Data mining implementation in multiparticulate unit systems characterization

Ivana Vasiljević, Erna Turković, Jelena Parojčić

Department of Pharmaceutical Technology and Cosmetology, University of Belgrade – Faculty of Pharmacy, Vojvode Stepe 450, 11221 Belgrade, Serbia

Introduction

Multiparticulate unit systems (MPUs) represent versatile drug delivery systems which provide flexible dosing and reproducible drug release profile (Al-Hashimi et al., 2018). They are usually filled in capsules/sachets or compressed into tablets, and, in order to ensure good processability, thorough MPU characterization is required. The aim of this work was to explore the impact of composition and preparation method on MPU characteristics and identify potential patterns between the data using advanced computational analysis.

Materials and methods

Multiparticulate unit systems were prepared using four distinct preparation methods, namely: extrusion/spheronization (P-series samples), manual kneading/granulation (G-series samples), liquisolid system MPU technology (L-series samples) and selective laser sintering (S-series samples). MPU composition was polymer-based, while either caffeine (CAF) or ibuprofen (IBU) was incorporated as model drug.

The prepared MPUs (38 samples) were characterized regarding: bulk and tapped density, Hausner ratio, Carr index, angle of repose, tensile strength obtained at 500 kg-compression load, inter-particulate porosity, moisture content, fraction of particles smaller than 350 μm and aspect ratio, assessed by dynamic image analysis.

In order to develop predictive models linking MPU characteristics with MPU composition and preparation method, Orange data mining software (version 3.34.0) was used. The generated data set was split into training (70%, i.e. 27 samples) and test subsets (30%, i.e. 11 samples). Different algorithms were applied (k-nearest neighbor algorithm (kNN), decision tree, random forest, neural network, adaptive boosting (AdaBoost)) and their

applicability was evaluated based on the relevant statistical parameter values. In the case of classification algorithms, classification accuracy (CA) and precision was used, while, in the case of regression algorithms, mean squared error (MSE) and coefficient of determination were assessed. In order to obtain models with higher accuracy, some parameters were discretized and the obtained categorical attributes were used. Model prediction accuracy was evaluated by comparing experimentally obtained data and predicted values, using relevant statistical parameters.

Results and discussion

The prepared data set was analyzed for patterns and high positive correlation between aspect ratio and both bulk and tapped density was found (correlation coefficients equal to 0.867 and 0.839, respectively). This indicated that spherical MPUs exhibited high bulk and tapped density, which is desirable for further processing, particularly capsule/sachet filling and coating (Chopra et al., 2002). Both Carr index and Hausner ratio were highly correlated with the angle of repose (correlation coefficients equal to 0.735 and 0.736, respectively), which implied the comparable trend in MPU flowability, evaluated by different methods. It may be postulated that, by reducing the number of methods employed, MPU characterization may be simplified without compromising the data quality. Aspect ratio was negatively correlated with both inter-particulate porosity and fraction of particles smaller than 350 μm (correlation coefficients equal to 0.756 and 0.721, respectively), indicating that irregular particles exhibited higher consolidation tendency and wider particle size distribution.

Bulk density, angle of repose and aspect ratio were thoroughly analyzed using data mining as relevant MPU properties, affecting MPU flowability and processability.

