Article:
Barjat, H, Checkley, S, Chitu, T et al. (11 more authors) (2021) Demonstration of the Feasibility of Predicting the Flow of Pharmaceutically Relevant Powders from Particle and Bulk Physical Properties. Journal of Pharmaceutical Innovation, 16 (1). pp. 181-196. ISSN 1872-5120

https://doi.org/10.1007/s12247-020-09433-5

1 Abstract

Purpose: Understanding and predicting the flow of bulk pharmaceutical materials could be key in enabling pharmaceutical manufacturing by continuous direct compression (CDC). This study examines whether, by taking powder and bulk measurements, and using statistical modelling, it would be possible the flow of a range of materials likely to be used in CDC.

Methods: More than 100 materials were selected for study, from four pharmaceutical companies. Particle properties were measured by static image analysis, powder surface area and surface energy techniques, and flow by shear cell measurements. The data was then analysed and a range of statistical modelling techniques were used, to build predictive models for flow.

Results: Using the results from static image analysis a model could be built which allowed the prediction of likely flow in a shear cell, which can be related to performance in a CDC system. Only a small amount of powder was required for the image analysis. Surface area did not add to the precision of the model, and the available surface energy technique did not correlate with flow.

Conclusions: A small sample of powder can be examined by Static image analysis, and this data can be used to give an early read on likely flow of a material in a CDC system or other pharmaceutical process, allowing early intervention (if necessary) to improve the characteristics of a material, early in development.

Keywords: Powder Flow, static image analysis, continuous direct compression, modelling, shear cell, powder surface area.

2 Introduction

The flow of powders, active pharmaceutical ingredients (API's) and excipients, is a key determinant of their performance in the pharmaceutical manufacture of solid oral dosage forms.

Continuous direct compression[1-6] is a particularly attractive route of development for solid oral dosage forms as it reduces some aspects of scale up, can be implemented in a relatively low footprint environment and, from the initial development of the technology, has had Process Analytical Tools (PAT) and process control designed into the implementation, meaning that the processes should be inherently understandable and controllable[7,8]. Continuous processing is now an established manufacturing route for pharmaceuticals[9], with more products being developed and approved using this technology, and is thus viable as a key route for development of new products.

A key component for continuous direct compression processes are loss-in-weight (LIW) feeders. These devices hold and transport the individual components for a formulation, and transmit the material to the blending component of a continuous process. A key determinant of the ability of a LIW feeder's ability to control the delivery of the material is the flow of the powder, defined by its material properties[10].

Flow is a notoriously difficult property to measure (and predict) for some powders, particularly when the particles are very small and the size and shape distributions are uncontrolled, and their flow performance is governed not only by gravity but by other forces and interactions[11,12]. Nevertheless a guide for potential performance of powders in LIW feeders suggests that a Flow Function Coefficient (FFC), measured on a shear cell or equivalent[13-16], of greater than 3 is a good predictor of acceptable (i.e. predictable and reproducible gravimetric feeding) performance in a feeder[7]. Historically a FFC of 2-4 is defined as cohesive, and therefore poorly flowing, and whilst it is

clear that LIW feeders can cope with some materials in this class, the most cohesive can be troublesome. It is also important to note that there is not a simple cut-off in performance at three, and that materials close to this FFC are likely to be difficult, regardless of whether they fall over or under the nominal FFC limit of 3. Recent data has suggested that, in some cases, materials with a FFC<3 do not perform in all circumstances in LIW feeder systems[17]. In some cases, for powders with unacceptable flow and performance, it has been necessary to pre-blend drug powders with flow aids to achieve the necessary performance in the continuous system, which adds a batch process (the pre-blending) to the nominally continuous one.

The flow of powders, and cohesive forces which can affect flow, has been demonstrated to be influenced by a range of factors, including but not limited to, particle size[18], particle shape, particle size distribution[19], surface area, surface energy and electrostatics[20]. Capturing all of these factors can be difficult, but multivariate and other modelling approaches are used to understand better the factors, and their interrelationships.[17,18] It is recognised that engineering of materials to ensure performance in LIW feeders is a possible, and perhaps necessary, adjunct to the successful use of such manufacturing approaches.[1]

The composition (material) and geometry of a feeding system also makes a significant difference to the flow that a material achieves through the system, as can other perturbations such as when the feeding system is topped up[21,22]. The amount of material that is being refilled into a hopper, the ability of the system to adjust feed rate and controller parameters and algorithms to control these factors are all-important in ensuring reproducible and predictable flow into the mixer part of the CDC system, and thus the production of a robust blended system.

