ELSEVIER

Journal of Pharmaceutical and Biomedical Analysis
17 (1998) 419–425

JOURNAL OF
PHARMACEUTICAL
AND BIOMEDICAL
ANALYSIS

Reproducibility of ^{1}H-NMR integrals: a collaborative study[1]

M. Bauer [a], A. Bertario [b], G. Boccardi [b,*], X. Fontaine [c], R. Rao [a], D. Verrier [a]

[a] *Sanofi Recherche, 195 Route d'Espagne, 31036 Toulouse Cedex, France*
[b] *Centro Ricerche Midy Sanofi, via Piranesi 38, 20137 Milano, Italy*
[c] *Sanofi Recherche, 371 Rue du Professeur J. Blayac, 34184 Montpellier Cedex, France*

Received 4 August 1997; accepted 3 September 1997

Abstract

The quantitative use of NMR spectroscopy was investigated by a reproducibility study of ^{1}H-NMR integrals involving five laboratories. A significant laboratory effect was found confirming the difficulty to obtain very precise data by integration of complex signals. The reproducibility of any NMR assay measurement, which requires a high precision should be validated by an interlaboratory study. © 1998 Elsevier Science B.V.

Keywords: Quantitative NMR; Reproducibility; Interlaboratory study

1. Introduction

Even though NMR spectroscopy is mainly a technique for structural analysis, it has many applications as a quantitative analytical tool. The most attractive feature of this spectroscopy is that under appropriate acquisition conditions the molar response factor is exactly the same for all resonating nuclei in a given solution. This allows to perform quantitative analysis without the analytical standard of the analyte, using simple and well characterised primary standard as in volumetric analysis. A quantitative determination by NMR is normally obtained from the ratio between the integral of a signal of the test compound and the integral of a signal of the primary standard. Furthermore, under quantitative conditions the matching of integrals with the structural formula is an index of purity. A number of papers describe NMR methods for the quantitative determination of active substances in pharmaceutical forms. Even recent papers describe analytical NMR methods performed using continuous wave spectrometers [1–3], but these equipments are obsolete in comparison with Fourier transform spectrometers. In many cases values of precision were reported for NMR analyses, [1–6], very similar to those of chromatographic methods. It is worth noting that very often either the experimental conditions and the optimisation procedure are not very detailed, leading the reader to think that these experiments are quite simple. It is also worth noting that in some cases the claimed 'reproducibility R.S.D.' is in fact repeatability R.S.D. and that internal standards better than

* Corresponding author.
[1] Dedicated to the memory of Renée Rao.

