Acta Poloniae Pharmaceutica – Drug Research, Vol. 65 No. 6 pp. 743–747, 2008 ISSN 0001-6837
Polish Pharmaceutical Society

SHORT COMMUNICATIONS

APPLICATION OF CONFIDENCE INTERVALS TO BIOANALYTICAL METHOD VALIDATION – DRUG STABILITY IN BIOLOGICAL MATRIX TESTING

PIOTR J. RUDZKI[1]* and ANDRZEJ LEŚ [2,3]

[1] Pharmacology Department, [2] Strategic Analysis Department, Pharmaceutical Research Institute, 8 Rydygiera St., 01-793 Warsaw, Poland,
[3] Faculty of Chemistry, Warsaw University, 1 Pasteura St., 02-093 Warsaw, Poland

Keywords: confidence intervals, bioanalysis, validation, stability, pharmacokinetic studies, statistics

Pharmacokinetic studies, e.g. innovative drug pharmacokinetic evaluation or generic drug bioequivalence testing, can be requested by the regulatory authorities during the registration process of medicinal products. The drug determination in biological matrices, e.g. blood, serum, plasma and urine, is performed by means of the instrumental analysis, including high-performance liquid chromatography (HPLC). To assure the reliability of results a selected bioanalytical method should be validated according to the specific requirements (1). A thorough drug stability testing in a biological matrix is one of the important steps within the validation. The U.S. Food and Drug Administration (FDA) Guidance describes a standard approach of direct comparison of analytical results for the studied and reference samples, but also suggests that other statistical approaches may be used. Generally, in recently published papers concerning bioanalytical methods that standard approach, without any statistical treatment, is applied to stability testing. It seems to be in opposition to the current trends in analytical chemistry, e.g. strong interest in chemical metrology and uncertainty of measurement.

The presented method of statistical stability evaluation is based on the application of confidence intervals. The idea was first reported and discussed by Timm et al. in 1985 (2). It was stressed that without confirmation of stability, the subsequent pharmacokinetic data are questionable. Thus the application of confidence intervals was proposed as a practical procedure based on statistical treatment, which takes into account the precision of measurement as well as assigns a degree of certainty to the conclusions concerning drug stability. It was declared that the application of the method indirectly enhances patient security, because the compounds are classified as stable only if relevant degradation can be excluded with high probability (95%). Therefore, the subsequent results of pharmacokinetic studies are more reliable, which enables more accurate conclusions concerning the clinical use of an investigated drug. It was argued that the application of confidence intervals in stability investigation is superior to the routine Student *t*-test (2). The latter might show statistically significant difference in stability, which might not be relevant in the case of pharmacokinetic studies.

The method of statistical stability evaluation presented by Timm et al. was based on the assumption of equal variances for the studied and reference samples (2). To allow application of confidence intervals – when this condition is not met – a new method is proposed.

EXPERIMENTAL

Theoretical background

Assumptions

The log-normal distribution of measurement results for studied (x) and reference (z) samples was assumed. This assumption was verified and confirmed with the use of two normality tests (Shapiro-Wilk and Kolmogorov-Smirnov) implemented in the SAS suite of programs (4).

Confidence Interval calculation method

The step by step confidence interval calculation method was presented in Figure 1. After determina-

* Corresponding author: e-mail: p.rudzki@ifarm.waw.pl

tion of drug and/or metabolite concentrations in studied and reference samples, the possible outliers in each group of measurement results were identified. The presence of outliers strongly influences variance calculation and causes widening of respective confidence interval. To avoid this situation, the suspected results were identified by application of two-sided Q-Dixon test at significance level $\alpha = 0.10$. Because the statistically significant outliers are not always relevant, the final decision concerning the inclusion or exclusion of a suspected result in further calculations was made by the analyst. In each individual case the judgement was made after studying the possible sources of error, e.g. poor chromatography, equipment malfunction or improper sample preparation. Irrespective to statistical analysis which pointed at some results to be rejected as outliers, all those results which were within the method accuracy limits [e.g. 85-115% of nominal value (1)], were included in further calculations. The next step was the logarithmic transformation of original data, which corrected the possible skew of their distribution.

