ELSEVIER

European Journal of Pharmaceutics and Biopharmaceutics 53 (2002) 9–14

European Journal of Pharmaceutics and Biopharmaceutics

www.elsevier.com/locate/ejphabio

Research paper

Typical variability and evaluation of sources of variability in drug dissolution testing

Martin Siewert, Ludwig Weinandy*, David Whiteman[a], Catherine Judkins[b]

[a]*Aventis Pharma AG, Frankfurt am Main, Germany*
[b]*Cambridge University, Cambridge, UK*

Received 2 January 2001; accepted in revised form 24 August 2001

Abstract

To investigate variability in dissolution testing an international collaborative study was performed by 29 laboratories. Glibenclamide (glyburide) tablets were used in the investigation in which multipoint dissolution profiles were established using USP paddle apparatus. In contrast to a previous report, the variability of the glibenclamide dissolution data was significantly lower. Total variances (s^2) were found to range from 18.34–44.18, Between Laboratory and Between Analyst variances (synthetic value) ranged from 12.9–38.7 and the Within Analyst variances ranged from 5.08–5.78. The dissolution profiles and corresponding variances obtained by laboratories with little or no experience in glibenclamide dissolution testing were similar to those obtained by more experienced laboratories, indicating that the test, especially when designed as multiple point dissolution testing, is sufficiently robust and capable of identifying differences in a manufacturing process or drug formulation. The smallest statistically detectable mean difference between two dissolution runs was calculated (95% CI) to be 7% for one analyst, or 5% if two analysts were to perform the dissolution tests.

Keywords: Variability in drug dissolution testing; Collaborative study with glibenclamide/glyburide tablets (marketed name is Daonil); Robustness and capability of dissolution testing

1. Introduction

Drug dissolution testing is a fundamental part of drug development and manufacture, being used primarily as a method of quality control. The in vitro dissolution characteristics of a given drug provides information on (a) how that drug may behave in vivo; (b) the consistency of batches; (c) the homogeneity of dosage forms within a given batch; and (d) the impact of changes in composition, process or site of manufacture.

A recent study [1] organised by the International Pharmacists Federation (FIP) reported high variability in dissolution testing, both with United States Pharmacopoeia (USP) calibrator tablets and a marketed drug product (Daonil, batch N765, was used as an example). The authors concluded that dissolution itself gives rise to an intrinsic variability and discussed that the method could *not* be used reliably to indicate changes in the manufacturing process or to determine whether dissolution test equipment complies with USP specifications.

In the above-mentioned FIP publication, 28 official or semi-official governmental laboratories in 23 countries submitted data on dissolution tests using Daonil (glibenclamide) tablets. The high variability reported in the study led Aventis Pharma AG to initiate an internal international collaborative study in order to investigate the sources of this variability and to determine the discriminatory power of the dissolution testing method used for glibenclamide tablets. Twenty-five manufacturing site quality control laboratories and four research and development laboratories in 19 countries were included in a collaborative study identical to the one outlined by the FIP. Of the 29 laboratories, 11 had no previous experience with Daonil testing and 8 of the laboratories typically tested less than one batch per month.

Tests with basket apparatus were not included in the study since all historical data and experience with dissolution testing was carried out with paddle equipment.

It was not in the interest of this study to re-evaluate the quality of the glibenclamide tablet batch or to qualify or disqualify the 'worst' or 'best' laboratories, nor was it intended to set new limits of acceptability for this product. Furthermore, it was not the objective to defend existing or to suggest alternative pharmacopoeial methods of calibration

* Corresponding author. Aventis Pharma AG, Ludwig Weinandy, Indrustriepark Höchst, H790, D-65926 Frankfurt am Main, Germany.
Tel.: +49-69-305-80042; fax: +49-69-305-20522.
E-mail address: ludwig.weinandy@aventis.com (L. Weinandy).

0939-6411/02/$ - see front matter
PII: S0939-6411(01)00213-2

or to propose a change in the quality control testing of glibenclamide tablets.

