Short communication

Christel Kamp*, Björn Becker, Walter Matheis, Volker Öppling and Isabelle Bekerredjian-Ding

How to draw the line – Raman spectroscopy as a tool for the assessment of biomedicines

<https://doi.org/10.1515/hsz-2020-0388>

Received December 17, 2020; accepted March 30, 2021;
published online April 13, 2021

Abstract: Biomedicines are complex biochemical formulations with multiple components that require extensive quality control during manufacturing and in subsequent batch testing. A proof-of-concept study has shown that an application of Raman spectroscopy can be beneficial for a classification of vaccines. However, the complexity of biomedicines introduces new challenges to spectroscopic methodology that require advanced experimental protocols. We further show the impact of analytical protocols on vaccine classification using R as an Open Source data analysis platform. In conclusion, we advocate for standardized and transparent experimental and analytical procedures and discuss current findings and open challenges.

Keywords: machine learning; pre-processing; quality control; Raman spectroscopy; standardisation; vaccine.

Biomedicines, such as vaccines or therapeutic allergens, are complex biochemical formulations containing multiple components. Biotechnological production or extraction from biological sources results in inherent variability and thus requires high levels of standardisation in production and quality control to ensure drug safety and efficacy. For many products, in particular vaccines, manufacturers' process controls are complemented by batch release testing through governmental medicines control laboratories. Batch release testing of vaccines and allergen products performed at the Paul-Ehrlich-Institut ranges from biochemical and immunological assays to *in vivo* testing, which is expensive and time-consuming.

Use of animals further raises ethical concerns. Vibrational spectroscopy (infrared- and Raman-spectroscopy) has already proven to be a valuable tool to address quality control for pharmaceutical products, largely specific, solid chemical compounds (Bunaciu and Aboul-Enein 2017; Ewing and Kazarian 2018; Pivonka et al. 2007). Applications are, on one side the identification of compounds and on the other the detection of counterfeit products. Spectral measurements can be calibrated through reference samples allowing for quantitative measurements (Bonnier and Byrne 2012; Bonnier et al. 2017; Byrne et al. 2020; Makki et al. 2019; Parachalil et al. 2020). Identification and quantification capabilities can be used for the detection of sub-potent formulations, contaminations and other failures to meet quality criteria. Optical methods are non-destructive, do not require specific labelling and are rapid and cost-efficient.

Our initial proof-of-concept studies have shown that an analogous application of vibrational spectroscopy can be equally beneficial in the quality control of vaccines (Silge et al. 2018). The approach can be extended to other biomedicines (Butler et al. 2016) and matches the need for rapid and reliable measurement techniques in quality control of these products. However, the complex nature of biomedicines introduces challenges to spectroscopic methods that need to be addressed by basic research (Baker et al. 2018). Complex, liquid formulations result in strong spectral signals from water and excipients that need to be differentiated from the often weaker signal of the pharmacologically relevant active substances (Bonnier et al. 2017; Parachalil et al. 2020; Zhao et al. 2015). Vaccines are often inhomogeneous and turbid formulations that may show highly localized signals with temporal changes due to sedimentation. Their essential compounds are immunogenic proteins, polysaccharides, and attenuated pathogens and immune enhancers (adjuvants). A work-around to reduce signals from water and excipients and to stabilize measurement conditions is to study dried products. This however, comes with its own challenges in standardisation, in particular with respect to drying protocols and crystallization patterns. No calibrations have as yet been established to link the findings in dried products

*Corresponding author: Christel Kamp, Paul-Ehrlich-Institut, Paul-Ehrlich-Straße 51-59, D-63225 Langen, Germany, E-mail: christel.kamp@pei.de. <https://orcid.org/0000-0001-9997-6772>

Björn Becker, Walter Matheis, Volker Öppling and Isabelle Bekerredjian-Ding, Paul-Ehrlich-Institut, Paul-Ehrlich-Straße 51-59, D-63225 Langen, Germany, E-mail: Bjoern.Becker@pei.de (B. Becker), Walter.Matheis@pei.de (W. Matheis), Volker.Oeppling@pei.de (V. Öppling), Isabelle.Bekerredjian-Ding@pei.de (I. Bekerredjian-Ding)

with information about the native product and its active components.