The proper determination of confidence interval limits requires variance equality verification. To compare variances in studied (s_x^2) and reference (s_z^2) samples, the *F*-test was applied at significance level $\alpha = 0.01$. The lower and upper limits of two-sided 90% confidence interval were calculated either using pooled variance (in the case of equal variances) or individual variances of the studied and reference samples (2-4).

Method A. Equal variances

The pooled variance (s_p^2) was calculated according to the formula F1, which included the influence of variances of studied (s_x^2) and reference (s_z^2) samples as well as sample size in both groups (n_x and n_z).

$$s_p^2 = \frac{(n_x - 1)\cdot s_x^2 + (n_z - 1)\cdot s_z^2}{n_x + n_z - 2} \qquad [F1]$$

The lower (LL) and upper (UL) confidence interval limits were calculated according to the formulas F2 and F3, respectively. The number of degrees of freedom was defined as $df = n_z + n_x - 2$, while $t_{\alpha/2,df}$ was the value of the Student *t*-distribution quantile at selected significance level.

$$LL = e^{\overline{\ln x} - \overline{\ln z} - t_{\alpha/2,df}\cdot\sqrt{s_p^2\cdot\left(\frac{1}{n_x}+\frac{1}{n_z}\right)}}\cdot 100\% \qquad [F2]$$

$$UL = e^{\overline{\ln x} - \overline{\ln z} + t_{\alpha/2,df}\cdot\sqrt{s_p^2\cdot\left(\frac{1}{n_x}+\frac{1}{n_z}\right)}}\cdot 100\% \qquad [F3]$$

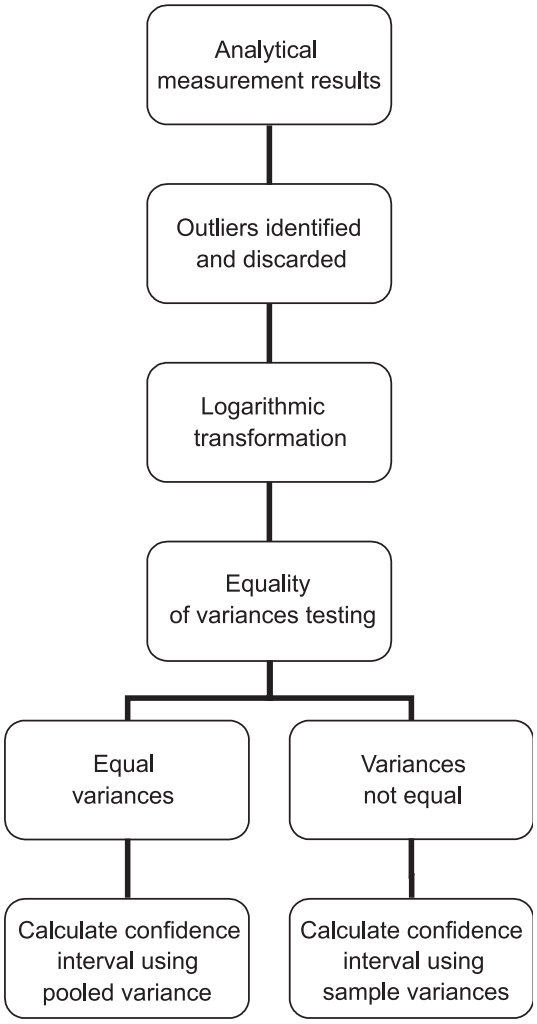

Figure 1. Confidence interval calculation block diagram.

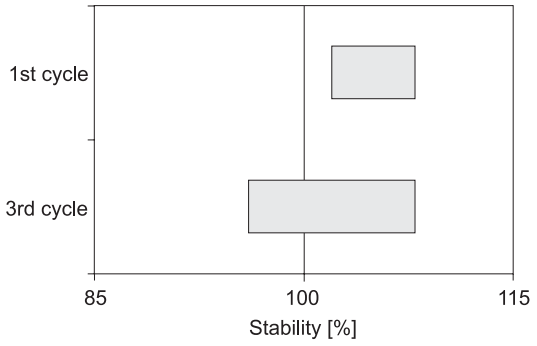

Figure 2. Calculated confidence intervals.