2. Experimental

2.1. Study Protocol

A study protocol identical to the one used in the FIP study was executed by the 29 study participants, including 25 quality control laboratories and 4 research and development laboratories. The participants were requested to submit the results of the dissolution testing and to include the results of their most recent USP [3] Apparatus Suitability Tests for paddle apparatus.

As batch N765 (which had been used in the previous study [1]) was no longer available in sufficient quantities at the time of the present study, another randomly selected commercial batch of Daonil was chosen and the tablets sent to each laboratory. In addition to the 20 glibenclamide tablets (5 mg, batch U740, manufactured in February 1998) each laboratory also received glibenclamide reference standard (Aventis Pharma, lot U329) together with the certificate of analysis. The participants were requested to specify the dissolution apparatus used, including the type of UV spectrophotometer and the level of the analyst's experience in Daonil dissolution testing.

The apparatus used was required to comply with the mechanical specifications for wobble, centering, levelness, verticality of shaft, rotation speed, temperature and vibration, as stated or implied in the USP [3] section $<711>$.

Analysts were asked to carry out the dissolution run with six individual tablets each, using paddle apparatus at 75 rpm. Absorbance values were recorded using UV spectrophotometers at a wavelength of 225 nm.

The dissolution medium used in the tests was prepared as follows: anhydrous monobasic potassium phosphate, KH_2PO_4 (68.05 g) and sodium hydroxide (5 M, 78.2 ml) were dissolved in water (10 l) and the pH was adjusted to 7.40 ± 0.05. It should be noted that this buffer solution is not the same as that used in the registered control test for Daonil tablets but it is the dissolution medium used in the FIP study and was thus chosen in this case to maintain consistency and comparability.

Laboratories were asked to use only reagents and solutions of the required analytical grade and freshly prepared, de-aerated dissolution media.

Tablet masses were recorded prior to commencing the analyses and an inert filter was used to isolate the specimens in order to avoid adsorption.

The percentage of glibenclamide dissolved in the phosphate buffer solution (900 ml) at pH 7.4 was measured at the following sampling times: 10, 20, 30, 60, 90 and 120 minutes. The results from all analysts at these sampling times were reported to one decimal place using the format provided.

In order to assess the analyst as a source of variability, laboratories were asked to have a second analyst repeat the entire procedure on a different day using the same equipment, freshly prepared reagents, reference standard solution and dissolution media. Apart from this, all other test parameters were identical to those outlined in the FIP study protocol.

2.2. Method of data analysis

Once all data had been collated an ANOVA variance component analysis was carried out to estimate the following sources of variation:

- Between Laboratory.
- Between Analyst (among the two analysts nested within a laboratory).
- Within Analyst (nested within a laboratory i.e. within runs).

All data from two of the laboratories were excluded from the ANOVA analysis due to inconsistencies within their dissolution profiles. These omissions were not merely statistical outliers, but the data showed implausible deviations (i.e. decreases in concentration over time) from the monotonic increase in the percentage of drug released.

Since the FIP study called for only one analyst at each laboratory to perform the testing, the statistical analysis employed by the authors was an ANOVA variance component analysis with the following sources of variation estimated:

- Between Laboratory.
- Within Analyst (or Within laboratory, since analyst and laboratory are confounded).

This study therefore has no 'Between Analyst' component. In order to make a direct comparison between the two studies it was necessary to generate a synthetic variance from the Aventis Pharma data by adding the Between Laboratory and Between Analyst variance components, which corresponds to the Between Laboratory variance component of the FIP study. The degrees of freedom associated with the synthetic variance were estimated using the Saitterwaite approximation (Table 1).