Aside from these probe-specific complications Raman spectroscopy is further challenging because it provides very detailed information about the sample and measurement conditions alike. Signals of interest are typically superimposed by interfering signals such as variable background signals and noise from measurement devices, sample fluorescence, substrate specific signals or sporadic peaks from cosmic radiation. Furthermore, changes that may occur in the wavenumber (x-axis) and intensity (y-axis) of spectra require calibration (Dörfer et al. 2011). Extraction of the signal of interest through pre-processing of raw spectral data is generally considered as mandatory for a subsequent robust and accurate classification (Lasch 2012).

For our study, we followed the experimental protocol for vaccines dried on CaF_2 slides as developed and described in Silge et al. (2018) for the vaccines listed in Table 1. For each dried vaccine spot (replicate) we have measured a grid of 100 spectra, each with an acquisition time of 30 s at a laser excitation wavelength of 785 nm and laser power of 100 mW (BioRam, CellTool GmbH). The choice of an excitation wavelength in the near infrared wavelength regime – as compared to the earlier study which used an excitation wavelength of 514 nm – can reduce fluorescence background but requires longer acquisition times. Measurements were started after an initial drying time of 30 min when vaccine dots were visibly dry on CaF_2 slides and six dots were measured consecutively (one per hour). In line with earlier findings (Silge et al. 2018), the first two to three measurements showed larger variability before equilibration of the spectra seen in dried samples, which was however minor as compared to between vaccine variation (cf. Supplementary Figure S2). In cases of availability of more than one manufacturing batch per vaccine we found negligible between batch variation as compared to between vaccine variation (cf. Supplementary Figures S2 and S3).

The focus of the current study was to explore the impact of pre-processing schemes on subsequent classification of

spectral representations – or fingerprints. To ensure transparent and reproducible pre-processing of raw spectra we used the Open Source environment of the statistical programming language R with the package hyperSpec (Beleites and Sergo) for the pre-processing and analysis of spectral data and the EMSC package for extended multiplicative scatter correction (Liland et al. 2016). The R Stats core function `prcomp()` was used for principal component analysis, `lda()` for linear discriminant analysis. All spectra were calibrated for potential shifts in the wavenumber axis using a Paracetamol reference measurement for each measurement day prior to further pre-processing (Dörfer et al. 2011). To assess the impact of different pre-processing procedures, we followed three variant procedures:

Procedure 1: Spectra with values more than four standard deviations from the mean spectrum were skipped as outliers or as contaminated by cosmic spikes. Subsequently, linear baselines were removed and spectra smoothed at a resolution of four wavenumbers in the range 400–3200 cm^{-1} (without down-sampling). This is followed by an intensity normalisation in which each spectrum is divided by its mean value over the considered wave number range (area normalisation).

Procedure 2: Procedure 1 is followed with an additional constraint to the wavenumber range of 400–1500 cm^{-1} which was chosen due to strong background signals in our setting above 1500 cm^{-1} .

Procedure 3: Procedure 2 is followed with additional rescaling within each measurement lot of 100 spectra applying extended multiplicative scatter correction using the mean spectrum of each replicate measurement of 100 spectra as a reference (EMSC [Liland et al. 2016]).

We analyzed spectra of five vaccines as listed in Table 1 and background spectra measured without vaccine as a control largely representing the background signal from the CaF_2 slide and measurement device. All spectra were subjected

Table 1: Vaccine products used in this study.

Vaccine product (antigen composition)	Type	Number of manufacturer's batches	Total number of replicates
DTaP-IPV-HepB	For primary vaccination	3	18
dTaP	For booster vaccination	3	24
dTaP-IPV	For booster vaccination	3	18
Pneu1	For primary vaccination	1	6
Pneu2	For primary vaccination	1	6

Vaccine antigens: T, tetanus, d; D, diphtheria (low and high antigen content); aP, acellular pertussis antigen; IPV, inactivated poliovirus; HepB, hepatitis B; Pneu, pneumococcal capsular polysaccharides.

to pre-processing procedures 1–3 and major variability in the resulting spectral data was subsequently assessed through principal component analysis. The reduction of spectral space to dimensions of highest variability in the data allows for an exploratory overview of (dis-)similarities between vaccine spectra. Mean spectra of each vaccine type together with the first four principal components are shown in Figure 1. The latter represent directions in spectral space showing highest variance in data (in consecutive order). This technique of unsupervised learning allows to explore between vaccine (and background) variation without prior knowledge, i.e. spectral measurements were colored in Figure 1 according to vaccine type for identification but this knowledge is not considered in the determination of principal axes.

