Validation and Compliance Issues

Helmut Schütz

Bioanalytical Method Validation (BMV)

The method should be *fit for the intended use* – no need to have a 'perfect' method

- **Reliable and reproducible according to the goalposts set in the BMV guidelines.**

- **Intended use in BE:**
  - **LLOQ**   **Possible to detect carry-over ($\leq$5% C_{max} in any subject).**
    $AUC_{0-t}/AUC_{0-\infty} \geq$80%.
  - **ULOQ**   **Covering the expected C_{max} in *any* subject.**
  - **A & P**   **Chromatography**   **LLOQ**   **20%**
                                      >LLOQ   **15%**
               LBAs   **LLOQ**   **30%**
                                         >LLOQ   **20%**
  - **Stability**   **Covering start of clinical phase to end of bioanalytics.**
- **Relevant guidelines: EMA (2011), FDA (2018), ICH (draft 2019)**

Bioanalytical Method Validation (BMV)

**No guideline comes that close to a 'cookbock'
than the ones about BMV**

- **Follow them literally and you are fine…**

- **However, there are some slight differences which have to be taken into account if submitting studies to different regions.**

- **Best approach:**

    – **Close communication with the clinical team already before method development (concentration range, parent and/or metabolites, co-medications, matrix, anticoagulant, duration of study, storage, sample shipment, chiral or achiral method).**

    – **After the method is developed, assess what is required by the EMA's GL (most detailed). Only if required:**

        – **Check whether there are differences in the FDA's.**

        – **Check the current state of affairs in the ICH's.**

Bioanalytical Method Validation (BMV)

Parts

- **Method development**
    - Although not covered in GLs, good documentation recommended.
- **Method validation**
    - **Full validation**
        - Selectivity
        - Carry-over
        - Sensitivity
        - Calibration curve
        - Accuracy
        - Precision
        - Dilution accuracy
        - Stability
        - Matrix effect

Bioanalytical Method Validation (BMV)

Parts

- **Method validation (cont'd)**
  - Partial validation
  - Cross validation
- **Analysis of study samples**
  - Analytical run, acceptance criteria
  - Calibration range
  - Reanalysis of samples
  - Integration
  - Incurred sample reassessment (reanalysis)
- **Validation report**
- **Analytical report**

Fleming.

BMV: Similarities, differences

Topic	EMA	FDA	ICH
Selectivity	Must be able to differentiate analyte and IS from endogenous compounds and other components (metabolites, co-administered drugs).		
	Six individual sources of matrix.		Six individual sources of matrix + one lipaemic + one haemolyzed.
	Response <20% of LLOQ for the analyte(s) and <5% for IS.		
Carry-over	Addressed and minimized in method development.		
	Analyze blank sample after a high calibrator.		
	Response <20% of LLOQ for the analyte(s) and <5% for IS. If carry-over unavoidable, inject blank between samples.		
Sensiti-vity (LLOQ)	Lowest concentration which can be quantified reliably (with acceptable A & P). Lowest nonzero standard of the calibration curve.		
	For BE ≤5% of the anticipated C_{max}.		
	≥ Five replicates in ≥ three runs.		
	Response ≥ five times the response of the zero calibrator. Accuracy ≤20%, Precision ≤20%.*		

* Sloppy terminology; actually
Inaccuracy ±20% = Accuracy 80 – 120%
Imprecision 20%

BMV: Similarities, differences

Topic	EMA	FDA	ICH
Recovery	Not required (nonsense)	For methods employing extraction.	
		Extracted samples at L, M, and H QCs versus extracts of blanks spiked with the analyte post extraction (at L, M, and H).	
		Does not need to be 100%, but the extent of recovery of analyte and the IS should be consistent.	
Calibration curve	Blank (no analyte, no IS), zero (no analyte), $\geq$ six calibrators (optionally in replicates). If multiple analytes, separate CCs (nonsense).		
	Back-calculated concentrations $\pm$15% of nominal (except at LLOQ, where $\pm$20% of nominal). $\geq$75% must pass this criterion. If replicates are used, $\geq$50% must pass at a given level.		
Accuracy	Quality control samples prepared from stock solution different from calibrators.		
	Four levels (LLOQ, L, M, H; $\geq$five replicates): L $\leq$3$\times$LLOQ, M 30–50% of ULOQ, H $\geq$75% of ULOQ. At least three runs (LLOQ needed in only one of them).		
	One run of prospective study's size	Not required	One run of prospective study's size
	Back-calculated conc's $\pm$15% of nominal (except at LLOQ, where $\pm$20% of nominal).		

BMV: Similarities, differences

Topic	EMA	FDA	ICH
Precision	QCs of accuracy runs.		
	CV ≤15% (except at LLOQ, where ≤20%).		
Dilution integrity	Spiked samples >ULOQ, diluted with blank matrix.		
	≥ Five replicates per dilution factor.		
	Accuracy ±15% of nominal, precision CV ≤15%.		
Stability	Stock solution and working solutions of analyte and IS. Whole blood (covering time interval from draw to freezing of matrix; not required for the FDA). Long term (covering time interval from first clincial sample to end of bioanalytics). Bench-top / short term (from thawing to extraction). Processed samples (dry extract or in injection phase). Auto-sampler (duration of prospective run). Three freeze-thaw cycles.		
	L and H QC levels (at least triplicates).		
	Accuracy ±15% of nominal (precision no required).		