Even though they are inherently compact and efficient systems, trials of LIW feeders in conjunction with formulations requires significantly more material than is often available during the initial phase of development, and that predicting the flow properties to determine whether a

material is *likely* to have the correct properties would be beneficial. One means to achieve this is to use surrogate materials for this purpose. The general methodology for identifying surrogate materials has been outlined[23], as well as its specific use in predicting flow in continuous systems[24]. However, in the published cases the surrogate materials are more similar to larger, relatively free flowing materials, rather than the more difficult materials representative of API's.

A potential block to the more widespread adoption of continuous direct compression in pharmaceutical development is a lack of knowledge, early in a development programme, of the likely performance of a new material in a putative continuous process. If the performance of the API is an unknown, or cannot readily be inferred or predicted, it is possible that a "conservative" choice of formulation route, albeit one which is ultimately less efficient, is chosen. This decision, made in early development in order to mitigate against the unknown, may persist long after improvements are made in the properties of the API.

If one assumes that most companies utilise a "platform" formulation for their continuous processing activities, i.e. use known excipients of low variability, the only variable that needs to be determined is the likely performance of the API. It would be valuable to know how an API is likely to perform in an LIW feeder, and whether the API could be engineered to meet necessary characteristics. It can be seen from recent data[25] that materials have a range of properties, which may necessitate exceptional formulation efforts, but that a decision on formulation may be made early in the development process, before all information is known.

The aim of this project is to extend the window of predictability of flow for materials outside the range of published information, to micronized systems and for materials with other adverse properties, as demonstrated by poor performance in LIW systems. The intention was to predict, when relatively little material is available (e.g. when manufacturing scales are at the 100g stage of development, at which time it is not feasible to commit **all** of the available material to a flow test),

whether a material is *likely* to have acceptable properties in the LIW feeder, or whether further intervention and engineering is necessary to make the API acceptable for processing by CDC.

It would be anticipated that excipients, at least those intended to aid flow and provide other bulk properties, would have better flow properties than those of the API's which they are "carrying". However, there appears to be a floor value, of an FFC around 3, which all materials must meet to ensure that flow is sufficiently predictable for consistent use in CDC systems.

The model would also be able to predict the properties of *feasible but unmanufactured* materials, i.e. materials that could be made using available technologies (size, shape and their respective distributions) but only exist *in silico* for the moment, to indicate which would have the most acceptable properties. The aim here would be identify particle properties *in silico,* which, if manufactured, would have acceptable properties. This would help direct engineering efforts.

In many cases API's, as initially tested, will have unacceptable characteristics, but capturing this in an objective manner will help frame the discussion on likely and successful interventions. In addition, the ability to develop a model, which will predict those particle properties that are likely to achieve sufficient flow, is a key goal for early stage examination of materials. This would give a target for Chemical Engineers and Chemists with an additional *in silico* experimental tool for translating putative materials from bench to manufacturing scales, to aim at during development of an API for this process.

Statistical models predicting flow from particle properties have been published previously[26,27]. Initial systems using this data set were built as data became available, using a data analysis workflow enabling data to be integrated into the prediction as new materials where made available by the pharmaceutical consortia. Whilst it was possible that additional compound analysis techniques could have been developed for this work, as a first pass it was agreed that if the modelling approach could achieve a predictive capability *across the wide range of particle*

properties, that was better than currently available models (which tend to have been built from materials with relatively "good" flow properties)[28] then the additional techniques would potentially not add additional meaningful granularity.

Although models for the prediction of flow in a range of pharmaceutical systems have begun to emerge[28] the data sets used in these previous studies were either smaller than the one envisaged for this programme and (or) did not cover the full range of powder flow characteristics observed in pharmaceutical oral solid dosage forms. Also those models may not have incorporated size and shape from robust instrumentation, capable of capturing size and shape, and their relevant distributions, which can be key in understanding material performance.[29]

A consortium of companies and a specialist data centre conducted this work as part of the Advanced Digital Design of Pharmaceutical Therapeutics (ADDoPT) programme.

