doi: 10.1255/tosf.37

Sampling in pharmaceutical manufacturing—Many opportunities to improve today's practice through the Theory of Sampling (TOS)

Rodolfo J. Romañach[a] and Kim H. Esbensen[b,c]
[a]Department of Chemistry, University of Puerto Rico, Mayagüez, Puerto Rico
[b]Geological Survey of Denmark and Greenland, Copehagen, Denmark
[c]ACABS Research Group, Aalborg University, Campus Esbjerg (AAUE), Denmark

This overview discusses sampling at different stages of pharmaceutical manufacturing—and why. The pharmaceutical industry primarily uses grab sampling. In spite of the need to know exactly the status of pharmaceutical processes and products, there are only a very few pharmaceutical applications where principles in TOS has been applied so far and representative sampling solutions are consequently often lacking. But this translates into many opportunities to improve pharmaceutical manufacturing. The authors have embarked on a large-scale programme to introduce proper sampling approaches within this important industry sector.

Introduction to pharmaceutical manufacturing

Pharmaceutical manufacturing is generically first concerned with production of a drug, usually referred to as the drug substance or *active pharmaceutical ingredient* (API), which is followed by a process where the API is mixed with excipients to manufacture the *dose units* of the drug product. API production typically includes reactions, crystallisation, solvent washes, centrifugation and drying steps. This may involve synthesis of a small molecule API, a fermentation process for an antibiotic, or the bioprocessing of large proteins. In-process characterisation of the API production constitutes the first stage of sampling in the five pharmaceutical processes shown schematically in Figure 1.

The second sampling is located at the end of the API production process. Sampling is here performed on a drug substance with a high purity. For example, many "small molecule" products are characterised by purity higher than 99% (w/w) to avoid possible secondary effects from impurities. Even though API purity is high, these are *not* "homogeneous products" and their detailed characterisation is essential in terms of both chemical and physical properties. The API must be analysed very carefully to determine its chemical properties, for their concentration of *impurities*, water content or and solvent *residues*.

Note that instead of what could appear to be a trivial case for sampling (a high purity substance) the focus is on the most difficult case: very low concentrations of impurities, necessarily with a significant heterogeneous distribution. The physical properties, i.e. particle size and crystal form are also needed. Many API have limited solubility, for which reason a reduced particle size is needed to improve the rate of dissolution. Crystal form also affects solubility as exemplified by the well-known *Ritonavir* case, involving a drug which failed key dissolution tests and for which the original crystal form could not be obtained after production of 240 lots—which caused a drug shortage of a life-saving medicine.[1] The Ritonavir case was likely caused by a low-level degradation product that served as a template for the lower solubility form. Thus, even if the drug was greater than 99% (w/w) pure, differences in low level impurities, water content or crystal form throughout the lot may very well have serious effects on the product's *performance*.