The results show that vaccines are separated in spectral space but that their arrangement is affected by the chosen pre-processing procedures. This means that subsequent classification based on supervised learning is feasible, but its outcome will be influenced by the chosen pre-processing procedures.

This is not surprising given that the differences in considered spectral ranges between procedures 1 and 2 correspond to a selection of features that might correspond to a focus on or neglect of features that are particularly relevant to distinguish certain vaccines. Feature selection is often necessary to allow for reliable classification of high dimensional data in which their complexity cannot be matched by the available data (Hastie et al. 2017). Extended multiplicative scatter correction used in pre-processing procedure 3 reduces fluctuations in spectral intensity among replicate measurements through model-based alignment with a reference spectrum, in this case a rescaling towards the mean spectrum of each replicate measurement of 100 spectra (Liland et al. 2016). This reduces the effective sample size of measured spectra but also variability in spectral intensity that may arise through inhomogeneous drying and crystallization as can be seen in the right column of Figure 1. Linear discriminant analysis (LDA) was applied as a classification model for the differences between vaccine products (Hastie et al. 2017). The results are summarized in a confusion table (Table 2) showing overall high percentages of correctly classified vaccines. Yet, differences in the model predictions for pre-processing procedures are evident illustrating the impact of spectral pre-processing on subsequent classification.

Given the complexity and variety of measurement setups and devices as well as measurement samples we have to acknowledge that there is no globally optimal or

preferable pre-processing procedure. Naturally, pre-processing is guided by the aim to selectively extract features relevant for the specific question addressed by the spectroscopic method (Gerretzen et al. 2015). Distinctive features may be expected within in specific spectral ranges or require a certain spectral resolution (Larkin 2011; McCreery 2000). However, chemometric prediction models are affected by the pre-processing of spectral data that are used for their training and validation, this holds particularly for samples with overlapping spectral characteristics. Therefore, the choice of pre-processing procedures is part of the design of experiment (Gerretzen et al. 2015) which should be made transparent as part of the chemometric model: knowledge about prior information guiding feature selection and the choice of algorithms and their parameters used in pre-processing are relevant to define the range of application and limitations of a chemometric model. A reproducible implementation of the analysis pipeline within an OpenSource framework is particularly favourable as its full transparency leaves little space for ambiguities and can support community efforts towards standards and best practices in spectral pre-processing.

Overall, we consider reproducibility and standardisation as one of the major challenges that need to be met at various levels to exploit the full potential of Raman spectroscopy for pharmaceutical applications in the context of biomedicines. While these challenges have already been met for many pharmaceutical applications, our case study on vaccines is a good showcase for the challenges encountered in the spectral characterization of biomedicines including instrumentation, sample preparation and chemometric modelling. Drying of vaccines increases the concentration of components of interest but can result in product- and preparation-specific, inhomogeneous crystallization patterns, which may not be stable over time and subject to environmental conditions (Silge et al. 2018). As the assessment of the native product is the eventual goal experimental protocols for measurements of native products or calibration procedures linking to substitute formulations (e.g. dried products) need to be developed (Byrne et al. 2020). The former involves on the one hand to capture low concentrations of medically relevant compounds such as proteins or polysaccharides in liquid solution and on the other hand to deal with complex, turbid formulations showing sedimentation. This requires sophisticated sample preparation and instrumentation – again, calling for standardisation.

