrip Rensburg et al

- Used Analyse-it® Software

- Laboratory based method comparison
  - Heparinized venous whole blood obtained from volunteers
  - Analysed within 60 min of collection
  - Total of 155 blood samples analyzed of the 4 glucose meters, then immediately centrifuged for measurement on the laboratory instrument Siemens Advia 1800
  - 20 of these samples spiked with various volumes of glucose concentrate to extend the range

- Regression Statistics
- Difference Plots (Bland-Altman plots)
Comparison of methods

Regression Statistics

<table>
<thead>
<tr>
<th>Glucose Meter</th>
<th>Slope</th>
<th>Intercept, mmol/L</th>
<th>R²</th>
<th>Median Bias, mmol/L</th>
</tr>
</thead>
<tbody>
<tr>
<td>StatStrip Xpress</td>
<td>1.013</td>
<td>0.005</td>
<td>0.997</td>
<td>0.1</td>
</tr>
<tr>
<td>StatStrip Connectivity</td>
<td>1.014</td>
<td>0.055</td>
<td>0.997</td>
<td>0.2</td>
</tr>
<tr>
<td>Accu-Chek 1</td>
<td>0.966</td>
<td>0.179</td>
<td>0.925</td>
<td>0.5</td>
</tr>
<tr>
<td>Accu-Chek 2</td>
<td>0.978</td>
<td>0.73</td>
<td>0.934</td>
<td>0.5</td>
</tr>
</tbody>
</table>

*Data of only 1 lot number displayed as similar result obtained with other lot number strips.

**Slope (proportional error)**
- Slope close to 1 = nearly no proportional error

**Intercept**
- Intercept close to "0" = nearly no constant error

**R²**
- R² > 0.99 = excellent correlation; regression line excellent for predicting Y
Comparison of methods

Bland-Altman plot (difference plot/bias plot)

ISO 15197 (2003) used to confirm accuracy:
- within 0.83 mmol/L at levels < 4.2 mmol/L
- ± 20% at levels above 4.2 mmol/L

POCT Method Validation

Test (POCT) – comparative method (Lab)

Difference

Bias (Xpress mmol/L minus Advia-Oxidase mmol/L)

Advia - Oxidase (mmol/L)

Comparative method (Laboratory)
Comparison of methods

A

B

C

D

POCT Method Validation
Comparison of methods

100% of measurements within allowable error (ISO 15197)
No bias trend

Only 80% of samples up to 4.2 mmol/L met ISO standard
Only 91% of samples >4.2 mmol/L met ISO 15197 standard
Positive bias trend
StatStrip Rensburg et al

- Used Analyse-it® Software

- **Clinical Method Validation**
  - Clinical Setting: Diabetic clinic
  - 110 patients
  - Finger-prick capillary blood sample measured on StatStrip Express and Connectivity meters
  - Simultaneously venous whole blood collected in Na-fluoride tube

- Regression Statistics
- Difference Plots (Bland-Altman plots)
Comparison of methods

Regression Statistics

<table>
<thead>
<tr>
<th>Glucose Meter</th>
<th>Slope</th>
<th>Intercept, mmol/L</th>
<th>$R^2$</th>
<th>Median Bias, mmol/L</th>
</tr>
</thead>
<tbody>
<tr>
<td>StatStrip Xpress</td>
<td>1.013</td>
<td>0.005</td>
<td>0.997</td>
<td>0.1</td>
</tr>
<tr>
<td>StatStrip Connectiv</td>
<td>1.014</td>
<td>0.055</td>
<td>0.997</td>
<td>0.2</td>
</tr>
</tbody>
</table>

Bland-Altman plot (difference/bias plot)  
With ISO 15197 (2003) standard limits
Interference testing

POCT Method Validation
Interference testing

• Interference testing performed to estimate the effect of specific materials on the accuracy of a method

• Pairs of test samples prepared
  – First sample: interfering material added to sample
  – Second sample: diluent added to sample
  – Both analyzed with the test method to see if there is any difference