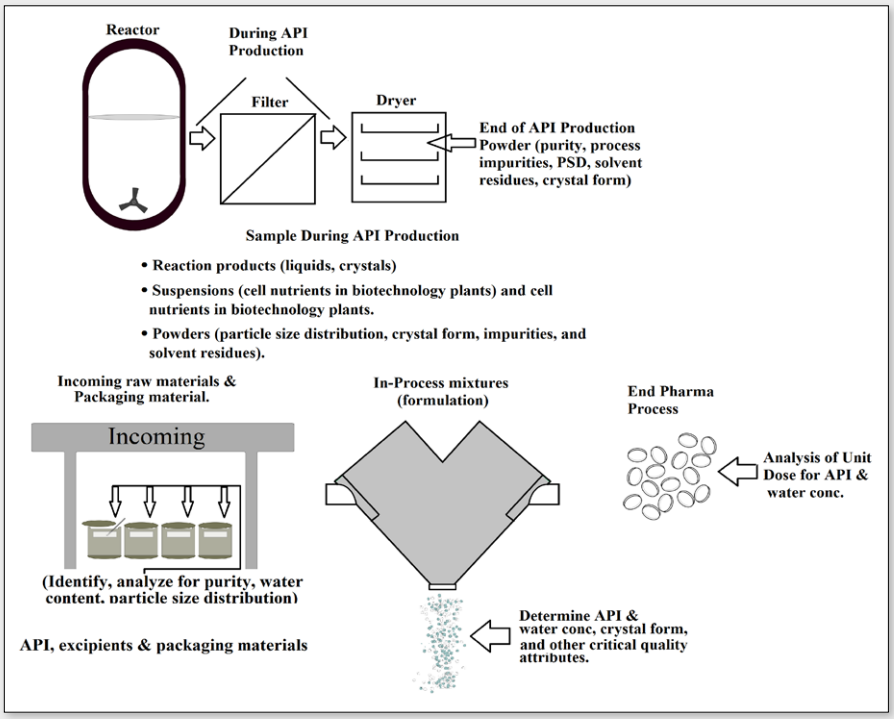

Figure 1. Flow path of the generic pharmaceutical manufacturing process with the principal sampling locations indicated.

The API is usually transported to a separate facility where the drug product, or *formulation*, is manufactured. The dose unit, which a patient receives, is only rarely the pure API. Instead the API is typically *mixed* with *excipients* to develop a formulation that is called the drug product. The third stage of sampling (Figure 1) is performed to identify the API *before* it is used at the manufacturing facility where the formulation is prepared. Sampling is also required here to identify and characterise the excipients that will be mixed in to obtain the desired formulation, and all packaging materials that will come in contact with the pharmaceutical product must also be characterised. The API and excipients are frequently identified by mid-infrared spectroscopy since each compound has a unique spectrum in this wavelength region.[2]

Near infrared (NIR) spectroscopy is used extensively in pharma to identify and characterise incoming *raw materials* even though the differences in NIR spectra are more subtle than those observed in mid-IR spectroscopy. This is a task which is greatly helped by involving chemometrics, especially applying multivariate calibration and proper validation.[3,4] NIR spectroscopy is also able to discern between raw materials on the basis of the differences in their physical properties.[5] The sampling and identification of these materials is a required cGMP regulation, and Section 211.84 (a) states: *"Each lot of components, drug product containers, and closures shall be withheld from use until the lot has been sampled, tested, or examined, as appropriate, and released by the quality control unit."*

History has proven that this sampling stage is vitally important. In 2006, at least 100 citizens (mostly children) tragically died in Panama after consuming cough syrup prepared with di-ethylene glycol instead of the specified glycerin. One mid-infrared or NIR spectrum could have avoided this tragedy, but none of the five companies that brought the material from China to Panama analysed the material.[6] We here emphasise this incident, as a stark reminder that proper sampling and proper analysis are key components of due diligence. TOS certainly has a key role to play within the pharmaceutical industry.

After they are satisfactorily identified, *excipients* play a major role in pharmaceutical formulations. For context, some API are extremely potent and 1 mg may be sufficient to obtain a therapeutic effect. But 1 mg as a direct drug delivery is not handled easily by a patient, which is why excipients are used as *diluents* to obtain a dose unit with a greater tablet mass. Pharmaceutical formulations very often involve mixing of excipients and one or more API; sampling of these *mixtures* constitutes the fourth stage of sampling in the pharmaceutical manufacturing pathway. Approximately 80% of pharmaceutical dose units sold are solid oral dosage forms (tablets, capsules), due to the convenience of administration of this type of drug delivery. The solid form is also important to obtain a drug product with a longer date of expiration, as API are typically more stable in the solid state than in the liquid state. It is typically *powder mixtures*, the most prevalent results from the mixing of the API and excipients that are used to form tablets and capsules. Sampling of in-process powder mixtures as well as of the final dose units are *mandated* by the current good manufacturing practices (cGMPs), and the API concentration must be determined in both stages.

The final drug product or dose unit is also analysed. Sampling of the drug product at the unit dose constitutes the final stage of sampling in pharmaceutical manufacturing. Dose units from throughout the entire production batch are sampled and sent to the quality control lab for analysis. Most of the analyses are mandated to be performed with High Performance Liquid Chromatography (HPLC), which requires breaking up tablets and capsules for extraction of the API from the formulation. These sample preparation steps require time and only 10–30 dose units are usually analysed per lot (typically a lot can have 3,000,000 or more tablets/capsules). The traditional discussion of sampling for final dosage forms characterisation has very much been focused on the number of samples that are needed to fully evaluate the drug content of the individual units. These are areas in which traditional statistics is well applied, but it is unfortunately not always the case that the analytical data supplied to statistical treatment are proven to be representative—TOS to the fore.