PII S0731-7085(97)00208-2

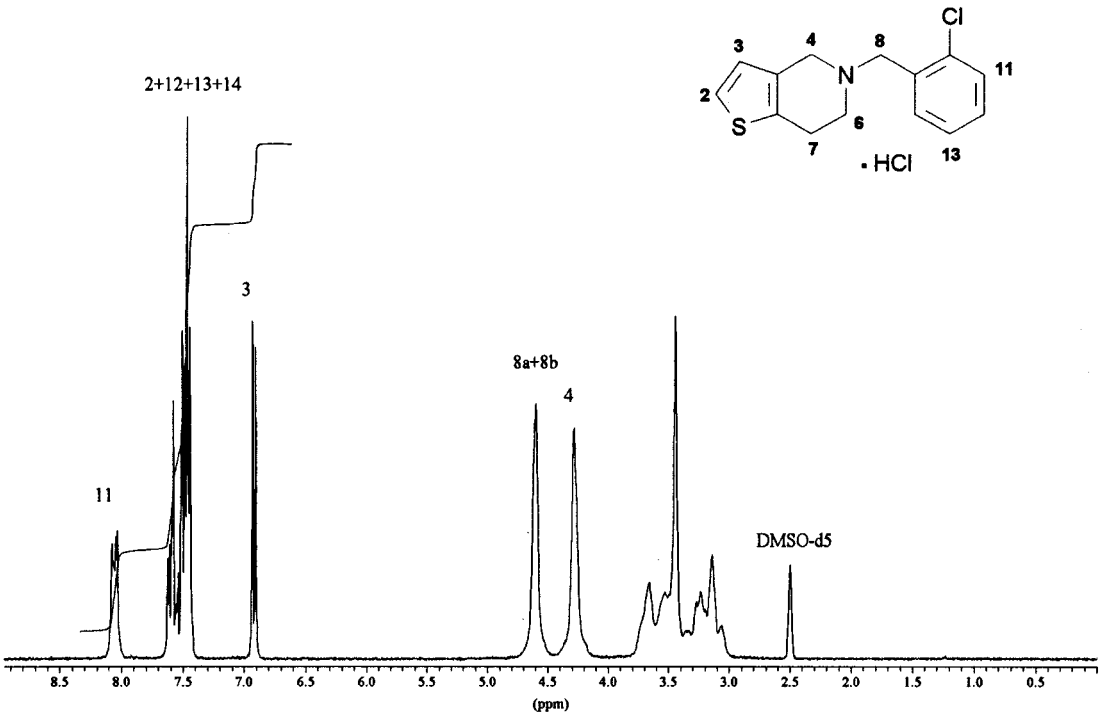

Fig. 1. Ticlopidine hydrochloride and its ^{1}H-NMR spectrum in dmso-d_6 (200 MHz).

'≥ 99' grade should be used for assessing accuracy in the range 99–101%.

In fact a number of factors make it difficult to obtain precise NMR integrals [7]. ^{1}H-NMR is the most useful quantitative technique for sensitivity and relaxation reasons, and the present discussion will deal only with this spectroscopy. The first source of bias to consider is the slow decay of the Lorentz-Cauchy curve of the NMR resonance line: the integration range must be 636-times the line width to obtain an accuracy of 99.90% [8]. This is practically impossible to achieve and for this reason the ratio of two NMR integrals can be accurate only because the errors cancel each other. ^{13}C satellite bands can also affect accuracy, and its (field dependent!) interference should be evaluated carefully.

Some processing operations in NMR (e.g. phase and baseline corrections) rely on a subjective judgement of the goodness of the shape of signals. According to the author's experience, the training of the operator in the processing of the spectra of a specific product leads to an improvement in accuracy and precision. Collaborative studies were reported on ^{1}H-decoupled deuterium NMR spectroscopy, based on height measurements [9]. Height measurement allows the avoidance of the integration problem due to the Lorentzian profile, but leads to the loss of the absolute response and is only reasonable in the case of singlets.

To the best of the author's knowledge, all papers on quantitative NMR based on integral measurement report precision in terms of repeatability, whereas one could expect that the dependence of the performance of a method to many experimental parameters and to subjective judgements will lead to reproducibility problems.

The object of the present collaborative study is to evaluate the reproducibility of a quantitative NMR method. The attention was focused only to the variability due to the NMR measurement, and not to that from sample preparation. For this reason the analytical parameter was the ratio