It was also of interest in the Aventis Pharma study to determine the smallest statistically detectable difference between two dissolution profiles. At each dissolution timepoint, the following procedure was adopted to calculate how large the true, but unknown, difference between two profiles means needed to be in order to have at least a 95% chance of detecting that mean difference:

The synthetic variance was computed using the following method:

- If one analyst at each of the two laboratories conducts one dissolution run the synthetic variance is computed by adding the Between Analyst variance component and the

Table 1
Variance component estimation and comparison[a]

Source	Aventis Daonil study						FIP study				Comparison	
	Mean percent released	df	Variance component	Synthetic variance	CV (%)	df	Mean percent released	Variance component	CV (%)	df	Ratio	OSL
10-min												
Between Labs		26	31.20	38.29	17.78	26		127.00	28.39	26	3.32	0.0016
Between Analysts (lab)		27	7.09									
Within Analysts (lab)		269	5.23		6.57			14.20	9.49	135	2.72	0.0000
Total		322	43.52		18.96			141.20	29.93		3.24	
	34.8						39.7					
20-min												
Between Labs		26	15.66	20.52	10.15	26		97.80	19.70	26	4.77	0.0001
Between Analysts (lab)		27	4.86									
Within Analysts (lab)		269	4.96		4.99			9.60	6.17	135	1.94	0.0000
Total		322	25.48		11.31			107.40	20.64		4.22	
	44.62						50.2					
30-min												
Between Labs		26	10.99	16.26	7.84	26		116.10	18.20	26	7.14	0.0000
Between Analysts (lab)		27	5.27									
Within Analysts (lab)		270	5.60		4.60			20.20	7.59	135	3.61	0.0000
Total		323	21.86		9.09			136.30	19.72		6.24	
	51.43						59.2					
60-min												
Between Labs		26	7.06	11.94	5.41	26		137.50	16.15	26	11.52	0.0000
Between Analysts (lab)		27	4.88									
Within Analysts (lab)		270	5.25		3.59			12.90	4.95	135	2.46	0.0000
Total		323	17.19		6.49			150.40	16.89		8.75	
	63.88						72.6					
90-min												
Between Labs		26	7.50	12.76	5.04	26		143.80	14.84	26	11.27	0.0000
Between Analysts (lab)		27	5.26									
Within Analysts (lab)		270	4.95		3.14			15.90	4.94	135	3.21	0.0000
Total		323	17.71		5.94			159.70	15.64		9.02	
	70.82						80.8					
120-*min*												
Between Labs		26	8.56	13.78	4.87	26		129.40	13.26	26	9.39	0.0000
Between Analysts (lab)		27	5.22									
Within Analysts (lab)		270	5.59		3.10			8.80	3.46	135	1.57	0.0016
Total		323	19.37		5.78			138.20	13.70		7.13	
	76.18						85.8					

[a] OSL, observed significant level (aka, *P*-value). If the OSL is less than 0.05, then the Aventis Daonil Study variance is statistically smaller than the FIP Study variance. Saitterwaite's approximation used to estimate the df for the synthesized variance.

Within Analyst variance component and dividing by six.

- If two analysts at each of the two laboratories each conduct one dissolution run the synthetic variance is computed by adding the Between Analyst variance component divided by 2 and the Within Analyst variance component and then dividing by 12.

The synthetic variance was multiplied by two and the square root taken. This was called the standard error of the mean difference.

The standard error of the mean difference was multiplied by 1.645 to yield the detectable difference.

3. Results

Twenty-nine Aventis Pharma laboratories (comprising 25 quality control laboratories and four research and development laboratories) submitted their results. Following a check for inconsistencies, all data submitted from two laboratories were excluded from further evaluation as discussed previously. The average percentage of drug dissolved at each of the sampling times and the variability in the percentage released were found to lie within the expected range.

From Table 1 it can be seen that the Total CV values in the Aventis Pharma study lay between 5.78 and 18.96%,

with the higher values being observed at the earlier sampling points. These values are considerably lower than those indicated by the FIP study, which ranged from 13.70 to 29.93%. However, the FIP study also showed higher variation at the earlier sampling points.

To determine the possible sources of variability in dissolution testing, the Between Laboratory and Between Analyst variances were investigated. In view of the modified Aventis Pharma study design, a synthetic variance term was calculated (see Section 2.2) from the Between Laboratory and Between Analyst data which could then be compared with the Between Laboratory variance term of the FIP study. The variance component estimation and the comparison of the data for both studies are shown in Table 1. In addition, the *F*-test for equality of variances and the associated observed significance level (*P*-value) are included in the table.