The role of sampling in pharmaceutical manufacturing

Within pharma the importance of sampling is clearly acknowledged, e.g. by Brittain, "Samples are therefore defined as the units upon which a program of testing is conducted."[7] The effectiveness of all quality control activities depends on the samples acquired, or, in clear text: if ever there were a context in which only representative samples are acceptable, the pharmaceutical manufacturing pathway must rank among the most important examples—health and lives are at stake.

But in spite of this general recognition, application of the Theory of Sampling (TOS) in pharmaceutical production is very limited.[8–10] Brittain does make reference to Gy's work in his widely accessible article, but TOS may in this, as well as in many other industrial sectors, often have been perceived as relating more to the mining and metallurgical industries, and not to pharmaceutical applications because at that time the major successful examples and case histories mostly still came from this sector. Regardless of the reasons and the very few articles that apply TOS to pharmaceutical formulations, TOS is unquestionably a critical asset also for pharmaceutical manufacturing as discussed further in this article and in several companions to be presented at WCSB7.

Regulatory requirements

This section describes the regulatory requirements in the current Good Manufacturing Practices and European Commission Rules related to sampling. There are many similarities between these regulations and those of the World Health Organisation and other agencies. Pharmaceutical cGMP and European regulatory requirements emphasise that "written procedures must be followed to obtain representative samples". The procedures must specify the number of containers to be sampled, the amount of material to be taken, and the need for appropriate statistical criteria for component variability, confidence intervals, and the degree of precision required. The cGMPs also indicate that if it is necessary to sample a component from the top, middle and bottom of its container, these sample sub-divisions should *not* be composited for testing.[11] The European rules indicate that *"the identity of a complete batch of starting materials can only be ensured if individual samples are taken from all the containers and an identity test performed on each sample."* [12] These requirements are quite understandable given concerns for the identity of incoming raw materials. Finally, the materials shall be withheld from use until the samples are analysed by the quality control unit.

In the cGMPs a representative sample means a *"sample that consists of a number of units that are drawn based on rational criteria such as random sampling and intended to assure that the sample accurately portrays the material being sampled."* The European rules do *not* define a representative sample, however. Representativeness criteria, as detailed in TOS, are missing in these regulations.[11,12] The cGMPs also never discuss "sampling correctness", and never make a distinction between samples and specimens. However, the cGMPs (CFR 210.1) does clearly specify that the regulation contains the *minimum* current good manufacturing practice. Thus, additional emphasis and scientific approaches proven (otherwise, elsewhere) to lead to unambiguous representative sampling, such as described by the Theory of Sampling, are principally not beyond the scope of the regulations.

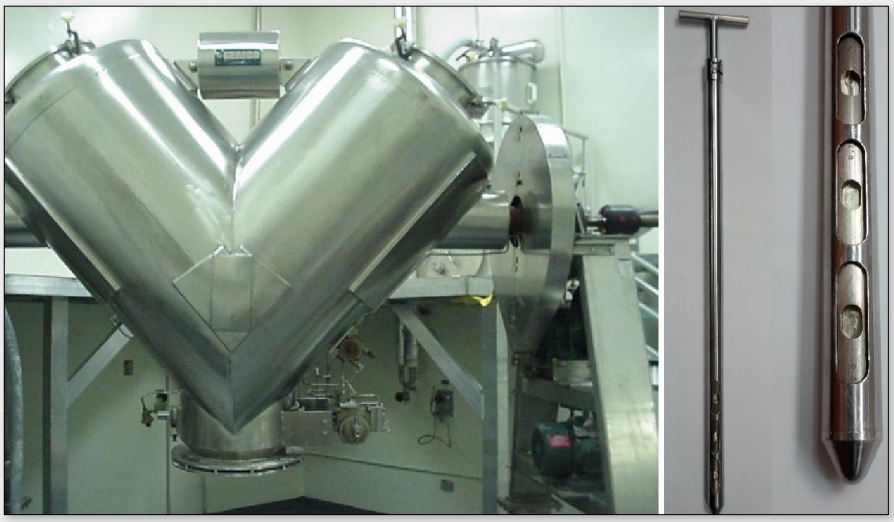

Figure 2. Sampling thieves are used extensively to extract single samples from blenders with various fixed geometrical schemes.[13] Left: large V-blender used in pharmaceutical industry; right: expanded view of sample thief and die cavity.