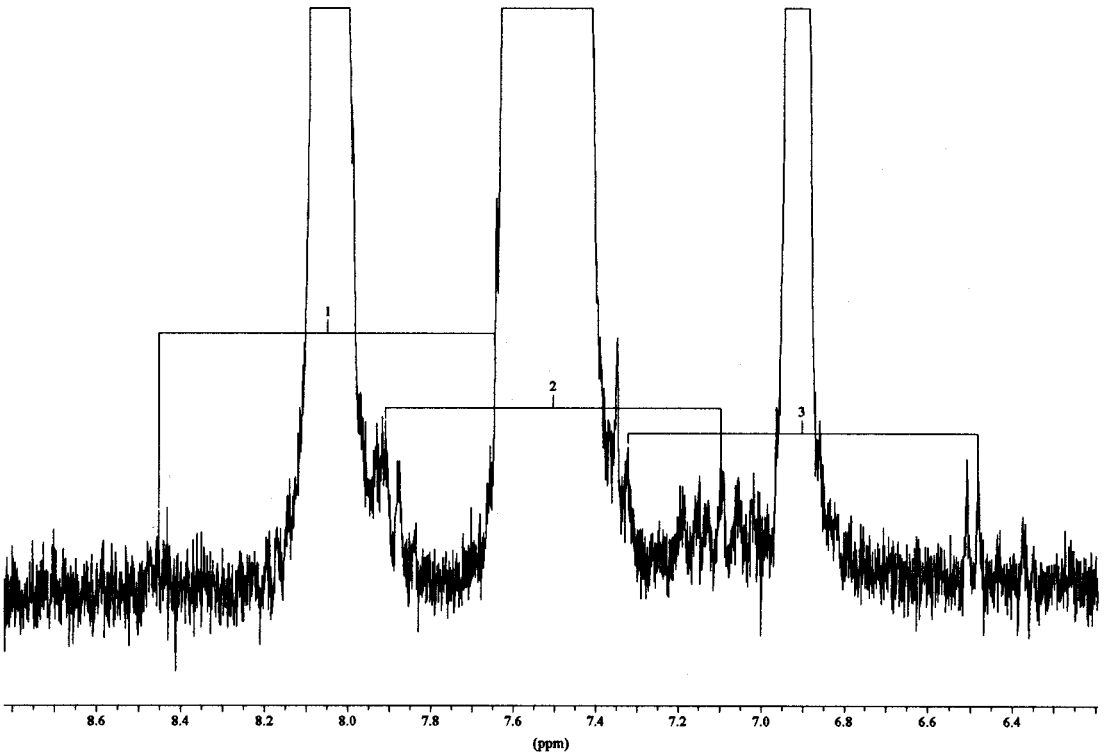

Fig. 2. Spectrum of ticlopidine hydrochloride (960 scans) showing the ^{13}C satellite bands. The splitting of each signal due to the one-bond $^{1}H-^{13}C$ coupling is shown as predicted using a standard value of 160 Hz. The position of the satellites is qualitatively confirmed by the experimental spectrum. Spinning bands were not detected.

between two signals from the same molecule. The operating field (4.7 T) and some experimental parameters (e.g. number of transients, no degassing) were chosen in order to keep the cost of the analysis within a reasonable range.

The investigated molecule, ticlopidine hydrochloride (Fig. 1) is a powerful inhibitor of ADP-induced platelet aggregation available as a very pure substance, with a simple but not trivial spectrum. The theoretical relative integrals of the aromatic signals can be easily calculated, taking into account the overcrowding of the ^{13}C satellite bands. Since the integration error can be very different considering a complex signal or a simple sharp one, two integrals ratios were measured, R_1 being the ratio between two sharp single proton signals and R_2 being the ratio between a sharp single proton signal and a broad signal arising from a group of 4-protons:

$$R_1 = 100 \times \frac{\mathrm{H}(11)}{\mathrm{H}(3)}$$

and

$$R_2 = 400 \times \frac{\mathrm{H}(11)}{\mathrm{H}(2+12+13+14)}$$

2. Experimental

2.1. Analytical method

The participating laboratories were asked to carry out the following protocol. The spectrome-

Table 1
Results of the collaborative study: descriptive statistics on R_1 data

Laboratory		R_1, manual integration			R_1, automatic integration		
		Sample 1	Sample 2	Sample 3	Sample 1	Sample 2	Sample 3
1	Mean	101.12	99.07	101.07	102.53	99.50	97.97
	R.S.D. %	2.19	0.73	1.06	0.57	0.15	0.29
2	Mean	100.70	99.77	99.15	99.12	99.56	98.75
	R.S.D. %	0.48	0.56	0.37	1.03	0.68	0.25
3	Mean	101.19	104.06	102.97	97.68	98.70	103.60
	R.S.D. %	0.39	0.41	0.83	0.55	1.84	0.17
4	Mean	102.13	101.16	101.22	100.98	99.54	101.87
	R.S.D. %	0.66	0.41	0.36	1.52	0.43	0.36
5	Mean	100.39	100.35	97.76	100.74	101.42	100.73
	R.S.D. %	3.9	1.23	2.00	1.67	1.21	2.76

