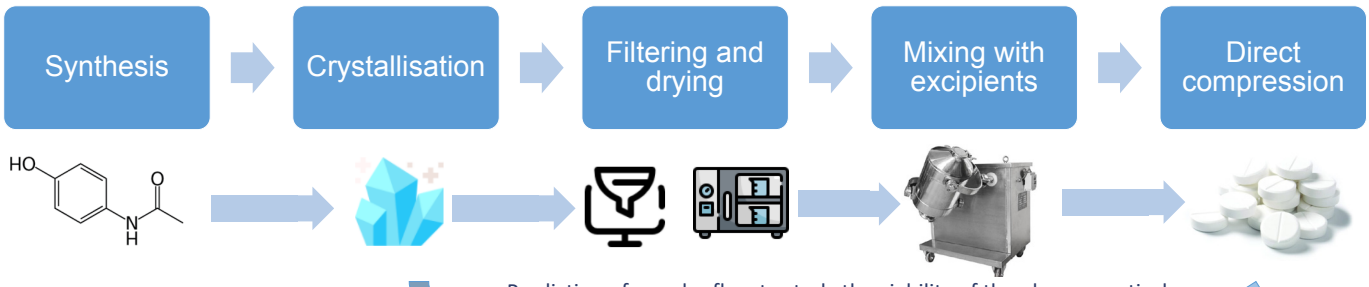
POSTER 76

Prediction of powder flow of pharmaceutical materials using machine learning

Laura Pereira Diaz, Cameron J. Brown, Alastair J. Florence
laura.pereira-diaz@strath.ac.uk
EPSRC CMAC Future Manufacturing Research Hub; Strathclyde Institute of Pharmacy & Biomedical Sciences

Introduction:
The lack of understanding of powder flow adds cost and time to the development of robust production routes and compromises manufacturing process performance in the pharmaceutical industry. In this work, implementing machine learning models enables rapid decision-making regarding manufacturing route selection, thus, minimizing the time and amount of material required. This work focuses on using ML models to predict powder flow behavior of pharmaceutical materials for routine, widely available materials.



Prediction of powder flow to study the viability of the pharmaceutical material for direct compression in early-stage development to save time and resources

Number of pharmaceutical powders:

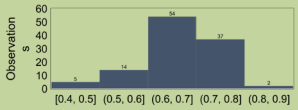
	Excipients	Blends
30	43	40

Data (pure materials and mixtures):

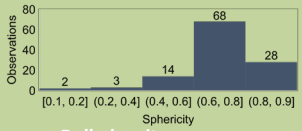
- **Particle size distribution**

Parameter	Ranges (µm)	Median (µm)
D10	9-225	54.84
D50	25-644	149.19
D90	53-1892	328.87
D[3,2]	19-393	94.63

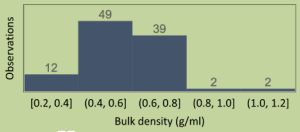
- **Aspect ratio distribution**



- **Sphericity distribution**



- **Bulk density**



- **FFc**

FFc	Behaviour	N. Observations
< 4	Cohesive	29
4 < FFc < 10	Easy flowing	34
> 10	Free flowing	55

- Particle size and shape analysis using QICPIC.
- Bulk density and flow function coefficient measurement using FT4.

Machine learning models

- Support Vector Machines (SVM), Random Forests, neural networks, Naïve Bayes, k-Nearest Neighbours (kNN), Logistic Regression, and Adaboost were all investigated for classification capabilities of powder flow into three categories: cohesive, easy-flowing, free-flowing.
- The performance of each algorithm was evaluated using area under the curve receiver operating characteristics (AUC – ROC).

Classification models

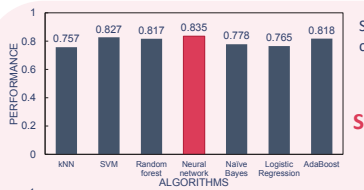
Data availability

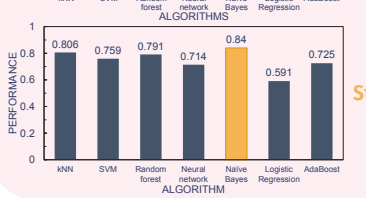
Algorithm selection

Data sampling for testing

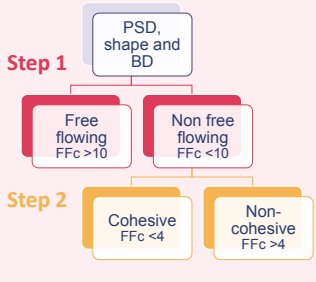
Training the model

Model validation and interpretability





Sampling method for testing: 10-fold cross-validation

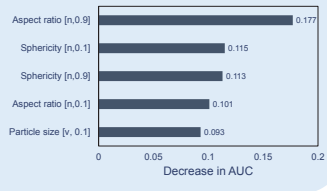


- Feature importance was analysed to improve the knowledge of how the model makes the predictions.
- Feature importance is calculated based on the decrease in AUC – ROC when each individual feature is replaced by noise after 10 permutations

External validation: 8 powders

Actual \ Predicted	Cohesive	Easy-flowing	Free-flowing
Cohesive	0	0	2
Easy-flowing	0	2	1
Free-flowing	0	0	3

Feature importance analysis



Conclusions:

1. The 118 materials analyzed exhibited a wider range of PSD, particle shape distributions, and bulk densities, and covered 3 classes of FFc (cohesive, easy-flowing, and free-flowing).
2. The best performing algorithm for Step 1 achieved a performance of 0.835, and for Step 2, 0.84.
3. The external validation of the classification models showed that 5/8 were correctly classified.
4. Implementing machine learning models in the early stages of drug development could help determine suitable manufacturing strategies for a given material, providing a rapid digital screening tool for advanced pharmaceutical development.

References

1. M. Capece, K. R. Silva, D. Sunkara, J. Strong and P. Gao, Int J Pharm, 2016, 511, 178-189.
2. H. Barjat, S. Checkley, T. Chitu, N. Dawson, A. Farshchi, A. Ferreira, J. Gamble, M. Leane, A. Mitchell, C. Morris, K. Pitt, R. Storey, F. Tahir and M. Tobyn, Journal of pharmaceutical innovation, 2021, 16, 181-196
3. 3. M. Leane, K. Pitt, G. Reynolds and M. C. S. W. Group, Pharmaceutical development and technology, 2015, 20, 12-21.