Method B. Variances not equal

The number of degrees of freedom was calculated according to the formula F4, with the result being rounded down to the nearest integer.

Table 1. Results of gemcitabine freeze-thaw stability study in rat plasma (–70°C).

Nominal GMC concentration [μg/mL]	Freshly prepared	1st cycle (24 h)	3rd cycle (72 h)
	[μg/mL]	[μg/mL]	[μg/mL]
7.50	7.10	7.20	6.92
	7.18	7.57	7.03
	7.24	7.67	7.23
	7.33	7.84	7.84
	7.33	7.86	7.93
Mean	7.236	7.628	7.390
s	0.099	0.268	0.466

Table 2. Confidence interval calculation after the 1st freeze-thaw cycle.

Nominal GMC concentration [μg/mL]	Freshly prepared		1st cycle (24 h)	
	z [μg/mL]	ln(z)	x [μg/mL]	ln(x)
7.50	7.10	1.96	7.20	1.97
	7.18	1.97	7.57	2.02
	7.24	1.98	7.67	2.04
	7.33	1.99	7.84	2.06
	7.33	1.99	7.86	2.06
Mean	7.24	1.98	7.63	2.03
s	0.10	0.01	0.27	0.04
s^2	0.00983	0.00019	0.07177	0.00127

Table 3. Confidence interval calculation after the 3rd freeze-thaw cycle.

Nominal GMC concentration [μg/mL]	Freshly prepared		3rd cycle (24 h)	
	z [μg/mL]	ln(z)	x [μg/mL]	ln(x)
7.50	7.10	1.96	6.92	1.93
	7.18	1.97	7.03	1.95
	7.24	1.98	7.23	1.98
	7.33	1.99	7.84	2.06
	7.33	1.99	7.93	2.07
Mean	7.24	1.98	7.39	2.00
s	0.10	0.01	0.47	0.06
s^2	0.00983	0.00019	0.21755	0.00394

$$df = \frac{\left(\frac{s_x^2}{n_x} + \frac{s_z^2}{n_z}\right)^2}{\frac{1}{n_x - 1} \cdot \left(\frac{s_x^2}{n_x}\right)^2 + \frac{1}{n_z - 1} \cdot \left(\frac{s_z^2}{n_z}\right)^2} \quad [F4]$$

The lower (LL) and upper (UL) confidence interval limits were calculated according to the formulas F5 and F6, respectively.

$$LL = e^{\overline{\ln x} - \overline{\ln z} - t_{\alpha/2,df} \cdot \sqrt{\frac{s_x^2}{n_x} + \frac{s_z^2}{n_z}}} \cdot 100\% \quad [F5]$$

$$UL = e^{\overline{ln\,x}-\overline{ln\,z}+t_{\alpha/2,df}\cdot\sqrt{\frac{s_x^2}{n_x}+\frac{s_z^2}{n_z}}}\cdot 100\% \qquad \text{[F6]}$$

RESULTS

As an example of confidence intervals application, the study of gemcitabine freeze and thaw stability in rat plasma was performed. The gemcitabine concentrations determined in reference samples (i.e. freshly prepared) as well as after 1st and 3rd freeze-thaw cycles are presented in Table 1.

After the 1st freeze-thaw cycle samples were analyzed, the original data were logarithmically transformed (Table 2) and then the variances were compared. The *F*-test did not reveal significant differences in variances for studied and reference samples, as experimental *F*-value ($F_{exp} = s_x^2 / s_z^2 = 6.73$) was lower than the corresponding quantile ($F_{(0.01, 4, 4)} = 15.98$). The confidence interval was calculated according to the Method A presented above. The pooled variance ($s_p^2 = 0.00073$) and number of degrees of freedom ($df = 8$) were calculated. To calculate 90% confidence interval, the Student *t*-distribution quantile at significance level $\alpha/2 = 0.05$ was used ($t_{(0.05, 8)} = 1.86$). The calculated confidence interval (*90% C.I.* = 102-109%) was found to be within the acceptance limits of 85-115%. As 100% value was not included within the confidence interval, one can notice a statistically significant increase in the response during the analysis of studied samples. However, this increase was not relevant from the pharmacokinetic point of view, because the calculated stability was embraced completely by predefined acceptance limits (1).