ters were checked for conformity to the manufacturer's specifications for chloroform lineshape and spinning sidebands and the transmitter 90° pulses were calibrated. Ticlopidine hydrochloride (20 mg) was dissolved in deuterated dimethylsulfoxide (1 ml). The spectrum was recorded on a spectrometer equipped with DISNMR software. The sample was rotated at 20 Hz. Shims' z and z^2 were carefully shimmed. Free induction decay (FID) was acquired under the following conditions (DISNMR acronyms are indicated): pulse width (PW) 30°; sweep width (SW) 4000 Hz; carrier frequency (O1) set at 7.5 ppm; number of transients (NS) 96; time domain (TD) 32 K; repetition time (RD + AQ) 20 s; room temperature not less than 20°C. The spectrum was processed without apodisation and phased manually (zero order on the peak H-11 at 8.1 ppm and first order on the peak of the pentadeuterated dimethylsulfoxide). The baseline was corrected by an automatic routine (ABS). For the manual integration procedure, the peaks were manually integrated between 7.75 and 8.25 ppm (H-11), between 7.21 and 7.75 ppm (H-2 + H-12 + H-13 + H-14) and between 6.78 and 7.04 ppm. The manual integration consists in digitally integrating each peak between the limits manually entered with a cursor. For the automatic integration procedure, the AZF routine was used, with AZFE = 100. AZF is the automatic integration routine of DISNMR software; its only critical parameters, AZFE, represents the number of points added to the right and to the left of the integration domain in order to include the onset of the peak.

2.2. Collaborative study

Five laboratories participated in the study. All the laboratories were equipped with a Bruker AC 200 spectrometer, working at 200 MHz ^{1}H. A ticlopidine hydrochloride sample from a single batch was distributed to the laboratories. Each laboratory was asked to record three FIDs on separate ticlopidine hydrochloride solutions, each on a different day. Each FID was processed three times with manual integration and three times with automatic integration, each time reprocessing phase and baseline.

2.3. Collaborators

A. Bertario (Research Cantre Midy Sanofi, Milano, leading laboratory), X. Fontaine (Sanofi Recherche, Bruxelles), C. Bonnel (Centre Applications Levallois, Levallois), C. Navette (Sanofi Recherche, Toulouse), G. Comminges (Laboratoire Chimie de Coordination, Toulouse).

2.4. Statistical analysis

Results were analysed according to ISO 5725. Relative standard deviations for repeatability (R.S.D. (r)), intermediate precision (R.S.D. (int)) and reproducibility (R.S.D. (R)) were calculated from an appropriate analysis of variance using a mixed model.

Table 2
Results of the collaborative study: descriptive statistics on R_2 data

Laboratory		R_2, manual integration			R_2, automatic integration		
		Sample 1	Sample 2	Sample 3	Sample 1	Sample 2	Sample 3
1	Mean	99.95	98.00	99.51	100.07	99.30[a]	98.10
	R.S.D. %	1.71	0.71	0.48	0.21	—	0.10
2	Mean	99.25	98.96	98.58	98.68	98.95	98.82
	R.S.D. %	0.32	0.44	0.28	0.57	0.43	0.21
3	Mean	98.11	99.54	97.92	100.07	100.62	100.76
	R.S.D. %	0.32	0.36	0.51	0.025	0.44	0.204
4	Mean	100.54	100.19	100.00	100.20	99.83	100.71
	R.S.D. %	0.41	0.28	0.23	1.29	0.21	0.32
5	Mean	97.55	98.16	95.72	98.45	99.39	98.46
	R.S.D. %	2.93	1.15	2.34	0.65	1.30	2.65

[a] This data was detected as an outlier ($P<0.01$).