Sampling associated with *sample* thieves—and its many difficulties

Adequate analysis of excipients and API powder mixtures is required by cGMP regulations. Powder mixtures are extracted from blenders, an operation that overwhelmingly has been performed with the use of sampling *spears*, called sample *thieves* in the pharmaceutical manufacturing realm. Figure 2 illustrates the traditional approach for using thief sampling from a V-blender. Usually a fixed number of samples are required (6 or 10); we treat the details of thief blender sampling in a WCSB7 contribution.[13] The sample thieves employ pre-set cavities to assure that the powder mixture extracted has approximately the mass of a single dose unit, which from a "consumer" point of view appears as a very reasonable demand: the analytical result must pertain to the dose unit the patient receives, but see Reference 13 regarding the fundamental sampling error (FSE).

The use of sample thieves emanates from the understanding that there *could* be "dead spots"—areas of incomplete mixing and drug agglomerates within the blender.[14] The emphasis has been on protecting the patient from a possible over-potent or sub-potent dose unit and identifying these units within the blender, and then improving the blending process to minimise the risk. The Blend Uniformity Work Group composed by members of industry, academia, and the FDA developed the stratified sampling guidance to address these concerns. Stratified sampling was defined as "the process of selecting units deliberately from various locations within a lot of batch or from various phases or periods of a process to obtain a sample.[33] Stratified sampling of the blend and dosage units specifically targets locations either in the blender or throughout the compression/filling operation, which have the higher risk of producing failing content uniformity results."[15]

It will come as no surprise for the TOS community that careful evaluation of sample thieves has shown that they are most often unable to furnish representative samples. Thus, there are many opportunities to improve the sampling and evaluation of powder mixtures through the use of TOS with respect to eliminating ISE and/or FSE.

TOS has for too long not been recognised as an essential component in modern pharmaceutical manufacturing implementations. The present authors are currently collaborating extensively in this endeavour, focusing on the liberating opportunities of basing process monitoring (mixing process in particular) on a rational basis of introducing variographic analysis and characterisation.

Sampling in PAT—sampling with non-destructive methods

In 2004, the FDA published the famous Process Analytical Technology (PAT) Guidance, starting a well-planned effort to bring the latest science and engineering principles into pharmaceutical manufacturing to improve the quality of pharmaceutical products.[16] PAT requires careful study of the API or formulation processes to understand what process parameters can affect the quality of the product. These *critical parameters* are then measured during the process with sufficient frequency so that the information obtained can be used for feed-forward process control and quality assurance. Such PAT measurements can be simple, e.g. as concerns a reaction where the critical parameters are temperature or pH, and these can then be controlled. They may, of course, also be more complex and typically require spectroscopic methods for determination of drug concentration in a powder mixture or suspension for example.[3,17] Spectroscopic methods are often described as real-time, non-destructive methods, since they are able to provide measurements quickly and do not require dissolving samples (which is common for the majority of chemical analyses).

Being able to perform measurements precisely of the critical process, or product parameters constitute a continuous quality assurance process. PAT is now seen as part of the wider Quality by Design (QbD) initiative since the objective is to *design* processes to achieve quality and avoid relying on inspection and reliable removal of non-compliant products.

PAT represents significant progress in pharmaceutical manufacturing. However, in this realm of "advanced manufacturing" the basic principles of TOS are still important,

