doi: 10.1255/tosf.66

Estimating total sampling error for near infrared spectroscopic analysis of pharmaceutical blends—theory of sampling to the rescue

A. Roman-Ospino[a], C. Ortega-Zuñiga[a], A. Sanchez-Paternina[a], S. Ortiz[a], K. Esbensen[b] and R.J. Romañach[a]
[a]andres.roman@upr.edu, [a]carlos.ortega4@upr.edu, [a]adriluz.sanchez@upr.edu, [a]stephanie.ortiz10@upr.edu, [b]ke@geus.dk, [a]rodolfoj.romanach@upr.edu

A replication experiment was performed to validate a stream sampling method for a pharmaceutical powder blend. A 1.5 kg powder blend was prepared and an in-house developed feeder was used to divide into six sub-samples of approximately 250 g. Each 250 g sub-sample (1/6 total blender lot volume) was deposited along a rig of 3 meter length. A validated near infrared (NIR) spectroscopic method was used to determine the drug concentration as the powder deposited in the rig moved at a linear velocity of 10 mm/sec. The depth of penetration of the NIR radiation was 1.2 mm and the sample volume analysed was approximately 180 mg. The MPE (minimum practical error) obtained with the system was 0.04% w/w acetaminophen (APAP), which was considered excellent for the system. The replicate analysis of the powder deposition provided 390 measurements of drug concentration, with a mean APAP concentration of 14.93% (w/w) and a relative standard deviation (RSD) of 5.20%. Replicate measurements (n = 650) of the powder deposited along a single rig of 3 m length ×10 provided an RSD of 2.23%, attributable to deposition (outflow) heterogeneity. Finally, static replicate analysis of the measurement error alone amounted to an RSD of 0.14%. The embedded replicate experiments elucidated all sources of variation in a sampling system for pharmaceutical powder blends, and proved reliable and highly sensitive in identifying areas of non-acceptable residual heterogeneity (dead zones).

Background

The analysis of drug concentration in pharmaceutical blends is mostly done through grab sampling where a sampling spear (called sampling thief in the pharmaceutical industry) is frequently inserted into a blender to extract 6–10 samples.[1,2] The extracted material is then taken to a laboratory where the drug concentration of the powder blend is determined. The sample thief is used to extract powder mixture from specific locations and transects through the blender volume, which based on previous studies, have shown a greater likelihood to represent "dead spots" (areas of residual incomplete mixing).[3] Thus, all the components of the blender volume, the lot, do not have the same probability of being extracted for analysis. This is a structural fault of the sampling system. If the areas of incomplete mixing are not those selected with this fixed location approach the sampling approach will fail to do what it is supposed to do and volumes with larger residual heterogeneities will go undetected. This is the exact opposite of the objective of end-of-mixing sampling and analysis.[1–3]

These flawed approaches are currently being complemented by non-destructive near infrared (NIR) spectroscopic methods developed to analyse the drug concentration within the blender (in-line), or at-line/off-line. The non-destructive spectroscopic methods are so far usually interfaced at a single location within the blending vessel (or interacting through a window in the vessel wall).[4] If powder moves in and out of the sampling interface there is a greater likelihood that larger parts of the lot will be analysed than with a powder thief, depending on the specific combination of analysis volume w.r.t. material through-flow in relation to the full vessel volume. But such solutions, despite having a clear potential of being significantly better than thief sampling, are by no means a complete solution for the desired blender material characterisation based on the full blender volume. To the degree that this is not achieved (yet), the present verification approaches cannot be said to be comprehensive.

However the powder mixture can alternatively be sampled after it leaves the blender, either using a physical sampling approach or by invoking the rapid, and more efficient NIR spectroscopic method for analysis.[1,2] In this approach the powder flows down a chute, or is ducted via a mini-conveyor belt, from which a NIR spectrometer can obtain spectra of the mixture. This is a Process Analytical Technology (PAT) approach, of great potential and considerable proved merit.[5–7] Based on a chemometrics multivariate calibration model it is possible to predict the drug concentration in the NIR-beam analytical volume.[8] This stream sampling approach has been followed experimentally in a limited number of pilot studies.[9–11]