3. Results and discussion

3.1. Method development

The method was developed in the leading laboratory. T_1 approximate measurements were performed on undegassed samples in order to set a correct repetition time. H-3 was found to be the slower relaxing proton ($T_1 = 2$ s). The ruggedness was tested using a two level incomplete factorial design on seven acquisition or processing parameters, with three replicates for each experiment. The following factors were considered: the use of a relaxation agent ($Cr(acac)_3$) to shorten and to level T_1 values (0 or 6.1 mmol mol^{-1}); the control of temperature (no control or 310 K); the number of transients to improve the signal to noise ratio (96 or 960); the carrier frequency (7.5 or 5.5 ppm); the memory size of the FID (TD = 32 K or TD = 64 K); zero-filling, cited as a critical point in quantitative NMR [10] (no or single order zero-filling); the use of an exponential multiplication to improve the signal to noise ratio (line broadening of 0 or 1 Hz). The number of transients was the sole factor showing a significant effect ($P = 0.008$). The use of relaxation agents and of exponential multiplication of the FID were excluded, because of their detrimental effect on the linewidth of the relevant signals. The other parameters were chosen in order to perform the most simple experiment. The number of transients was set to a value compatible with a reasonable analysis time.

3.2. The 'true' value of R_1 and R_2

R_1 and R_2 are defined so that their value is approximately 100. In order to calculate the exact value of these ratios the ^{13}C satellite bands have to be taken into account, given by the presence of 1.1% of the NMR active nuclide ^{13}C. Fig. 2 shows the spectrum of the relevant region, acquired with 960 scans to enhance sensitivity and expanded in the ordinate scale to show the satellite bands. A fairly 'true value' of R_2 at 200 MHz can easily be obtained by the following considerations. Using the standard value of 160 Hz for the $^1J_{H-C}$, the ^{13}C satellite bands of H(2 + 12 + 13 + 14) are fused with the H-11 peak and the satellite bands of H-11 and H-3 overlap with the H(2 + 12 + 13 + 14) peak (Fig. 2). Hence:

$$R_2 = \frac{100 + 4 \cdot 0.55}{400 + 2 \cdot 0.55} \times 400 = 101.9$$

The hump of the narrowest of the considered peaks, H-3, is 24 Hz: this allows to neglect $^2J_{H-C}$ (1–4 Hz from the literature) and $^3J_{H-C}$ (7–10 Hz).

The interference on H-3 is less clear because of the closeness of the satellites of H(2 + 12 + 13 + 14) and the signal of H-3, so R_1 can only be evaluated between 101 and 102%. The spectra obtained by the leading laboratory did not show

Table 3
Statistical analysis

Measured ratio	Integration method	R.S.D. (r)[a]	R.S.D. (int)[a]	R.S.D. (R)[a]	*P* (sample)*	*P* (laboratory)**
R_1	Manual	1.40	1.61	1.97	0.07	0.036
R_1	Auto	1.17	2.08	1.93	<0.001	0.76
R_2	Manual	1.15	1.27	1.63	0.13	0.016
R_2	Auto	0.89	0.92	1.19	0.32	0.010

[a] See Section 2.4.
* Significance of the sample factor.
** Significance of the laboratory factor.

any spinning band, even with 960 scans. Furthermore spinning bands were at least partially included in the integral of the corresponding peak. For these reasons spinning bands were not taken into account.

3.3. Results of the collaborative study

Descriptive statistics of the results obtained from the participants are presented in Table 1 and Table 2. In each cell the standard deviation represents the precision of the three processings of the same FID, performed with the given method, manual or automatic. The three samples each laboratory analysed were obtained from different solutions and different tubes and measured on different days, and therefore the sample effect represents the influence of the sum of these parameters. No problems were reported from participants. Detection of outliers was performed using the Dixon test. One point from laboratory 1 and sample 2 was detected as an outlier ($P < 0.01$) and removed. A qualitative inspection of the data could suggest that laboratory 5 gave more dispersed results for all the signals. The statistical analysis showed no significant difference of the mean of the laboratory 5, even though its S.D. was significantly larger than the value of the other laboratories. Laboratory 5 submitted calibration spectra ($CHCl_3$ lineshape and ethylbenzene sensitivity) and an example of integrated spectrum: no abnormality was found. Laboratory 5 has an older spectrometer with a long 90° pulse (about 13 μs, whereas the other spectrometers have 6 μs 90° pulses). This fact, however, should not be significant, taking into account the narrow range of frequency in which all the relevant signals are located. Therefore data from laboratory 5 were not removed.