3 Selecting Materials and Techniques for the Study

The study was conducted by a team from the ADDoPT Consortium, consisting of members from four multinational drug companies, "Primes", experienced in continuous manufacturing (AstraZeneca, Bristol-Myers Squibb, GlaxoSmithKline and Pfizer), experts in the control and understanding of continuous processing systems (Perceptive Engineering), and the Science and Technology Facilities Council's Hartree Centre, providing expertise in data science.

The techniques chosen for characterisation of the bulk powders were agreed between the four Pharmaceutical Companies, and the organization doing the data analysis, based on the following agreed principles:

1. The techniques had to be commonly available, and in regular use, at all of the companies, i.e. were a regular part of their workflow;
2. There had to be some evidence that the techniques should be predictive of flow characteristics and/or a good surrogate for performance in an LIW feeder (recent developments in this area have been reported[30]);
3. The data, raw or summarised, had to be capable of being outputted in a machine-readable format suitable for use in statistical methods such as machine learning.

Following extensive discussion, it was agreed that the techniques meeting these criteria were:

1. Bulk flow properties: Data generated from the Shear Cell, using an agreed protocol (outlined below) was chosen as the mechanism to generate FFC data[16]. Shear cells, which have a fundamental underpinning in the physics and understanding of flow, are widely used in the Pharmaceutical Industry to understand relevant flow measurements;

2. Particle characterisation by Static Image Analysis: Image analysis has the ability to generate both size and shape information from a wide range of relevant powders, even for small (micronized) particles, and the data can also capture the presence of agglomerated particles and their impact on the particle size distribution[31];
3. Surface area from Brunauer-Emmett-Teller (BET) theory. BET measurements[11,32]; thought to be an important factor in predicting material flow properties and performance[25] and is widely measured on robust apparatus. A subset of the data was tested for this parameter;
4. Surface energy measurements. Considered a factor in particle-particle interactions. Surface energy measurements via surface energy analysis (SEA) was generated[23, 26], as this has been shown, in some cases, to provide some small advances in the predictive capability of models in this area. A subset of the data was tested for this parameter.

Specific apparatus was not prescribed for the shear cell measurements or BET measurements, subject to the use of standard equipment for this purpose, as these measurements were felt to be robust across a range of apparatus.

For particle size and shape by static image analysis the Morphologi G3 instrument was chosen[29,31,33,34]. This instrument can accommodate a wide range of particle sizes and shapes relevant to pharmaceutical processing[27], from micronized materials to granules and agglomerates, and also captures shape factors. Protocols for generating the data were agreed and updated, along with those for data analysis.

The selection of materials for the testing was made by the Pharmaceutical primes, from their portfolio of batches, but also including excipients and model drugs with specific properties. The aim was to represent the widest range of materials likely to be incorporated into formulations for oral solid dosage forms, in particular for CDC. Materials for inclusion were intended to encompass the

range of materials that each of the Pharmaceutical companies envisaged for possible use in continuous direct compression. Active API's were selected to have a broad range of properties, aligned with what was expected from CDC systems. Unmilled, milled, and micronized powders were included, as were a range of particle morphologies, including needles. The aim was to have materials both above and below the FFC of 3, which has been shown elsewhere to be a key parameter in determining acceptable performance in LIW feeders[35]. In addition, excipient materials were also chosen and were all expected to have higher FFC's than 3, as demonstrated in previous studies. It is to be expected that bulk excipients (i.e. those added to improve bulk properties rather than perform another function such as lubricity) will have good flow, and this would drive their choice in CDC. Table *1* gives a breakdown of the materials tested.

A subset of samples, in this case 50, were tested by SEA and BET. For this work, the SEA system instrumentation from Surface Measurement Systems was utilised, as per previous publications.[36,37] These samples represented the tested samples from one of the Primes. The samples had properties representative of the spectrum of materials tested overall.

A statistical modelling approach, outlined below, was used to generate predictions of flow. In general, an R^2>0.80 for model predictions was thought to be a good first step towards an adequate model. This level of prediction has been possible for narrow ranges of materials (e.g. free flowing materials) but has not been published for a wider range of particle size and shape.

4 Method Protocols and Data Analysis

4.1 Materials included in the study

Materials, API's and excipients, anticipated to have flow properties across the spectrum of what could be accepted to feed into a LIW feeder in a CDC system, were included in the study. A significant number of materials Micronized and other milled materials were included. Materials

within the study had been demonstrated, in some cases, to have acceptable or unacceptable properties in trials.