We here report on pioneering laboratory validation of a PAT stream sampling approach where the active drug concentration is determined by NIR spectroscopy. Previous studies have involved thorough validations of NIR analytical methods obtaining accurate estimates of the Total Analytical Error (TAE), but have not addressed the accompanying sampling errors.[4] This study describes the result of a first systematic Replication Experiment approach[12] in a realistic laboratory setting. The systematic replication experiments represent a new approach to the analysis of blends and to estimating the *effective* sampling and measurement uncertainty within pharma.[12–14] We are aware of only two other forays within pharma, in which TOS is also an important element, both focusing on product analysis uncertainty[17,18]

Experimental

Materials: The blends were prepared from lactose monohydrate Granulac (Meggle Pharma), microcrystalline cellulose Vivapur 102 (JRS Pharma) and semi-fine acetaminophen (APAP) from Mallinckrodt Inc. (Raleigh, NC). The lactose monohydrate was passed through a U.S. Standard Sieve 60 (250 µm opening) before mixing.

Calibration Model: An experimental design was followed to minimize correlation between components and obtain a robust

Table 1. Composition of calibration and test set blends for NIR calibration model.

Blend	1	2	3	4	5	6	7	8	Test set
APAP (% w/w)	7.50	7.50	7.50	14.00	15.00	16.25	22.50	22.50	15.0
MCC (% w/w)	30.00	90.00	60.00	63.50	30.00	83.75	77.50	30.00	66.67
LAC (% w/w)	62.50	2.50	32.50	22.50	55.00	0.00	0.00	47.50	18.33

calibration model. Three component blends were prepared (correlation between majority components is unavoidable, and this process reduces the other two). The experimental design software MODDE 8.0.0.0 Umetrics (Umeå, Sweden) was used. Settings were 14 runs, objective: screening, in a D-optimal design linear model. The concentration range was 50% above and below the 15.0% w/w APAP target concentration, resulting in a calibration set spanning 7.5–22.5% w/w. Table 1 shows the concentrations of the eight calibration blends prepared.

Preparation of Test Set Blend: A 1.5 kg blend with an APAP concentration of 15.0% (w/w) was prepared as shown in Table 1. This blend was used for the entire replicate study.

Description of Fourier Transform Near Infrared (FT-NIR) system and software to develop the calibration model: A Bruker Optics (Billerica, MA) Matrix FT-NIR spectrometer was used to obtain spectra. Calibration and test set spectra were obtained at a spectral resolution of 8 cm^{-1} and a total of 32 scans were averaged. Each spectrum (average of 32 scans) requires about 4.4 seconds. All spectra were obtained as the powder moved at a linear velocity of 10 mm/s, except for the static repeatability test (see below). Under these conditions, each spectrum can be estimated to represent approximately 180 mg of powder mixture as shown in Figure 1.[11] Calibration models were developed in SIMCA 13.0 Umetrics (Umeå, Sweden), partial least squares algorithm. NIR spectra were pre-treated with a standard normal variate transformation and a first derivative based on 17 points. The chemometric model was performed on the 9100–5000 cm^{-1} NIR spectral range. The performance of the calibration model was evaluated with independent test blends, aka test set validation.[13–15]

A sampling system was designed to deposit blends over the conveyor belt for simulating a 1-dim industrial blender outflow sampling/analysis system. Each powder mixture (both calibration – and validation blends) was deposited in a 3 m long, 4 cm wide and 3 cm deep rig by the use of an in-house developed screw feeder, as shown in Figure 2. The feeder was operated so as to provide a thick powder bed on the rig. FT-NIR spectra were obtained along the entire 3 m length rig corresponded to approximately 250 g of the 1.5 kg lot powder mixture. The powder surface was left uneven and no attempt was made to obtain a flat surface of powder in the recipient, aiming to produce a highly realistic industrial situation.

Figure 2 shows a photograph of the system for Replication Experiment studies (six successive rig depositions, 10 times to-and-fro over just one outflow. The Matrix FT-NIR spectrometer is situated at a height ~10 cm to obtain spectra as the rig moves at 10 mm/sec. The replicate experiment was first conducted by performing 6 outflow depositions each of approximately 250 g along the 3 m rig. This setup yielded approximately 65 spectra per outflow stream. The APAP drug concentration was predicted for each spectrum using the validated FT-NIR calibration model (multivariate calibration prediction).[8]

The second replication experiment consisted of moving one of the full length outflow deposition over the conveyor belt to and fro 10 times, obtaining spectra from one end to the other. The final part consisted of a repeatability study, where six consecutive spectra were obtained at one fixed location without moving the powder mixture or the spectrometer. This repeatability study was itself performed a total of 6 times. All replication experiment results are shown in Table 2.