A recent proficiency study (Guo et al. 2020) has made valuable contributions in analyzing the impact of various spectroscopic platforms and devices on the primary acquisition of raw spectral data. Similar care is mandatory for

Figure 1: Presentation of vaccine spectra along four principal component axes (PC1–PC4) for pre-processing procedures 1–3. Principal component axes represent highest variance seen in the data (percent explained variance in brackets). Vaccine spectra show distinct clustering in spectral space, however, their particular arrangement is influenced by the chosen pre-processing procedures. Corresponding mean spectra are shown in the bottom panel with bands indicating the 16th and 84th percentile (corresponding to $\pm$ one standard deviation for normally distributed data).

Table 2: Confusion table for vaccine products used in this study.

		Vaccine					Sensitivity	Specificity
Procedure 1 96% correctly classified		dTaP	dTaP-IPV	DTaP-IPV-HepB	Pneu1	Pneu2		
Prediction	dTaP	888	5	16	3	1	96%	99%
	dTaP-IPV	5	1159	4	3	1	94%	100%
	DTaP-IPV-HepB	30	1	1002	0	3	98%	99%
	Pneu1	5	66	1	292	2	96%	98%
	Pneu2	0	3	2	5	318	98%	100%
Procedure 2 96% correctly classified		dTaP	dTaP-IPV	DTaP-IPV-HepB	Pneu1	Pneu2	Sensitivity	Specificity
Prediction	dTaP	902	5	28	8	3	97%	98%
	dTaP-IPV	5	1162	10	0	0	94%	99%
	DTaP-IPV-HepB	13	2	979	0	1	96%	99%
	Pneu1	5	60	1	290	1	96%	98%
	Pneu2	3	5	7	5	320	98%	99%
Procedure 3 97% correctly classified		dTaP	dTaP-IPV	DTaP-IPV-HepB	Pneu1	Pneu2	Sensitivity	Specificity
Prediction	dTaP	904	5	23	0	0	97%	99%
	dTaP-IPV	4	1163	8	0	0	94%	100%
	DTaP-IPV-HepB	17	3	989	0	0	96%	99%
	Pneu1	2	63	1	303	1	100%	98%
	Pneu2	1	0	4	0	324	100%	100%

Classifications were comparatively tabulated following a 6-fold cross validation: models are trained on five vaccine spots (each with a lot of 100 spectra, as used in EMSC of pre-processing procedure 3) and predictions were made on the remaining vaccine spot based on on spectra that were treated with pre-processing procedures 1–3 (number of spectra, percentages of sensitivity and specificity are rounded to whole percentages). Overall there is a high percentage of correctly classified spectra. However, changes in pre-processing procedures introduce shifts in misclassified spectra.

sample preparation of biomedicines, i.e. heterogeneous and often liquid formulations, for which reproducible and stable measurement conditions have to be defined. Finally, a well-defined and transparent integration of pre-processing and chemometric modelling will help to develop a reliable and reproducible data- and knowledge-base of spectral information. This will ultimately allow to define spectral fingerprints of biomedical products and to draw the line at falsified products or products of low quality.

Acknowledgements: The authors would like to acknowledge Claudia Beleites for many insightful discussions and advice on the R package hyperSpec as well as general aspects of spectroscopic techniques and intricacies.

Author contributions: All the authors have accepted responsibility for the entire content of this submitted manuscript and approved submission.

Research funding: None declared.

Conflict of interest statement: The authors declare no conflicts of interest regarding this article.

References

- Baker, M.J., Byrne, H.J., Chalmers, J., Gardner, P., Goodacre, R., Henderson, A., Kazarian, S.G., Martin, F.L., Moger, J., Stone, N., et al. (2018). Clinical applications of infrared and Raman spectroscopy: state of play and future challenges. *The Analyst* 143: 1735–1757.
- Beleites, C. and Sergo, V. *hyperSpec. a package to handle hyperspectral data sets in R*.
- Bonnier, F., Blasco, H., Wasselet, C., Brachet, G., Respaud, R., Carvalho, L.F.C.S., Bertrand, D., Baker, M.J., Byrne, H.J., and Chourpa, I. (2017). Ultra-filtration of human serum for improved quantitative analysis of low molecular weight biomarkers using ATR-IR spectroscopy. *Analyst* 142: 1285–1298.
- Bonnier, F. and Byrne, H.J. (2012). Understanding the molecular information contained in principal component analysis of vibrational spectra of biological systems. *Analyst* 137: 322–332.

