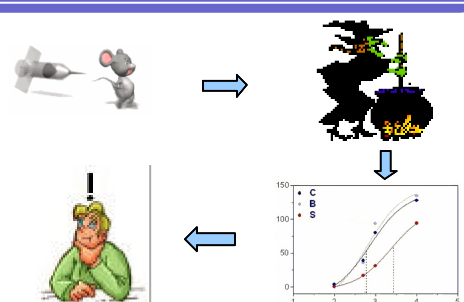
To be precise… and accurate!

Yossi Levy
ISA Annual Meeting
May 9, 2006

Agenda

- What is a biological method?
- Accuracy and precision
- "The way they do it at Chemistry"
- Estimation of accuracy and precision
  - Modeling
  - Experimental design
  - Reporting

Biological method on a shoestring



Definitions – ICH Guideline Q2A

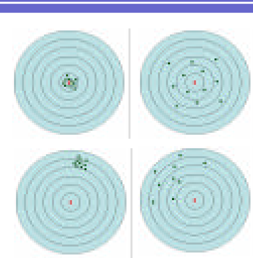
The accuracy … expresses the closeness of agreement between … an accepted reference value and the value found.

The precision … expresses the closeness of agreement (degree of scatter) between a series of measurements

Definitions – ICH Guideline Q2A

- *Repeatability expresses the precision under the same operating conditions over a short interval of time.*
- *Intermediate precision expresses within-laboratories variations: different days, different analysts, different equipment, etc.*
- *Reproducibility expresses the precision between laboratories.*

A look at accuracy and precision



Accuracy=Bias *Precision=Variance*

"The way they do it at Chemistry "

1. Measure accuracy and repeatability using 6 runs by the same analyst on the same day – report CV.
2. Measure reproducibility using another 6 runs by another analyst on another day – report "Reproducibility Difference "

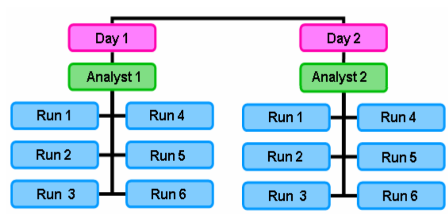
"The way they do it" advantage

- No experimental design
- No modeling
- No complex calculations
- Simple reporting

"The way they do it" problems

- Biological methods are more complicated to implement, therefore the numbers of possible runs in a single day is limited.
- Variation of biological methods is generally higher compared to chemical methods.
- Measuring intermediate precision is not enabled.
- No statistical sense.

"The way they do it" Experimental Design



Example - Biological data

Day/ Analyst	Run 1	Run2	Run 3	Mean	STD
1	0.768	0.601	0.887	0.752	0.144
2	0.460	0.398	0.519	0.459	0.061

$$\text{Accuracy} = 100 \cdot \frac{0.752}{0.7} = 107.4\%$$

$$\text{Repeatability} = 100 \cdot \frac{0.144}{0.752} = 19.1\%$$

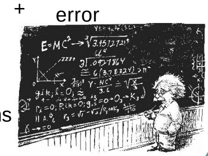
$$\text{ReproducibilityDifference} = 100 \cdot \frac{|0.752 - 0.459|}{\frac{0.752 + 0.459}{2}} = 48.4\%$$

The way we would do it, at Statistics

$$Y_{ij} = m + b_i + c_j + e_{ij}$$

Signal = fixed parameter + random effects + random error

Assumptions:
- •Independence
- •Normal distribution
- •Zero mean deviations
- •STDs: s_b, s_c, s



SAS syntax

```
data example1;
  input day y @@;
cards;
1 0.768 1 0.601 1 0.887
2 0.460 2 0.398 2 0.519
;
run;
proc mixed method=reml covtest cl;
  class day;
  model y= / solution cl;
  random day;
run;
```

Example - SAS output

Covariance Parameter Estimates							
Cov Parm	Estimate	Standard Error	Z Value	Pr Z	Alpha	Lower	Upper
day	0.03887	0.06077	0.64	0.2612	0.05	0.007062	175.52
Residual	0.01215	0.008592	1.41	0.0786	0.05	0.004362	0.1003

Repeatability

Between Day precision

Solution for Fixed Effects								
Effect	Estimate	Standard Error	DF	t Value	Pr > |t|	Alpha	Lower	Upper
Intercept	0.6055	0.1465	1	4.13	0.1511	0.05	-1.2560	2.4670