Results and Discussion

The above replication experiment was performed to validate a specific PAT sampling/analysis facility for a realistic 1.5 kg powder blend

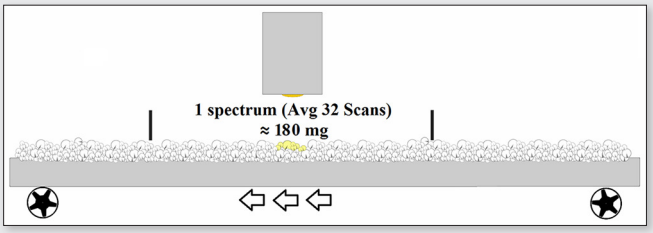

Figure 1. Schematic rig illustration of PAT sampling by a NIR spectrometer along conveyor belt material stream. Observe how the NIR beam only interacts with the top layers of the material stream, giving rise to structural IDE/IME contributions to the total measurement system error in the vertical direction[depth of penetration is 1.2 mm]. The estimated analytical mass is about 180 mg.

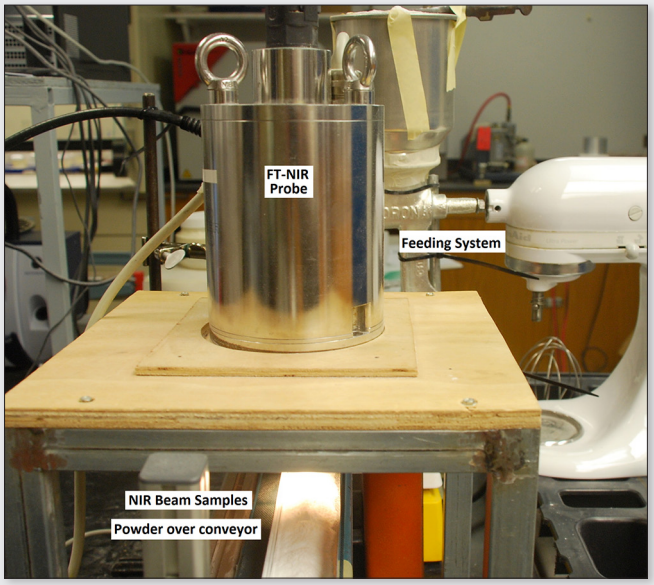

Figure 2. Conveyor belt assembly (total length 3 m) with FT-NIR spectrometer positioned at a height of 10 cm and powder feeding system (background). Note that the NIR beam covers the entire width of the conveyor belt, suppressing a potential IDE contribution to the total measurement system error in the cross-stream direction.

Table 2. Results of Replication Experiments.

	Deposition[a] (n = 6)	Replicates[b] of Single Deposition (n = 10)	Repeatability Study[c] (n = 6)
Ave.	14.93	15.21	15.78
Std. Dev.	0.78	0.34	0.14
RSD (%)	5.20	2.23	1.3
Spectra (#)	390	647	36

[a]Deposition = one deposition length (3 m)

[b]Spectra were collected 10 times along the complete length of the rig for a total of 647 spectra

[c]Static NIR beam footprint on unmoving rig; six replicated NIR spectra acquisition

prepared with a 15.0% (w/w) APAP concentration. The lot in question was the full 1.5 kg prepared blend, from which the six repetitions of a full length (3 m) 250 g rig experiment could be performed. Each 250 gram sub-sample (1/6 total blender lot volume) allowed about 65 analyses (based on NIR spectra) to be made along the rig length, Figures 1 and 2. This enables evaluation of both full and partial blender outflow analysis performance.