Table 3 summarises the statistical results. R.S.D. (r) were generally about 1%, whereas R.S.D. (R) varied from 1.2 to 2%.

Three of the four considered measurements (R_1/R_2/manual/auto) show a significant laboratory effect and no acquisition effect. This means that repeatability is quite good but systematic error may be significant. Only the R_1 ratio from automatic integration shows a strong sample effect, which hides a possible laboratory effect. R_1 is the ratio between the integrals of two sharp signals: in this case the manual integration can probably be better adapted to small linewidth or lineshape differences due to the acquisition from different tubes and different days.

For each ratio, the correlation between the manual integration and the automatic integration was checked. No significant correlation exists both for R_1 ($r = 0.14$, slope $= 0.15/0.30$) and for R_2 ($r = 0.41$, slope $= 0.29/0.18$). If the sample error was relevant, a correlation between the two elaborations of the same sample should be found.

4. Conclusions.

Some of the participating laboratories reported within sample precision less than 0.5%, and these individual data could suggest a great confidence in the quantitative use of proton NMR. In fact, the laboratory effect was found very significant in all measurements. The poor matching of the general mean of R_1 and R_2 with the theoretically expected

values confirms the presence of bias. Surprisingly, R_1 measurement also shows a quite significant sample effect, indicating that true replicates are to be preferred to simple reprocessing of the same FID.

The collaborative study has been performed using spectrometers of the same field and from the same manufacturer. The use of less homogeneous instrumentation would presumably give worse reproducibility.

The study seems to confirm the difficulty to obtain, in terms of reproducibility, very precise data by integration of complex signals. Analytical NMR is probably less useful in the case of an assay than in impurity determinations, where a high precision is not so mandatory and where the absolute response factor and the reproducibility of chemical shifts (in contrast to the variability of relative retention times in chromatography) are very attractive features [4,11]

Several items could be identified as possible ways to improve the accuracy and precision of NMR integrals: the use of larger field (with a huge increase in costs!), the use of more modern electronics, better integration algorithms, a more sophisticated experimental protocol (e.g. degassing of solutions) or better training of the laboratory staff. In any case thorough analytical development and validation with a complete description of the instrument requirements are mandatory for NMR methods. As the systematic errors are the more critical problem, the reproducibility should be confirmed by a collaborative study.

Acknowledgements

The authors are very grateful to Claudine Bonnel, Gérard Commenges, Clement van Meerbeeck, Christelle Navette and Luigi Uberti for participating to the collaborative study.

References

[1] G.M. Hanna, C.A. Lau-Cam, Drug Dev. Ind. Pharm. 17 (1991) 975–984.
[2] G.M. Hanna, C.A. Lau-Cam, J. Pharm. Biomed. Anal. 11 (1993) 855–859.
[3] G.M. Hanna, C.A. Lau-Cam, J. Assoc. Off. Anal. Chem. Int. 79 (1996) 833–838.
[4] P.P. Lankhorst, M.M. Poot, M.P.A. de Lange, Pharmacopeial Forum 22 (1996) 2414–2422.
[5] L.K. Revelle, D. André d'Avignon, J.C. Reepmeyer, R.C. Zerfing, J. Assoc. Off. Anal. Chem. Int. 78 (1995) 353–358.
[6] G. Fardella, P. Barbetti, I. Chiappini, G. Grandolini, Int. J. Pharm. 121 (1995) 123–127.
[7] A.E. Derome, Modern NMR Techniques for Chemistry Research, Pergamon Press, Oxford, 1987, pp. 168–170.
[8] C.H. Sotak, C.L. Dumoulin, G.C. Levy, Topics in Carbon-13 NMR Spectroscopy 4 (1984) 91–121.
[9] G.G. Martin, R. Wood, G.J. Martin, J. Assoc. Off. Anal. Chem. Int. 79 (1996) 917–928.
[10] K. McLeod, M.B. Comisarow, J. Magn. Res. 84 (1989) 490–500.
[11] B. Lindgren, J.R. Martin, Pharmeuropa 5 (1993) 51–54.