A similar approach was applied for the data after the 3rd freeze-thaw cycle. The samples were analyzed, the original data were logarithmically transformed (Table 3) and then the variances were compared. The *F*-test showed significant differences in variances for studied and reference samples, as experimental *F*-value ($F_{exp} = s_x^2 / s_z^2 = 20.89$) was higher than the corresponding quantile ($F_{(0.01, 4, 4)} = 15.98$). Thus the confidence interval was calculated according to the Method B presented above. The number of degrees of freedom was calculated according to the formula F4 ($df_{calc} = 4.38$) and then rounded down to the nearest integer ($df_{round} = 4$). To calculate 90% confidence interval, the Student *t*-quantile at significance level $\alpha/2 = 0.05$ was used ($t_{(0.05, 4)} = 2.13$). The calculated confidence interval (*90% C.I.* = 96-108%) was also found to be within the acceptance limits of 85-115%. As 100% value was included within the confidence interval, the change in response was neither statistically significant, nor relevant. The graphical illustration of calculated confidence intervals was presented in Figure 2.

DISCUSSION AND CONCLUSION

Two decades after the publication of Timm et al. paper (2), the topic of their research is still up-to-date. Since then, the bioanalytical method validation has become the subject of FDA regulation (1), but some details of the guidance are continuously under discussion. One of those topics is drug stability testing in biological samples, where the above mentioned over 20-year-old paper presents a very interesting point of view. Furthermore, this view corresponds to the recent trends in analytical chemistry, especially strong interest in the uncertainty of measurement.

The new method proposed in this paper enables wider application of the approach presented by Timm et al. The new method is in accordance with current FDA regulations (1) and therefore can be applied to the validation of any bioanalitycal method used in human studies requiring pharmacokinetic evaluation, e.g. clinical pharmacology, bioavailability and bioequivalence studies. The confidence intervals may be used in different stability investigations including, but not limited to: long-term stability under sample storage conditions, short-term stability in biological matrix at room temperature, stability after freezing and thawing, as well as post-preparative stability (in an autosampler). The approach can be also applied to the solutions stability testing, provided narrower acceptance limits are pre-defined.

It was observed that the statistically incorrect application of the previously described method under the conditions of unequal variances resulted in a narrower confidence interval. It could be illustrated by the confidence interval calculation for stability after 3rd freeze-thaw cycle, where the condition of equal variances was not met. The confidence interval calculated according to the previous method was $90\%\ C.I._{prev}$ = 96.7-107.6%, while the calculation according to the new method (Method B) resulted in $90\%\ C.I._{new}$ = 95.9-108.4%. In the case of a confidence interval lying near the acceptance limit, the difference of 1.6% might lead to the misinterpretation of results.

The described method of stability testing could be considered superior to the standard approach of direct comparison of analytical results for the studied and reference samples, because confidence inter-

vals take into account method precision as well as assign a degree of certainty to the conclusions concerning drug stability or instability. The application of confidence intervals during bioanalytical method validation can improve the quality of data obtained during pharmacokinetic studies and therefore seems to be an interesting area for further studies.

REFERENCES

1. Guidance for Industry: *Bioanalytical Method Validation*. U.S. Department of Health and Human Services. FDA. CDER. CVM. May 2001.
2. Timm U., Wall M., Dell D.: J. Pharm. Sci. 74, 972 (1985).
3. Bolton S.: Pharmaceutical Statistics, Practical and Clinical Applications, 2nd Ed., Marcel Dekker, Inc., New York and Basel 1990.
4. SAS ver. 9.1, SAS Institute Inc., Cary, NC, U.S.A.