Table 2 shows that the grand average concentration predicted by the NIR calibration model was 14.93% (w/w), based on all 390 analyses performed for the lot, i.e. a situation in which the *entire* outflow material stream has been analysed. The relative standard deviation of this complete lot volume results was 5.20%. These results must be considered excellent as these involve the maximal combined variation effects stemming from i) the outflow deposition (flow segregation), ii) residual blend heterogeneity and iii) TAE of the PAT NIR analytical method. The relative standard deviation is termed the relative sampling variability (RSV) for the replication experiment approach.[12]

Table 2 also shows the results from replicate analysis of a single deposition (i.e. a single conveyor belt pass but repeated to-fro 10 times). This experiment addresses the specific blend heterogeneity in one 1/6 total lot stream only (including the attendant TAE). As expected, this RSV variation is significantly lower, 2.23%. The average drug concentration is here 15.21% (w/w). Thus, the average concentration is different from that when the entire lot was analysed. There is thus a difference of +0.28% APAP, due to that only 1/6 part of the lot is being analysed.

The static analytical repeatability studies results (the NIR beam was focused on a single unmoving area of the powder blend and six consecutive spectra were acquired) are also shown in Table 2. The relative standard deviation in the repeatability study is approximately 0.2%, attesting to TAE only.

Variability larger than this analytical baseline represents i) residual blend heterogeneity (imperfect mixing), ii) specific outflow variability ("deposition" above) as well as iii) possible process sampling errors for the PAT sensor system. The variance of this analytical repeatability study $(0.2)^2$ may be subtracted from the square of the standard deviation of the replicate analysis of the single deposition to obtain a measure of the blend heterogeneity. The replicates of single deposition show a standard deviation of 0.34, and after subtracting the measurement repeatability the blend heterogeneity is

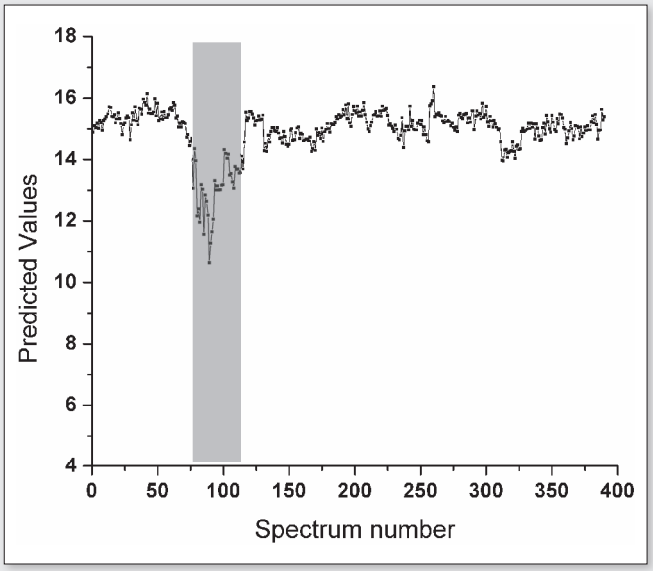

Figure 3. Prediction of drug concentration for the 390 individual analysis of the complete lot (six ×3 m rig lengths).

0.31. These values could be used as baseline level to improvement the sampling and measurement systems.

Figure 3 shows the plot of the drug concentration values throughout the entire run, revealing a significant drop in drug concentration from approximately spectrum #78 to 116. This simple plot is crucial in showing that a certain part of the blend was responsible for the overwhelming part the heterogeneity observed—a dead spot. The drug concentration from spectrum #81 to 100 averaged 12.5% instead of the 15.0% target level. Thus, the stream sampling approach was very capable to identify incomplete mixing process without the use of sampling spear.

The main feature of the replication experiment studies concerns the possibility to apply a variographic characterisation of the outflow stream. The variogram function V(j) was determined based on the drug concentration values predicted by the NIR calibration model. A lag of 1 was based on consecutive predictions of drug concentration, each concentration corresponding to approximately 180 mg as shown in in Figure 1. The maximum lag shown in the variogram is 190, since the total number of drug concentration predictions was ~390. From Figure 4 it is obvious that the total PAT measurement system error is very small (nugget effect) compared to the level of drug content variance (sill) along the full 3m outflow stream. The range is approx. 30–36, i.e. the distance within each predicted drug concentration is increasingly auto-correlated for smaller lags than this.