- Bunaciu, A.A. and Aboul-Enein, H.Y. (2017). Vibrational spectroscopy applications in drugs analysis. In: *Encyclopedia of Spectroscopy and spectrometry*. Elsevier, pp. 575–581.
- Butler, H.J., Ashton, L., Bird, B., Cinque, G., Curtis, K., Dorney, J., Esmonde-White, K., Fullwood, N.J., Gardner, B., Martin-Hirsch, P.L., et al. (2016). Using Raman spectroscopy to characterize biological materials. *Nat. Protoc.* 11: 664–687.
- Byrne, H.J., Bonnier, F., McIntyre, J., and Parachalil, D.R. (2020). Quantitative analysis of human blood serum using vibrational spectroscopy. *Clin. Spectrosc.* 2: 100004.
- Dörfer, T., Bocklitz, T., Tarcea, N., Schmitt, M., and Popp, J. (2011). Checking and improving calibration of Raman spectra using chemometric approaches. *Z. Phys. Chem.* 225: 753–764.
- Ewing, A.V. and Kazarian, S.G. (2018). Recent advances in the applications of vibrational spectroscopic imaging and mapping to pharmaceutical formulations. *Spectrochim. Acta Mol. Biomol. Spectrosc.* 197: 10–29.
- Gerretzen, J., Szymańska, E., Jansen, J.J., Bart, J., van Manen, H.-J., van den Heuvel, E.R., and Buydens, L.M.C. (2015). Simple and effective way for data preprocessing selection based on design of experiments. *Anal. Chem.* 87: 12096–12103.
- Guo, S., Beleites, C., Neugebauer, U., Abalde-Cela, S., Afseth, N.K., Alsamad, F., Anand, S., Araujo-Andrade, C., Aškrabić, S., Avci, E., et al. (2020). Comparability of Raman Spectroscopic configurations: a large scale cross-laboratory study. *Anal. Chem.* 92: 15745–15756.
- Hastie, T., Tibshirani, R., and Friedman, J.H. (2017). *The elements of statistical learning. Data mining, inference, and prediction*. New York, NY: Springer.
- Larkin, P. (2011). Introduction. In: *Infrared and Raman spectroscopy*. Elsevier, pp. 1–5.
- Lasch, P. (2012). Spectral pre-processing for biomedical vibrational spectroscopy and microspectroscopic imaging. *Chemometr. Intell. Lab. Syst.* 117: 100–114.
- Liland, K.H., Kohler, A., and Afseth, N.K. (2016). Model-based pre-processing in Raman spectroscopy of biological samples. *J. Raman Spectrosc.* 47: 643–650.
- Makki, A.A., Bonnier, F., Respaud, R., Chtara, F., Tfayli, A., Tauber, C., Bertrand, D., Byrne, H.J., Mohammed, E., and Chourpa, I. (2019). Qualitative and quantitative analysis of therapeutic solutions using Raman and infrared spectroscopy. *Spectrochim. Acta Mol. Biomol. Spectrosc.* 218: 97–108.
- McCreery, R.L. (2000). *Raman spectroscopy for chemical analysis*. New York: Wiley-Interscience.
- Parachalil, D.R., McIntyre, J., and Byrne, H.J. (2020). Potential of Raman spectroscopy for the analysis of plasma/serum in the liquid state: recent advances. *Anal. Bioanal. Chem.* 412: 1993–2007.
- Pivonka, D.E., Chalmers, J.M., and Griffiths, P.R. (Eds.) (2007) *Applications of vibrational spectroscopy in pharmaceutical research and development*. Chichester, West Sussex, UK: Wiley.
- Silge, A., Bocklitz, T., Becker, B., Matheis, W., Popp, J., and Bekereditian-Ding, I. (2018). Raman spectroscopy-based identification of toxoid vaccine products. *NPJ Vaccines* 3: 50.
- Zhao, Y., Ji, N., Yin, L., and Wang, J. (2015). A non-invasive method for the determination of liquid injectables by Raman spectroscopy. *AAPS PharmSciTech* 16: 914–921.

Supplementary Material: The online version of this article offers supplementary material (<https://doi.org/10.1515/hsz-2020-0388>).