The Between Analyst variance (Aventis study only) was observed to lie between 4.86 and 7.09 and remained relatively constant (ca. 5) from the 20-min time point onwards. This indicates that the variation between analysts within the same laboratory is *not* the major contributor to the Total Variance found in dissolution testing.

The synthetic variance terms (= Between Laboratory variance + Between Analyst variance) of the Aventis study ranged from 38.29 at the earliest sampling time to 13.78 at the 120-min time point. The higher value found at the 10-min interval dropped significantly upon reaching the 20 and 30 min sampling times and then remained at around 13 for the remainder of the testing. These values are considerably smaller than those indicated in the FIP study in which the Between Laboratory variance ranged from 127 at the beginning to 129 at the end of the test. At 10 and 20 min, the Between Laboratory variance components for the FIP study were ca. four times larger than the corresponding variance components in the Aventis Pharma study and at later time points this increased to values which were 7–10 times larger than those of the current study. The trend observed in the Aventis Pharma study of higher variances at earlier sampling times was not reflected in the FIP study. No correlation between variability and dissolution time point was discovered and, in contrast to the present study, the FIP results showed that the greatest variations between laboratories were at the 60 and 90 min intervals. The variability at these sampling times in the Aventis study was at its lowest and was reasonably constant. This comparison clearly shows that the variation between laboratories in the FIP study was very large indeed. Although the Aventis study also shows significant variation between laboratories, the values are not nearly as high as those found previously.

The Within Analyst variance components for both studies are also listed in Table 1. These values were found to be approximately constant in the Aventis Pharma study across all sampling points, ranging from 4.95 to 5.60. Again, these figures indicate that variability between the test runs carried out by each analyst is *not* a major contributor to the overall variability in dissolution testing. The corresponding Within Analyst values of the FIP study were 1.5–3 times larger than the present study values and, as with the Aventis study, showed no correlation between variance and sampling time.

The Total Variability in dissolution testing can be expressed as the sum of the incremental variances: Between Laboratory, Between Analyst and Within Analyst. From the above analysis it can be deduced that the major contribution to the Total Variability comes from the Between Laboratory variance component. Over the six dissolution time-points, variances between 31.20 and 7.06 were observed, with the higher values being at the earlier sampling times. This was in agreement with the findings of the previous study, which also found inter-laboratory variance to be the most significant. However, the values found in the current study were considerably lower.

The smallest detectable difference between the dissolution profiles for each single dissolution time point was calculated from the Between Analyst and Within Analyst variance components (Section 2.2). The resulting standard error of mean difference and the detectable differences are shown in Table 2. If one analyst in each laboratory was to carry out the dissolution testing, then a difference of 7% in the amount of drug released after 10 min would be determinable and a difference of 6% in the amount of drug released would be discernible at the remaining sampling times (95% CI). However, if two analysts in each laboratory perform the dissolution test, then a difference of 5% (at the 10 min interval) or 4% (at the remaining intervals) could be detected (95% CI).

4. Discussion

The results of this study have led us to believe that dissolution testing is robust. In the Aventis Pharma study no evidence could be found for the high variability observed in previous findings (CV from 13.7% up to 29.9%) – the CV values observed in the current study were considerably lower, ranging from 5.8% to 19.0%.

The Between Laboratory variability factor appeared to contribute the most to the Total Variability of the dissolution testing and the values of this component were higher at the earlier dissolution time points. However, this variability was found to be statistically much smaller in the Aventis Pharma study than in the FIP report. Hence, although variability does exist between the Aventis Pharma laboratories (and this fact does require some attention) there is considerably less variation between these laboratories than those that took part in the FIP study.