This run also allows a simulation of the variographic outflow approach for NOC (normal operation conditions), by excluding the samples in the interval #78–116 (resulting in a seamless outflow only characterised by the NOC residual heterogeneity). A renewed variogram for this data series is presented in Figure 4 (right), in which can be seen that the nugget effect is identical, while there is a very notable reduction of the sill level – both features as expected. Renewed estimation of the $RSV_{1\text{-}dim}$ results in 2.6%. This run is fully realistic w.r.t. to its industrial counterpart to the degree that the blender used is reasonably up-scalable; all other system elements

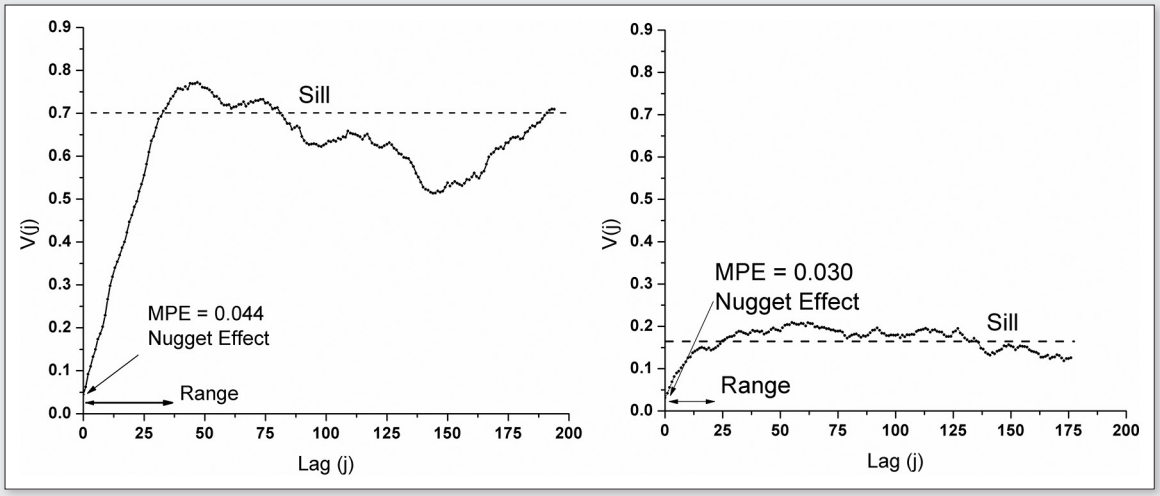

Figure 4. Left: Variogram based on the total of 390 individual analyses of the complete lot (six ×3 m rig lengths). The range is ~30-36; nugget effect = 0.04; sill = 0.7. The total measurement system uncertainty, $RSV_{1\text{-}dim}$, is therefore ~5.2% (rel).[12] Right: Same variogram excluding shaded area in Figure 3.

would be identical: outflow facility, NIR spectrometer, chemometric prediction model.

A recently withdrawn draft guidance which describes the analysis of powder blends by thief sampling requires the analysis of drug concentration for at least 10 blends from a tumble blender with: 1) a relative standard deviation ≤5%, and 2) all individual results within 10.0 percent (relative) of the mean drug concentration.[16] The 390 determinations of drug concentration display a RSD of 5.20% slightly exceeding the first requirement and did not meet the second requirement due to the dead spot drop in concentration shown in Figure 3. Thus, the outflow stream sampling system is eminently capable of finding areas of heterogeneity in the entire blend lot. If the blending process were improved by eliminating the sudden drug concentration drop shown in Figure 3, then the RSD in drug concentration reduces to approximately 2.6% and all values are now within 10% of the mean drug concentration stipulation.

To the degree that a complete, up-scalable measurement system can be established in the laboratory, the present approach will be able to guide rational product development, to some considerable degree without pilot—or full scale plant demonstration—until the manufacturing process has been brought into complete statistical control in the laboratory.

The value of an outflow variographic facility has been demonstrated and its merits exemplified. This is the first time a TOS-based approach (variographic and replication experiment) for the characterisation of a pharmaceutical manufacturing process has been applied with illustrative and highly satisfactory results.

Acknowledgements

This collaboration has been possible thanks to the support of the National Science Foundation (ERC research grant EEC-054085).