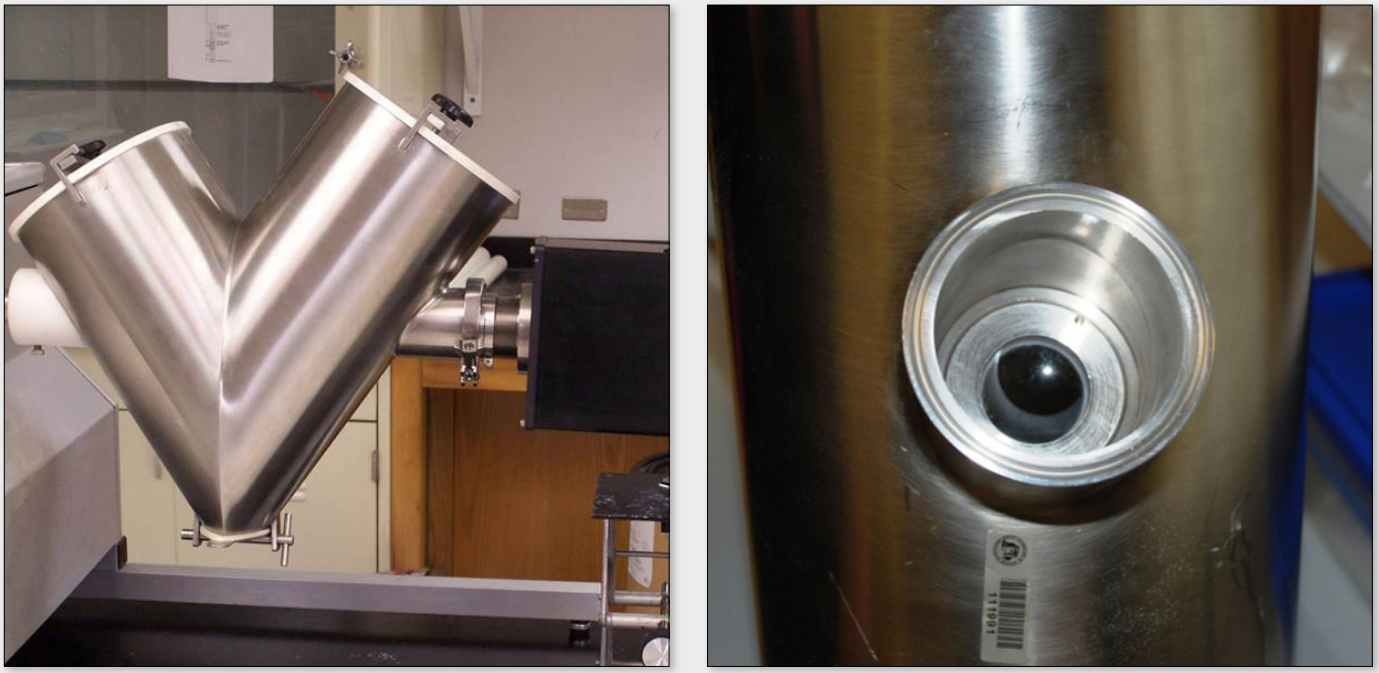

Figure 3. The "no sampling" fallacy in PAT. The NIR spectrometer is optically sampling the powder within the blender through a sapphire window. Left: tumble blender and wireless NIR spectrometer that rotates with the blender. The spectrometer obtains a signal for the material but only to a depth of less than 2 mm below of the window shown on the right. The remaining part of the material is not analysed. The assumption is that the vigorous mixing/blending allows for a meaningful averaging of the signal characterisation of the whole volume. This is a lab-scale system; typically manufacturing tumble blenders are much larger, which introduces ever greater scale-up issues.

and sadly, still almost lacking.[3] A spectroscopic method is still analysing only a very small part of the mass of the entire lot. Thus, the spectroscopic method is in effect doing "optical sampling" of the lot, through the interaction of electromagnetic radiation with a particulate matter, but the acquired spectra represents nothing but grab sampling of very small masses; the likelihood of a significant FSE is very high indeed. Figure 3 shows a wireless NIR spectrometer affixed to a tumble blender. The NIR method is able to obtain spectra of the powder passing across the sapphire sensor window but only to a depth of approximately 2 mm. Spectroscopic signal acquisition is eliminating the physical process of removing a sample from the process, but does *not* at all eliminate sampling errors to influence the final analytical results. The issue revolves around to which degree a "signal" represents a full cross-section or the pertinent volume of the moving matter, which has been treated in detail elsewhere.[3]

The system shown in Figure 3 leads to a very interesting situation since the heterogeneity of the system is being reduced as the blending progresses. Although the system is not able to sample the entire cross-section, as blending progresses new material will reach the window of the spectrometer. To the degree that blending is sufficiently effective, the Grouping and Segregation (GSE) error will be significantly reduced.