BMV: Similarities, differences

Topic	EMA	FDA	ICH
Re-injec-tion repro-ducibility	Recommended (QC levels)	Not mentioned	Recommended (QC levels)
	Back-calculated conc's $\pm15\%$ of nominal (except at LLOQ, where $\pm20\%$ of nominal).		
Matrix effect	Potential alteration of the analyte response due to interfering component(s) in the sample matrix.		
	At least six individual sources of matrix.		
	Case-by-case + one lipaemic + one haemolyzed	Lipaemic / haemolyzed not required	Recommended + one lipaemic + one haemolyzed
	At least triplicates at L and H QC levels.		
	Accuracy $\pm15\%$ of nominal, precision CV $\leq15\%$.		

BMV: Similarities

Partial validation

- **Required if study's samples not covered by the validated method**
  - Unexpected clustering of samples at one end of the calibration range
    - Re-analysis of samples (i.e., obtained with the original method) is not required.
    - Revise CC and QCs.
    - Revalidate the new range.
    - Open issue:
      » If the new range is lower than the original one, how 'far' should one go?
      » Whole blood stability and long term stability? The latter is a show-stopper.
  - Analytical site changes.
  - Change in sample volume, anticoagulant, storage conditions.
  - Change in sample processing.
  - Not mentioned in the GLs but logical for EMA and ICH.
    - Change in the size of a prospective run (A & P).

Fleming.

BMV: Similarities

Cross validation

- Data within a study from different fully validated methods.
- As above but different bioanalytical sites.
  - Not required if the same method is used.
- If possible done in advance.
- Same set of QCs analyzed.
  - Mean accuracy ±15% of nominal (wider if justified).

BMV: Similarities

Analysis of study samples

- **Analytical run.**
  - Blank sample (processed matrix without analyte and without IS).
  - Zero sample (processed matrix without analyte and with IS).
  - At least six calibrators.
  - At least three QC samples (L, M, H) in at least duplicate.
  - Study samples.
  - Preferrably processed in one batch.
    - If more than one batch (*e.g.*, limited by 96-well plates or more than one analyst), full set of calibrators and QCs in each batch.
  - Acceptance criteria applicable for the whole run.
  - In BE and crossover studies all samples of each subject should be analyzed in the same run.

BMV: Similarities

Analysis of study samples

- **Analytical run.**
  - **Acceptance criteria (AC).**
    - **Defined in the analytical protocol or in an SOP.**
    - **If a run consists of several batches, AC applicable to both the batches and the run (overall).**
    - **The latter takes presedence over the former (*i.e.*, the run might be still acceptable although one of the batches fails).**
    - **Accuracy of calibrators.**
      - » **Back-calculated concentrations within ±15% of nominal (±20% at LLOQ).**
      - » **At least 75% of calibrators must pass (≥6). Exclusion and re-evaluation possible.**
    - **Accuracy of QC samples.**
      - » **Back-calculated concentrations within ±15% of nominal.**
      - » **At least 67% of QC samples must pass (if replicates, exclusion is possible but not more than 50%).**

BMV: Similarities

Analysis of study samples

- **Analytical run.**
  - **Acceptance criteria (AC).**
    - **Accuracy and Precision of QC samples.**
      » **Should be reported for all accepted runs.**
      » **If A and/or P $>15\%$ additional investigation justifying this deviation. In case of BE this may result in rejection of the study.**
  - **(Re-) Integration.**
    - **Should be described in an SOP.**
      » **Original and final integration data documented at the analytical site and available upon request.**
      » **Cave! In many data systems the original integration is *not* saved, only the change is documented in the audit trail**

BMV: Similarities

Analysis of study samples

- Incurred sample reassessment (reanalysis) – ISR.
  - Validation based on spiked sample may not reflect the behavior of 'real world' samples (metabolites incl. back-conversion to the parent, co-medications, …).
  - ISR mandatory for BE.
    - Extent of testing depends on the analyte and the study samples, and should be based upon in-depth understanding of the analytical method and analyte(s).
    - However, as a guide, 10% of the samples should be reanalysed in case the number of samples is less than 1,000 samples and 5% of the number of samples exceeding 1,000 samples.
      Example: 1,200 samples. $ISR = 1,000 \times 10\% + 200 \times 5\% = 110$.

BMV: Similarities

Analysis of study samples

- **Incurred sample reassessment (reanalysis) – ISR.**
  - Assessment of the percent difference.

$$\%\text{difference} = 100 \frac{C_{repeated} - C_{initial}}{(C_{repeated} + C_{initial})/2}$$

  - %difference should not be >20% for at least 67% of ISRs.
  - Larger differences should be investigated.
    - » *Theoretically* that should not lead to rejection of a BE study.
    - » *Practically* expect a lot of problems.
  - However, this might be an artifical problem.
  - PhilipTimmerman of the European Bioanalysis Forum reported at the BioBridges meeting (Prague, September 2019) a survey where in only 2.1% of studies larger deviations were found.
  - Is this an artifact?