References

1. R. J. Romañach and K. H. Esbensen, “Sampling in pharmaceutical manufacturing - Many opportunities to improve today’s practice through the Theory of Sampling (TOS).”, *TOS Forum.* **4,** 5-9 (2015).
2. K. H. Esbensen and R. J. Romañach, “Proper sampling, total measurement uncertainty, variographic analysis & fit-for-purpose acceptance levels for pharmaceutical mixing monitoring”, in *Proceedings of the 7th International Conference on Sampling and Blending*, *TOS forum* **Issue 5,** 25–30 (2015). doi: 10.1255/tosf.68
3. G. Boehm, J. Clark, J. Dietrick, L. Foust, T. Garcia, M. Gavini, L. Gelber, J.-M. Geoffroy, J. Hoblitzell, P. Jimenez, G. Mergen, F. Muzzio, J. Planchard, J. Prescott, J. Timmermans and N. Takiar, “The Use of Stratified Sampling of Blend and Dosage Units to Demonstrate Adequacy of Mix for Powder Blends1”, *PDA Journal of Pharmaceutical Science and Technology.* **57,** 64-74 (2003).
4. C. V. Liew, A. D. Karande and P. W. S. Heng, “In-line quantification of drug and excipients in cohesive powder blends by near infrared spectroscopy”, *International Journal of Pharmaceutics.* **386,** 138-148 (2010). http://dx.doi.org/10.1016/j.ijpharm.2009.11.011
5. K. A. Bakeev, *Process Analytical Technology: Spectroscopic Tools and Implementation Strategies for the Chemical and Pharmaceutical Industries*, Second. Wiley, (2010)
6. K. H. Esbensen and P. Paasch-Mortensen, in *Process Analytical Technology* (John Wiley & Sons, Ltd, 2010) 37-80
7. K. H. Esbensen and L. P. Julius, “Representative Sampling, Data Quality, Validation - A Necessary Trinity in Chemometrics”, *Comprehensive Chemometrics: Chemical and Biochemical Data Analysis, Vols 1-4.* C1-C20 (2009).
8. H. Martens and T. Naes, *Multivariate Calibration*, Wiley, (1992)
9. M. Popo, S. Romero-Torres, C. Conde and R. J. Romanach, “Blend uniformity analysis using stream sampling and near infrared spectroscopy”, *AAPS PharmSciTech.* **3,** E24 (2002). 10.1208/pt030324
10. A. U. Vanarase, M. Alcalà, J. I. Jerez Rozo, F. J. Muzzio and R. J. Romañach, “Real-time monitoring of drug concentration in a continuous powder mixing process using NIR spectroscopy”, *Chemical Engineering Science.* **65,** 5728-5733 (2010). http://dx.doi.org/10.1016/j.ces.2010.01.036
11. Y. Colón, M. Florian, D. Acevedo, R. Méndez and R. Romañach, “Near Infrared Method Development for a Continuous Manufacturing Blending Process”, *Journal of Pharmaceutical Innovation.* **9,** 291-301 (2014). 10.1007/s12247-014-9194-1
12. DS 3077, Danish Standards Foundation, (2013)

13. K. Esbensen, P. Geladi and A. Larsen, “The Replication Myth 1”, *NIR news.* **24,** 17-20 (2013).
14. K. Esbensen, P. Geladi and A. Larsen, “The Replication Myth 2: Quantifying empirical sampling plus analysis variability”, *NIR news.* **24,** 15-19 (2013).
15. K. H. Esbensen and P. Geladi, “Principles of Proper Validation: use and abuse of re-sampling for validation”, *Journal of Chemometrics.* **24,** 168-187 (2010). 10.1002/cem.1310
16. Guidance for Industry Powder Blends and Finished Dosage Units-Stratified In-Process Dosage Unit Sampling and Assessment, (2003)
17. M. Paakkunainen, S. Matero, J. Ketolainen, M. Lahtela-Kakkonen, A. Poso and S. P. Reinikainen, “Uncertainty in dissolution test of drug release”, *Chemometrics and Intelligent Laboratory Systems*. **97,** 82-90 (2009).
18. M. Paakkunainen, J. Kohonen and S. P. Reinikainen, “Measurement uncertainty of lactase-containing tablets analyzed with FTIR”, *Journal of Pharmaceutical and Biomedical Analysis*. **88,** 513-518 (2014).