Thus, in order to evaluate any changes in manufacturing process or site by the dissolution method it would be best to have the products from the two sites tested in the *same* laboratory. This would eliminate the more significant source of variability and hence greatly reduce the overall Total Variability. In order to further reduce this total value, only one analyst should be allowed to perform the test. This

Table 2
Detectable difference estimation

Source		One Analyst per lab		Two Analysts per lab	
	Variance component	Standard error of mean difference	Detectable difference	Standard error of mean difference	Detectable difference
10-min dissolution					
Between Analysts (lab)	7.09				
Within Analysts (lab)	5.23	3.98	7	2.82	5
20-min dissolution					
Between Analysts (lab)	4.86				
Within Analysts (lab)	4.96	3.39	6	2.40	4
30-min dissolution					
Between Analysts (lab)	5.27				
Within Analysts (lab)	5.60	3.54	6	2.50	4
60-min dissolution					
Between Analysts (lab)	4.88				
Within Analysts (lab)	5.25	3.41	6	2.41	4
90-min dissolution					
Between Analysts (lab)	5.26				
Within Analysts (lab)	4.95	3.52	6	2.49	4
120-min dissolution					
Between Analysts (lab)	5.22				
Within Analysts (lab)	5.59	3.49	6	2.47	4

would rule out the Between Analyst variability factor, which was found to be the second greatest contributor to the overall variability, albeit a much smaller one.

The smallest statistically detectable difference between two profiles was found to be 6–7% depending on the time point, with one analyst carrying out the dissolution testing. If two analysts were to perform the tests, a dissolution mean difference of at least 4–5% would be discernible (95% CI). Although evaluation of the study results did show that variability was present in the dissolution method, the variances are in such a range that would still allow the use of glibenclamide tablets to detect changes in manufacturing site or processes.

Other possible sources of variability in dissolution testing, which have not been investigated in this study, could be the differences in the types of apparatus used and also in the preparation and de-aeration of the dissolution media. It was not surprising that the highest variability occurred with the Between Laboratory component since although all apparatus were USP compliant, differences in makes and models of equipment is bound to play some role in variability. Differences in climate etc. must also be taken into consideration. The general trend in both studies was that the highest variance figures were found to occur at the earliest sampling time (10 min) and, after this, the individual components of variability (Between Analyst and Within Analyst variance) seemed to contribute equally to the Total Variability. Perhaps this is an indication that the 10 min time point is too early to take measurements.

In answering the question as to why the Aventis Pharma study results were so different from those obtained by the previous study of identical design, the following points could be possible factors: Firstly, the analysts in the Aventis laboratories are not trained in Daonil testing (i.e. tablet-specific testing), but rather they receive specific training in the method of *dissolution* testing. This would offer an explanation of the lower observed variances (Between Laboratory, Between Analyst, Within Analyst) in the Aventis study and the consistency of the results obtained by laboratories that were experienced/inexperienced in Daonil testing. It could thus be deduced that dissolution testing *is* reproducible providing the necessary training is given.

The second influence in the Aventis Pharma study design may have been due to the fact that a reference standard was supplied to the laboratories together with the samples of drug product and this could also have contributed to the consistency of the Aventis Pharma study values.

The different findings of this study in comparison to the previous report [1] cannot be attributed to different batches being used since the two batches are considered to be essentially similar (i.e. differences in their dissolution profiles did not exceed 10% at any given sampling time point).

5. Conclusion

In conclusion, the dissolution method *is* a suitable technique for detecting the possible impact of formulation or manufacturing changes [2] for this particular immediate

release dosage form provided that the necessary precautions are taken. It appears to be most successful when designed as a multiple-point test using USP compliant paddle apparatus and, above all, appropriate and consistent training in the dissolution method must be given in order to obtain accurate and reproducible results.

Acknowledgements

We would like to thank the Dissolution Working Group of the International Pharmacists Federation (FIP) for their valuable comments on our study. In particular, we wish to thank Dr Vinod Shah, Rockeville and Dr Saeed Qureshi, Ottawa, for their valuable scientific input and support.

References

[1] S.A. Qureshi, I. McGilveray, Typical variability in drug dissolution testing: study with USP and FDA calibrator tablets and a marketed drug (glibenclamide) Product, Eur. J. Pharm. Sci. 7 (1999) 249–258.
[2] FDA, Guidance for Industry, Immediate Release Solid Oral Dosage Forms – Scale-up and Post-Approval Changes: Chemistry, Manufacturing and Controls, In Vitro Dissolution Testing and In Vivo Bioequivalence Documentation (1995).
[3] USP 24. The United States Pharmacopeial Convention, Inc. << 711 >> Dissolution, 1999 pages 1941-1944.