The intention of the study was to capture in the models those materials with clearly acceptable properties of flow, but also those whose characteristics would currently preclude them for use, without pre-processing, in a CDC formulation approach using currently available LIW feeder systems in CDC.

4.2 Morphologi G3 measurements - image based particle characterisation

Particle size and shape analysis for each batch was determined using a Malvern Morphologi G3 particle characterisation system (Malvern Panalytical, Malvern, UK). Powder analysis requires less than 500mg of material per test, but sampling characteristics and particle counts for effective understanding of distributions can be assessed[33]. The use of this technique is aligned with proposals for improving the understanding of particle performance, and product critical quality attributes (CQA's), in developing pharmaceuticals.[38]

Sample dispersion and imaging approaches were aligned to the nature of the material being analysed, as published in previous work[29,39-42]. In order to remove partially imaged and overlapping particles morphological filtering of the raw image data was conducted using a combination of convexity and solidity filters, according to previous studies[29,33,38,39,43].

Where samples were observed to contain aggregates (loosely associated particle assemblages) and/or agglomerates (fused particle assemblages), size distributions for both aggregated/agglomerated and primary particles were obtained. For such samples, shape data was only reported for the primary particle dataset.

4.3 Surface Area Measurements

The samples were analysed using a Gemini 2390A surface area analyser (Micromeritics, Norcross, USA) or equivalent, with nitrogen as the probe gas. Samples were typically out-gassed for 12 hours at 50°C under nitrogen gas prior to analysis. Samples were then evacuated at a rate of 500 mmHg/min for 5 minutes and equilibrated for 5 minutes. Multipoint measurements (8 points) over the range of 0.05-0.3 P/P_0 were performed, and linearity within the B.E.T. range confirmed.

4.4 Shear Cell Measurements

4.4.1 Instrument

The specific instrument used by the consortium was not stated, but an example instrument was the Ring Shear Tester (RST-XS, Dietmar Schulze, Wolfenbüttel, Germany). In each case, the included software was used to determine specific terms for further analysis. It was not possible to control completely the humidity in all of the laboratories, so this factor was not directly controlled in the testing. Relative humidities (RH) in the labs ranged from 40% RH-60% RH.

4.4.2 Bulk density

The reported bulk density was determined from the shear cell tester. This value may not be consistent with other measures of bulk density[44], but was consistently used in this study.

4.4.3 Wall Friction Testing

Shear cell conditions were set, but the specific instrument was not controlled in the study. A Ring Shear Tester (for example, RST-XS, Dietmar Schulze, Wolfenbüttel, Germany) was used for measuring the wall friction properties of the powders using a XS-Wm standard cell. Spacers were used to leave a 4 mm recess at the top of the cell to fill a layer of the test powder. A disk of 0.4 RA

316 stainless steel was used as the top disk to provide the test surface against which the test was performed. The consolidation load used was 4000 Pa and the test loads were 400, 800, 1200, 2400, 3200 and 4000 Pa.

The shear cell is rotated while the lid is fixed in position by the tie rods and wall shear stress increases (θ,w). This decreases over time until a constant wall stress is achieved. The values at *wall normal stress* (σ,w) and *constant wall shear stress* (θ,w) describe the kinematic wall friction. The process begins with the greatest wall normal stress and once steady-state is reached, the normal load is reduced while the shear cell continues rotating. With each decrease in wall normal stress, the wall shear stress also decreases. The pairs of values of wall normal stress and wall shear stress were plotted to give the Wall Yield Locus.

4.5 SEA Measurements

50 samples were tested with the SEA apparatus, using a range of dispersive probes; n-decane, n-nonane, n-octane, n-heptane, and n-hexane were injected at a range of 10 fractional surface coverages between 0.2% and 20%. Determination of the concentration free dispersive surface energy (i.e. the dispersive surface energy at 0% probe coverage) was calculated using a line of best fit through the data points. Specific (polar) probes, ethyl acetate and chloroform were also analysed at an equivalent range of fractional surface coverages, as used for the dispersive probes. The column dead time was determined using an inert probe, methane (0.208 cm^3 injection volume). The dispersive component was obtained using the Dorris and Gray approach and the polar components obtained using the polarization approach [37,45,46]. For the purposes of comparison, the values utilised were nominally the 5% surface coverage results although for dispersive surface energy results for 3% and 10% were additionally included.