• Use patient samples and analyse in each sample in duplicate

• Concentration of interfering material should be near maximum concentration expected in patient population

• Substances to be tested identified from performance claims, literature reports etc.
Interference testing

- Nr of paired samples will be much less than the 40 used for method comparisons

- Regression statistics not appropriate (data not likely to demonstrate wide analytical range)

- Compare observed error to allowable error (e.g. CLIA); if it exceeds allowable error = not acceptable

- Common interferences:
  - Bilirubin (addition of standard bilirubin solution)
  - Haemolysis (mechanically haemolyzing part of one of pair by freezing and thawing and adding back)
  - Lipaemia (commercial fat emulsion)

- Glucose specific:
  - Acetaminophen
  - Ascorbic Acid
  - Maltose
  - Xylose

Evaluated by Rensburg et al.
Manufacturers claim: no interference from these substances
Interference testing

• CLSI C30-A2: POC Blood Glucose Testing
  – Does not specify that interference testing should be performed
  – “refer to manufacturer’s information regarding potential interferences for a specific meter”

• AACB Guide for the Evaluation of PoCT instruments
  – Analyse pairs in duplicate
  – Good practice to test interference of bilirubin, haemolysis and lipaemia
  – Exogenous analytes/drugs
Interference testing

StatStrip Rensburg et al

- 150ml heparinized whole blood collected from single-donor volunteer with normal Hct
- Room temp on roller overnight (glycolysis)
- Spiked with various concentrations of glucose
  - 5 different glucose concentrations
  - Range 3.1 mmol/L – 23.7 mmol/L
- Each of 5 glucose levels spiked with:
  - Ascorbic acid
    (final concentration 0.29 mmol/L + 0.59 mmol/L)
  - Acetaminophen
  - Maltose
    (final concentration 2.8 and 5.6 mmol/L)
  - Xylose
    (final concentration 5.6 and 11.1 mmol/L)
Interference testing

Example of ascorbic acid at 0.29 mmol/L and 0.9 mmol/L
at a glucose level of ± 3 mmol/L

Roche Active meters affected by presence of ascorbic acid in low and normal blood glucose samples

Accuracy for the low and reference range glucose samples from 10% to 81%

StatStrip accuracy errors averaged ± 5% - 8%
Interference testing

Example of Maltose (2.8 mmol/L and 5.8 mmol/L) at a glucose of 1.9 mmol/L and Xylose 5.6 mmol/L and 11.1 mmol/L at a glucose of ± 6.6 mmol/L

Roche severely affected by Maltose and Xylose

Accuracy errors: 12% (mid-range glucose) and 42% (high range glucose)

Errors of 77% to 65% across reference- and low range glucose samples

StatStrip not affected (accuracy error < 2%)
Hematocrit effect
Hematocrit effect

- Various Hct levels prepared from the single donor sample using fresh heparinized whole blood
- Allowed to sit at room temp for ± 24 hrs
- Divided into 3 aliquots
- Spiked to obtain 3 different glucose levels
- Further divided into 5 aliquots of 1ml
- Centrifuged
- Plasma adjusted (taking plasma from 1 tube to another) to obtain 5 samples each of a different Hct (%) for each of the 3 glucose levels
- Rocked for 10 min
- Analysed on glucose meters
- Centrifuged and plasma glucose determined on Siemens Advia 1800
- Hct values measured using mini Hct centrifuge
Hematocrit effect

- Nova StatStrip had average accuracy errors of < 5% when compared to reference method.

- Roche Active had accuracy errors of up to 17% at low Hct levels and 21% at high Hct levels across low and high glucose levels.

- Roche meters results at low- and high-Hct levels differed from itself at normal levels by up to 22% on the same glucose samples.
Conclusion

- POCT meter use on the increase
- Ensures rapid turnaround times and timely treatment in various clinical settings
- Method evaluation, validation and verification provides objective evidence that a method is fit for purpose, meaning the particular requirements for a specific intended use are fulfilled