EUROPEAN MEDICINES AGENCY
SCIENCE MEDICINES HEALTH

21 September 2016
EMA/582648/2016
European Medicines Agency

Annex I

a. Approach compatible with CHMP guideline (Method A)

The approach envisaged when the current guideline was written was to simply use the same analysis method for replicate designs as is used for 2×2 trials.

```
proc glm data=replicate;
class formulation subject period sequence;
model logDATA= sequence subject (sequence) period formulation;
estimate "test-ref" formulation -1+1;
test h=sequence e=subject(sequence);
lsmeans formulation / adjust=t pdiff=control("R") CL alpha=0.10;
run;
```

For this model there is only one variance term estimated, $\sigma^2_{w,}$ the within subject variability.

b. Slight modification to approach compatible with CHMP guideline (Method B)

The same model as specified above could be used in PROC MIXED and subject specified as a random effect.

```
proc mixed data=replicate;
class formulation subject period sequence;
model logDATA= sequence period formulation;
random subject(sequence);
estimate "test-ref" formulation -1 1 / CL alpha=0.10;
run;
```

This means there are two variance terms estimated σ^2_w and σ^2_b, as a distribution is also fitted to the between subject variability. If subject is a fixed effect (as in the previous model) each subject is treated as being selected in some way rather than being sampled from a random distribution and a subject effect is estimated individually for each patient as is done for the period effect.

30 Churchill Place ● Canary Wharf ● London E14 5EU ● United Kingdom
Telephone +44 (0)20 3660 6000 **Facsimile** +44 (0)20 3660 5555
Send a question via our website www.ema.europa.eu/contact

This model will give the same results as Method A if all subjects included in the analysis provide data for all treatment periods.

c. Method C

The FDA Guidance for Industry document "Statistical approaches to establishing bioequivalence" specifies the code to be used for the analysis of replicate designs using PROC MIXED.

```
proc mixed data=replicate;
classes sequence subject period formulation;
model logDATA= sequence period formulation / ddfm=satterth;
random formulation/type=FA0(2) sub=subject G;
repeated/grp=formulation sub=subject;
estimate 'test-ref' formulation -1 1/ CL alpha=0.10;
run;
```

This model allows a different subject effect for each formulation (i.e. a subject by formulation interaction), and therefore has 5 variance terms (within subject for reference, within subject for test, between subject for test, between subject for reference, covariance for between subject test and reference – the last three are combined to give the subject ×formulation interaction variance component.)

This model will provide the same point estimate as methods A and B if all subjects provide data for all treatment periods. However it will generally give wider confidence intervals than those produced by methods A and B.

Results

Data set I

The following data reflect a four period crossover study where subjects receive both test and reference twice, with some subjects providing data for only a subset of the treatment periods. Results obtained with methods A, B and C are shown in the following table.

	Point estimate	**90% confidence interval**
Method A (guideline recommended)	115.66	107.11, 124.89
Method B (random effects)	115.73	107.17, 124.97
Method C (random effects with interaction)	115.66	107.10, 124.89

Within subject CV% (from method C) – reference 47.3%, test 35.3%

The results are generally very similar although missing treatment periods for some subjects causes the results to be different for all three approaches.

Data set II

Data of a three period crossover study where all subjects receive reference twice and test once were analysed using Methods A, B and C.

The results are given in the Table below:

	Point estimate	**90% confidence interval**
Method A (guideline recommended)	102.26	97.32, 107.46

Method B (random effects)	102.26	97.32, 107.46
Method C (random effects with interaction)	102.26	97.05, 107.76

Within subject CV% (from method C) – reference 11.5%

As there are no subjects with missing treatment periods the results from methods A and B are identical, and the point estimate is the same for all three approaches. Method C gives wider intervals.

Alternative computer programs

SAS (version 9.1, SAS Institute Inc., NC) was used in the previous computations. Results obtained by alternative, validated statistical programs are also acceptable except spreadsheets because outputs of spreadsheets are not suitable for secondary assessment.

Estimating the within subject variability

The guideline introduces the possibility of widening the acceptance limits for C_{max} if the within-subject variability for the reference product is greater than 30%. This is calculated using:

$$CV(\%) = 100\sqrt{e^{s^2_{WR}} - 1}$$

The widening is on a smooth function, i.e. the permitted widening increases as the variability increases (to a maximum of 50%). It is not an all or nothing criteria with 30% being a critical point.

An advantage of Method C is that it directly calculates s^2_{wr}. However, sometimes the algorithm fails to converge. For that reason the preferred way to get an unbiased estimate of σ^2_{wr} is using the data from the reference product only.

The following code removes all the test data from the data-set and then fits a model where the residual variance corresponds to the within subject variance for the reference product.

```
data var;
set replicate;
if formulation='R';
run;

proc glm data=var;
class subject period sequence;
model logDATA= sequence subject (sequence) period;
run;
```

Results obtained with the different methods for Data Set I and II are summarised in the table below.

Reference within subject CV%

	Model A/B	Model C
Data set I	47.0%	47.3%
Data set II	11.2%	11.5%

The data shows that the variability estimates given by the two approaches are very similar for these examples. There is no dependence on random effects mixed models to estimate within subject variability for a formulation.