Regulatory agencies will also expect an estimate of the sample mass analysed. This sample mass may be estimated taking into consideration the depth of penetration of the light, the density of the sample and the number of spectra that are averaged in the analysis.[8]

The location and placement of the spectroscopic equipment and its interfacing with the process is essential for the success of any PAT implementation. The resulting spectra could be sampling only the materials next to the interface while the bulk (inner) composition is not analysed (a clear breach of the fundamental sampling principle, FSP). The spectroscopic method would then be analysing some of the sample all the time generating an increment delimitation error (IDE), opening up for an inconstant sampling bias. If this happens the spectroscopic sampling is again no better than physical grab sampling. There is a real need to improve the installations for the spectroscopic methods, and install systems capable of evaluating an entire cross-section of the material.[3]

Assuming a spectroscopic signal can be made representative, each would then correspond to an *increment* (classical TOS style). Now the spectroscopic sampling approach is opened up for a strategy of aggregating several increments to form a problem-dependent *composite sample*; all spectroscopic methods permit averaging any number of spectra ("scans"). This opens up for regular process sampling approaches, well-known from TOS, in which variographic analysis allows estimating the necessary number of increments to be composited to force the total sampling error (TSE) below a desired threshold (fit-for-purpose representativeness), e.g. References 18–21.

The PAT initiative has brought significant challenges in the validation of these non-destructive, real-time methods. Regulatory authorities require that analytical methods be validated. Validation is the term used in the pharmaceutical industry for the study and documentation of method accuracy, precision, range and scope of use of the method—which is precisely the total analytical error (TAE) in TOS parlance. Valid and reliable determination of TAE is a challenge

for any dynamic process where samples are continuously changing, such as a mixing, drying or milling processes. Validation will involve efforts to compare the results of the real-time method to those of an off-line method where samples are sent to the lab. This comparison appears simple, but too often the sample analysed with the PAT method is not the same that is brought to the lab.[9] Examples of this particular issue are legion in many PAT implementations in industry, where the focus has all too often mostly been on the new on-line analytical possibilities and their effective calibration/validation within a chemometric context.[3]

Current PAT methods include thorough evaluations of the total analytical error (TAE), but not of the minimum possible error, MPE, which includes the much more dominant TSE (total sampling error). Thus, there are many opportunities for improving PAT approaches through the use of *variographic analysis* providing estimates of the nugget effect and thus the sum-total (TSE + TAE). TOS has not at all been recognised as an essential component of modern pharmaceutical applications.

The present authors intend to change this perception by a systematic innovation program collaboration, three presentations of which will appear at WCSB7.[13,22,23]

Conclusions

TOS has not been incorporated in the pharmaceutical industry to any significant degree—yet. There exist numerous opportunities for improving existing manufacturing practices, e.g. by eliminating bias-generating errors (ISE) both regarding physical sample extraction and/or when the PAT initiative is brought to bear. The most immediate advantage would appear, though, to be introduction of *variographic analysis* for optimal process monitoring and TSE + TAE control.

In the longer term we expect to see PAT methods where the entire cross-section of moving streams of matter is analysed. This will require careful design of instrument sample interfaces and improvements in the design of spectrometers.

Acknowledgements

This collaboration has been possible thanks to the support of the National Science Foundation (ERC research grant EEC-054085) to RR. Andres Roman Ospino is thanked for preparation of figures.