4.6 Data Analysis Principles

Yu *et al*[19] used partial least squares methods to predict the flow of common pharmaceutical excipients based on particle size data from laser diffraction and particle imaging data from dynamic image analysis. In that study only one sample had an *FFC* of less than 3. In another study Sandler *et al*[26] also used particle imaging data from dynamic image analysis of granular materials to predict flowability amongst other things. All materials in that study had *FFC* values of 6.9 or higher, representing easy or free flowing materials. Both these studies used particle size (and shape where applicable) distributions as input to the modelling.

The cumulative distribution function (e.g. of particle length) is a high dimensional object, not adequately summarized by a few numbers. This study sought to use the full distributions, rather than point measurements such as D_{50} or D_{90}, as input to the predictions, rather than digesting them to a small-dimensional set of descriptors.

Support vector regression is a standard machine learning method. In least squares regression the investigator uses prior knowledge to select features that are suspected to be predictive. In support vector regression, the investigator provides instead a "kernel function" that indicates how similar two inputs are

Kernel approaches are suitable for this purpose. A suitable kernel function captures the notion of “similar” samples in a precise function:

$$k(d, d') \Longrightarrow R \quad \textit{Equation 1}$$

A kernel function shows large values for similar samples, and small values for samples which are different. It therefore increases as a metric (distance) function decreases. Given a metric function, this most commonly used kernel function, and the one used here, is the Radial Basis Function

$$k(d, d') \implies exp\left(-\frac{\|d-d\prime\|^2}{2\alpha^2}\right) \quad Equation\ 2$$

Where the parameter α is large if the property to be predicted is very local in distribution space, and smaller correlations persist for larger differences.

To use this approach, we used a function that represents the distance between distributions, meeting the mathematical definition of a metric, including obeying the triangle inequality. The method chosen for this study was the Wasserstein Distance[47], also known as Earth Movers' distance[48] There are several ways to measure of the difference between two distributions. Our approach required one that obeys the triangle inequality, to be suitable for forming a kernel function for a support vector machine, which excludes Kullback-Leiber diverence. The Wasserstein (Kantorovich–Rubinstein) metric is informally known as the Earth Mover's Distance: if the probability density function were the profile of a sand pile, it is a measure of how much sand has to be moved to transform one into the other. For the current purposes a sample with 90% fines and 10% of particles of radius 0.1mm is similar to one with 10% of particles of radius 0.11µm, more so than it is similar to one with 10% of size 0.12µm. The Wasserstein metric captures this, in way that the alternative L1 or L2 metrics do not

Given a kernel function, there are varieties of statistical learning methods available. Of all the kernel methods available, Support Vector Machines are especially suitable when the amount of training data is low, because they are well regularized and their regularization is well understood[49].

Support vector regression (SVR) is used with the data sets' particle size (or particle size and shape) similarity matrix as input[50] and hyper-parameters ν, and *C*.

All hyper-parameters (α, ν, *C* and, if applicable, d_s) are optimised during model training using leave-one-out cross-validation. When the powder sample left out is one of the six samples with duplicate particle size and shape measurements, both of these are left out of the training set and two predicted flowability measures are computed, one for each measured particle size (or particle

size and shape) distribution. Least absolute deviation between measured and predicted values is used as the criterion for model optimisation to limit the sensitivity to potential outliers (points with large residuals).

For each SVR system with optimised hyper-parameters, the predicted values were fitted against the measured ones using linear regression and the R^2 value is reported.

In order to assess whether additional bulk powder properties, such as specific surface area and properties measured by surface energy analyser, might help explain part of the residuals from the models obtained, linear regression of the models' residuals against those measured properties were performed. R^2 is reported.

An analysis of variance model is then performed on the models obtained to assess the effect of the following experimental design factors on the R^2 values obtained:

1. number of particle size bins *n*,
2. particle size only or particle size and shape model,
3. and flowability at low-σ_{pre} , mid-σ_{pre}, high-σ_{pre}, low σ_1 or mid σ_1.

The principle of binning materials into classes is captured in figure 1.

Tukey's Honest Significant Differences were computed between the levels of factors that have a significant effect.

5 Summary of data generated

The data set consisted of 106 fully characterised materials, in terms of static image analysis measurements and associated shear cell data. 50 samples, from one Prime, were additionally tested for surface area (BET) and by surface energy analysis (SEA) according to the protocol. The modelling

methods were validated by leave-one-out cross-validation, which is appropriate for a data set of this size. As larger data sets become available it may be possible to come up with cross validation protocols leaving out more data, however this approach is robust for the data set in this case.