Accuracy

Results that make biological sense

$$\text{Accuracy} = 100 \cdot \frac{0.6055}{0.7} = 86.4\%$$

$$\text{Repeatability} = 100 \cdot \frac{\sqrt{0.01215}}{0.6055} = 18.2\%$$

$$\text{Reproducibility} = 100 \cdot \frac{\sqrt{0.03887 + 0.01215}}{0.6055} = 37.3\%$$

Results that make statistical sense

	Parameter	Estimate	95% confidence interval
Accuracy	m	0.6055	-1.2560 - 2.4670
Repeatability	s	0.0122	0.004362 - 0.1003
Between Days precision	$\sqrt{s_D^2 + s^2}$	0.2259	????

Satterthwaite Confidence Intervals

$$\frac{n \hat{s}^2}{c^2_{n,1-a/2}} \leq s^2 \leq \frac{n \hat{s}^2}{c^2_{n,a/2}}$$

where $n = 2Z^2$

and Z is the Wald statistic: $Z = \frac{\hat{s}^2}{SE(\hat{s}^2)}$

CI for variance components sum

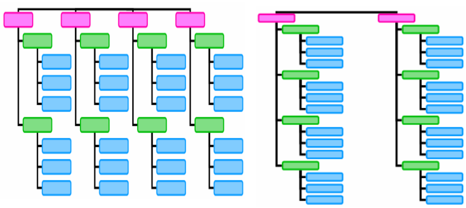
since the variance components are assumed to be independent, a confidence interval for $s^2 + s_b^2$ would be

$$\frac{m \cdot (\hat{s}^2 + \hat{s}_b^2)}{c^2_{n+n_b,1-a/2}} \leq s^2 + s_b^2 \leq \frac{m \cdot (\hat{s}^2 + \hat{s}_b^2)}{c^2_{n+n_b,a/2}}$$

where $n_b = 2Z_b^2$ $\quad Z_b = \frac{\hat{s}_b^2}{SE(\hat{s}_b^2)}$

and $m = \frac{n^2 + n_b^2}{\frac{s^2}{n^2} + \frac{s_b^2}{n_b^2}}$

DOE to measure intermediate precisions



4 Days, 2 Analysts or 2 Days, 4 Analysts

Reporting intermediate precisions

$$\text{Accuracy} = 100 \cdot \frac{m}{m_0}$$

$$\text{Between Day Precision} = 100 \cdot \frac{\sqrt{s_{\text{Day}}^2 + s^2}}{m}$$

$$\text{Between Analyst Precision} = 100 \cdot \frac{\sqrt{s_{\text{Analyst}}^2 + s^2}}{m}$$

$$\text{Repeatability} = 100 \cdot \frac{s}{m}$$

$$\text{CI for } \mu,\ s,\ \sqrt{s_{\text{Day}}^2 + s^2},\ \sqrt{s_{\text{Analyst}}^2 + s^2}$$

Example 2

Covariance Parameter Estimates

Cov Parm	Estimate	Standard Error	Z Value	Pr Z	Alpha	Lower	Upper
Analyst	0.000548	0.000796	0.69	0.2455	0.05	0.000106	0.8027
Day	0.002582	0.002132	1.21	0.1130	0.05	0.000821	0.03765
Residual	0.000177	0.000057	3.08	0.0010	0.05	0.000102	0.000377

Repeatability — Between Day precision — Between Analyst precision

Solution for Fixed Effects

Effect	Estimate	Standard Error	DF	t Value	Pr > |t|	Alpha	Lower	Upper
Intercept	0.5621	0.03045	1	18.46	0.0344	0.05	0.1753	0.9490

Accuracy

Example 2 - results

	Parameter	Estimate	95% confidence interval	
Accuracy	m	0.5621	0.1753 - 0.9490	80.3%
Between Analyst precision	$\sqrt{s_{\text{Analyst}}^2 + s^2}$	0.02692	0.02059 - 0.03890	4.8%
Between Days precision	$\sqrt{s_{\text{Day}}^2 + s^2}$	0.05252	0.04061 - 0.07439	9.3%
Repeatability	s	0.01329	0.01011 - 0.01942	2.4%

Acknowledgments

- Prof. Paul Feigin - The Technion
- Dr. David Lansky - Lansky Consulting