Open Issues

If in doubt

- **BEBA Forum**
  https://forum.bebac.at/

  – **Bioanalytics**
    https://forum.bebac.at/mix.php?category=7

  – **GxP / QC / QA**
    https://forum.bebac.at/mix.php?category=20

- **To post / reply you have to register first**
  https://forum.bebac.at/register.php

Hardware

Pentium FDIV bug (INTEL 1993)

- **Flaw in the x86 assembly language floating point divison.**
  - **Example**

$$\frac{4,195,835}{3,145,727} = 1.333\,739068902037589$$

$$\frac{4,195,835}{3,145,727} = 1.333820449136241002$$

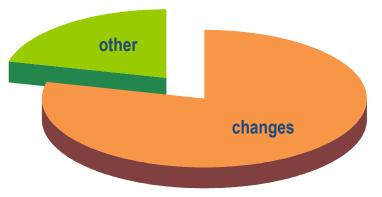 - **Costs for replacement: $475 million.**

Software

General Principles of Software Validation (FDA 2002)

- **Section 2.4: Regulatory Requirements for Software Validation**
  - 242 FDA Medical Device Recalls attributed to software failures (1992 – 1998).
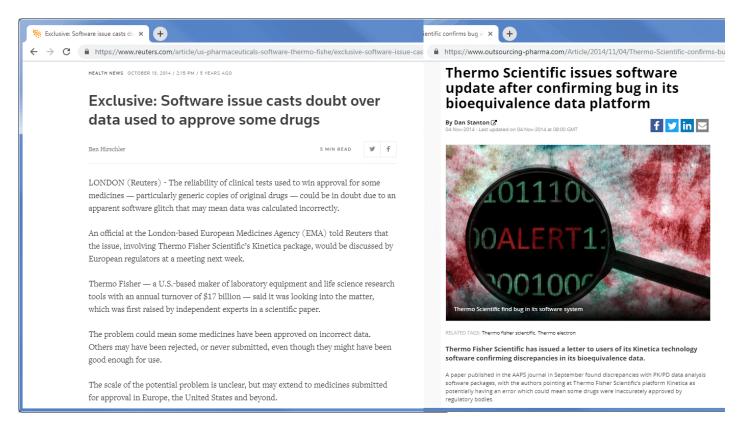 - 192 (79%) caused by software defects that were introduced when *changes* were made to the software after its initial pro-duction and distribution.

Software

… in bioequivalence: Is it validated?

Software

Reference data-sets in the public domain which allow users to PQ their software installations

design	sequences/ groups	vari-ances	R	SAS	PHX/ WNL	JMP	Stata	SPSS	OO Calc	Kinetica	Equiv-Test	Thoth-Pro	Statis-tica
2×2×2 Xover[1,2]	balanced	identical	☑	☑	☑	☑	☑	NT	☑	☑	☑	☑[a]	☑[b]
	imbalanced		☑	☑	☑	☑	☑	NT	☑	☒	☑	☒	☑[b]
2 groups parallel[3]	equal	equal	☑	☑	☑	☑	☑	NT	☑	☑	☑	–	☑
		unequal	☑	☑	☑	☑	☑	NT	☑	–	–	–	☑
	unequal	equal	☑	☑	☑	☑	☑	NT	☑	☒	–	–	☑
		unequal	☑	☑	☑[c]	☑	☑	NT	☑	–	–	–	☑
replicate, reference-scaling[4]	balanced, imbalanced, incomplete	equal, unequal	☑	☑	☑	☑	☑	☑	NT	–	–	–	☑

☑ passed   NT Not tested (yet)   [a] ≤100 subjects   [b] ≤500 subjects   [c] ≤ 1,000 subjects / group
☒ incorrect   –   Not implemented (design cannot be evaluated)

1. Schütz H, Labes D, Fuglsang A. *Reference Datasets for 2-Treatment, 2-Sequence, 2-Period Bioequivalence Studies.* AAPS J. 2014;16(6):1292–97. doi:10.1208/s12248-014-9661-0.
2. Moralez-Acelay S, de la Torre de Alvarado JM, García-Arieta A. *On the Incorrect Statistical Calculations of the Kinetica Software Package in Imbalanced Designs.* AAPS J. 2015;17(4):1033–4. doi:10.1208/s12248-015-9749-1.
3. Fuglsang A, Schütz H, Labes D. 2015. *Reference Datasets for Bioequivalence Trials in a Two-Group Parallel Design.* AAPS J. 2015;17(2):400–4. doi:10.1208/s12248-014-9704-6.
4. Schütz H, Tomashevskiy M, Labes D, Shitova A, González-de la Parra M, Fuglsang A. *Reference Datasets for Studies in a Replicate Design intended for Average Bioequivalence with Expanding Limits.* Manuscript submitted for publication 2019.

Thank You!
Open Questions?

Helmut Schütz
BEBAC
Consultancy Services for
Bioequivalence and Bioavailability Studies
1070 Vienna, Austria
helmut.schuetz@bebac.at