References

1. J. Bauer, S. Spanton, R. Henry, J. Quick, W. Dziki, W. Porter and J. Morris, "Ritonavir: An extraordinary example of conformational polymorphism", *Pharm. Res.* **18,** 859–866 (2001). doi: http://dx.doi.org/10.1023/a:1011052932607
2. D. Clark and A. Pysik, "The analysis of pharmaceutical substances and formulated products by vibrational spectroscopy", in *Handbook of Vibrational Spectroscopy*, Ed by J.M. Chalmers and P.R. Griffiths. John Wiley & Sons, Chichester, UK (2006). doi: http://dx.doi.org/10.1002/9780470027325.s8909
3. K.H. Esbensen and P. Paasch-Mortensen, "Process sampling: theory of sampling—the missing link in process analytical technologies (PAT)", in *Process Analytical Technology*, Ed by K.A. Bakeev. John Wiley & Sons, Chichester, UK, pp. 37–80 (2010). doi: http://dx.doi.org/10.1002/9780470689592.ch3
4. K.H. Esbensen and P. Geladi, "Principles of Proper Validation: use and abuse of re-sampling for validation", *J. Chemometr.* **24,** 168–187 (2010). http://dx.doi.org/10.1002/cem.1310
5. M. Blanco and M.A. Romero, "Near-infrared libraries in the pharmaceutical industry: a solution for identity confirmation", *Analyst* **126,** 2212–2217 (2001). http://dx.doi.org/10.1039/b105012p
6. W. Bogdanich and J. Hooker, "From China to Panama, a trail of poisoned medicine", *New York Times* 6 May (2007).
7. H.G. Brittain, "Particle-size distribution II: the problem of sampling powdered solids. (pharmaceutical physics)", *Pharm. Technol.* **26,** 67–67 (2002).
8. Y. Colón, M. Florian, D. Acevedo, R. Méndez and R. Romañach, "Near infrared method development for a continuous manufacturing blending process", *J. Pharm. Innov.* **9,** 291–301 (2014). doi: http://dx.doi.org/10.1007/s12247-014-9194-1
9. R.L. Green, G. Thurau, N.C. Pixley, A. Mateos, R.A. Reed and J.P. Higgins, "In-line monitoring of moisture content in fluid bed dryers using Near-IR spectroscopy with consideration of sampling effects on method accuracy", *Anal. Chem.* **77,** 4515–4522 (2005). doi: http://dx.doi.org/10.1021/ac050272q
10. A. Ghaderi, "On characterization of continuous mixing of particulate materials", *Particul. Sci. Technol.* **21,** 271–282 (2003). doi: http://dx.doi.org/10.1080/02726350307484a
11. *Current Good Manufacturing Practice in Manufacturing, Processing, Packing, or Holding of Drugs*. General, USDoHaH Services, US Federal Register (2014).
12. "Annex 8 sampling of starting and packaging materials", in *EudraLex—Volume 4 Good Manufacturing Practice (GMP) Guidelines*. http://ec.europa.eu/health/files/eudralex/vol-4/pdfs-en/anx08_en.pdf
13. K.H. Esbensen and R.J. Romañach, "Proper sampling, variographic analysis, total measurement uncertainty & fit-for-purpose acceptance levels for pharmaceutical mixing processes monitoring", in *Proceedings WCSB7*, in press (2015).
14. J.H. Timmermans, "A Report of the PQRI Workshop on Blend Uniformity", *Pharmaceutical Technology.* **25,** 76-76-84 (2001).
15. G. Boehm, "Report on the industry blend uniformity practices survey", *Pharm. Technol.* **25,** 20–26 (2001).
16. *Guidance for Industry PAT—A Framework for Innovative Pharmaceutical Development, Manufacturing, and Quality Assurance*. US FDSA (2004).
17. 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", *Chem. Eng. Sci.* **65,** 5728–5733 (2010). doi: http://dx.doi.org/10.1016/j.ces.2010.01.036
18. F.F. Pitard, *Pierre Gy's Sampling Theory and Sampling Practice. Heterogeneity, Sampling Correctness, and Statistical Process Control*. CRC Press (1993).
19. P. Gy, *Sampling for Analytical Purposes*, 1st Edn. Wiley, New York (1998).
20. K.H. Esbensen, H.H. Friis-Petersen, L. Petersen, J.B. Holm-Nielsen and P.P. Mortensen, "Representative process sampling—in practice: Variographic analysis and estimation of total sampling errors (TSE)", *Chemometr. Intell. Lab. Sys.* **88,** 41—59 (2007). http://dx.doi.org/10.1016/j.chemolab.2006.09.011
21. *DS 3077*. Danish Standards Foundation (2013).
22. A. Roman Ospino, C. Ortega, A. Sanchez, S. Ortiz, R. J. Romañach and K. H. Esbensen, "TOS to the rescue—determining TSE for near infrared spectroscopic analysis of pharmaceutical blends", in *Proceedings WCSB7*, in press (2015).
23. A. Sánchez Paternina, A. Roman Ospino, B. Alvarado, K.H. Esbensen and R.J. Romanach, "TOS when 'homogeneity' is expected—TOS in Pharmaceutical Manufacturing", in *Proceedings WCSB7*, in press (2015).