*ivanav@pharmacy.bg.ac.rs

In the case of bulk density, the algorithm which described the trend most accurately was AdaBoost. However, in the second phase performed on the test data subset, decision tree algorithm provided notably higher prediction accuracy and lower difference between experimentally obtained and predicted values (R^2 0.956, MSE 0.002, compared to R^2 0.721, MSE 0.015, for decision tree and AdaBoost predictions, respectively). The investigated MPUs bulk density was primarily affected by preparation method: P- and L-series samples exhibited higher bulk density in comparison to G- and S-series samples. Other relevant parameters were aspect ratio and model drug incorporated: MPUs with irregular shape and low aspect ratio exhibited also low bulk density values, while CAF contributed to the bulk density increase. Moisture content also impacted bulk density values (Fig. 1).
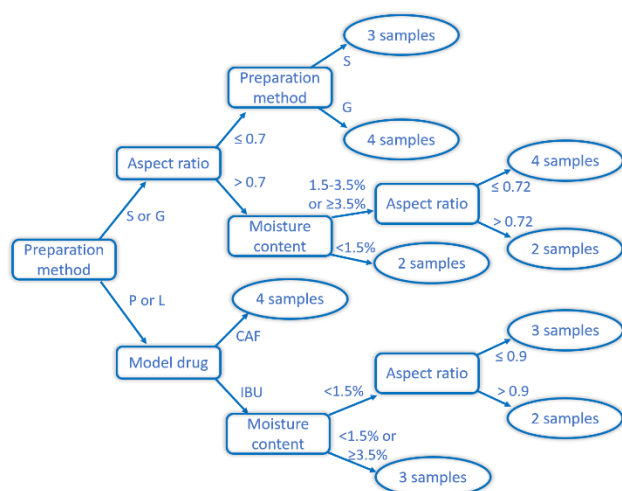

Fig. 1. Decision tree describing MPU bulk density

Angle of repose data were discretized in two groups: lower and higher values than 35°, corresponding to excellent/good and fair/poor flowability, as described in European Pharmacopoeia. Angle of repose tendency was best described by random forest classification algorithm (CA 0.883, precision 0.876). MPU preparation method, polymer and model drug incorporated, aspect ratio and bulk density affected angle of repose. Random forest also provided high level of prediction accuracy, analyzed on the test data subset (CA 0.800, precision 0.800).

Particle shape, evaluated using aspect ratio, was recognized as the most important material property affecting processability (Leane et al., 2018), particularly in the case of MPUs. Random forest was distinguished as the algorithm most suitable for describing the aspect ratio trend in the case of investigated MPUs (R^2 0.854, MSE 0.001). Aspect ratio was primarily affected by fraction of particles smaller than 350 μm, but also MPU preparation method, polymer and model drug incorporated, microcrystalline cellulose and polymer concentration, liquid phase concentration and moisture content. The developed model

was used for aspect ratio prediction on the test data subset and it provided relatively high accuracy (R^2 0.733, MSE 0.005), with regards to model predictability. Lower difference between experimentally obtained and predicted values was observed in the case of P- and L-series samples, in comparison to G- and S-series samples.

Conclusion

MPUs with high aspect ratio exhibited high bulk and tapped density, while their inter-particulate porosity and fraction of particles smaller than 350 μm was lower. Carr index and Hausner ratio correlated with the angle of repose, indicating the possibility of characterization simplification and avoiding data abundance. MPU preparation method notably affected all the investigated MPU properties. Additionally, sample composition, particularly model drug and polymer incorporated, as well as moisture content and fraction of particles smaller than 350 μm affected bulk density, angle of repose and aspect ratio of the obtained MPUs.

The data mining application enabled model development, describing the relationship between MPU preparation method, MPU composition and MPU characteristics. Additionally, relatively high prediction accuracy was observed. However, these conclusions are limited by relatively small sample size used in the present study and further analysis is required for confirmation of correlations detected.

References

Al-Hashimi, N., Begg, N., Alany, R.G., Hassanin, H. and Elshaer, A., 2018. Oral modified release multiple-unit particulate systems: compressed pellets, microparticles and nanoparticles. Pharmaceutics, 10(4), 176. https://doi.org/10.3390/pharmaceutics10040176

Chopra, R., Podczeck, F., Newton, J.M. and Alderborn, G., 2002. The influence of pellet shape and film coating on the filling of pellets into hard shell capsules. Eur. J. Pharm. Biopharm. 53(3), 327-333. https://doi.org/10.1016/S0939-6411(02)00015-2

Leane, M., Pitt, K., Reynolds, G.K., Dawson, N., Ziegler, I., Szepes, A., Crean, A.M., Dall Agnol, R. and Manufacturing Classification System (MCS) Working Group, 2018. Manufacturing classification system in the real world: factors influencing manufacturing process choices for filed commercial oral solid dosage formulations, case studies from industry and considerations for continuous processing. Pharm. Dev. Technol. 23(10), 964-977. https://doi.org/10.1080/10837450.2018.1534863