Initial analysis confirmed that these materials had the full spectrum of flow characteristics of materials expected to be included in solid oral dosage forms manufactured by continuous direct compression. FFC values range from 1 to 22 across the range of shear cell measurements (table 2), a range of surface areas from 0.05 to 10.23m^2/g were also observed. Median particle sizes from 2.4 µm to 443 µm were seen, and a range of shapes of particles was also observed in the samples. Particle size distributions included narrow particle size distributions, wider distributions and some bimodal distributions with wide spans.

There were a significant number of batches with a measured *FFC* of less than 3, ensuring that models could be trained using the range of powder flow behaviours encountered in pharmaceutical development laboratories, i.e. materials above and below the proposed threshold for acceptable flow were included in the study.

6 Model Building and Outcomes

6.1 Particle Size and Shape Information

Individual particle size and shape summary measures were obtained from 100 powder samples. These measures were obtained in duplicate for six powder samples, resulting in 106 powder size and shape samples, and a total of 16.6 million particles in the pre-processed files obtained from the industrial partners. All particles are included in the size and shape summaries described below and during the modelling.

Volume based distributions have previously been shown to be better predictors of flow, as the proportion of large, free flowing, particles within a system is an important determinant of flow.[34]

In the presence of fines, the distributions of particule size, mass, and surface area can be very different: 99% of particles might account for half the surface area but only 10% of the mass. In the light of this, we defined the buckets for our histograms by particle mass.

In order to summarise particle size information for all powder samples, the range of sphere equivalent volumes across all samples was calculated, and from that the particles were binned into n = 30, 40 or 50 sphere equivalent volume bins S_i. The total volume V_1 of particles in each size bin S_i is then calculated for each powder sample. In addition, the median particle sphere equivalent volume across all samples is calculated per size bin.

The aspect ratio of a particle is defined as its width divided by its length. It is the inverse of particle elongation. For each particle size bin S_i, and each powder sample, the mean aspect ratio AR_i of the particles in that bin is calculated, yielding a maximum value of one if the particles have equal widths and lengths, or a lower value, decreasing with increasing elongation.

6.2 Comparing Morphologi G3 Data Sets

For each powder sample, the summary of size and shape information described above is contained in n pairs of values, representing a size and shape fingerprint (or signature) for this sample. These values summarising the distributions of particle size and particle mean aspect ratio obtained are compared. These fingerprints can then be used to compare samples as described next.

6.2.1 Using Particle Volume Information

The Earth mover's distance (EMD)[48] was used to compare the fingerprint of each sample to every other sample in the data set.

For every bin S_i, the particle size bin index is used as the location of a particular data point in a one-dimensional space, and the total volume of particles in that bin V_i is used as the weight. The information for all bins is used to calculate the distance metric.

6.2.2 Using particle Volume and Aspect Ratio Information

In addition to particle volume information in the first dimension, a second dimension is added, a particle shape dimension representing the mean aspect ratio AR_i per size bin S_i. This results in each data set being represented by n points in a two-dimensional space. When compared to the sample summary obtained using particle size information only, adding particle aspect ratio information shifts each point at size location i to the value AR_i in the shape dimension. The total volume of particles per bin V_i remains the same. The EMD distance is calculated using points located in the 2-dimensional "particle size – particle mean aspect ratio" space.

Figure 1 illustrates the method for the hypothetical case of one powder sample where the particles are classified into only four size bins. The insets illustrate the sizes and shapes of the particles in these bins at the same hypothetical scale. In this case the fraction of the total powder volume that is made up of the smallest particles (left test tube) is the largest of the four fractions as shown by the amount of powder in this test tube. Particles in this fraction also have the lowest average aspect ratio (highest average elongation) as shown by the extreme position of the test tube along the aspect ratio axis. The test tube representing the next particle size class (second from the left) contains a lower volume of powder, and is made up of nearly spherical particles so it is located at the other extreme of the aspect ratio axis. In the next two size classes, the powder volume again differs, with that in the fourth class similar to that in the second class, and the average aspect ratio decreases from one class to the next. This powder "fingerprint" in two dimensions, size and aspect ratio, is used as the basis for comparing assessing similarities between different powder samples during modelling

6.3 Powder flow Measurements

The flowability of each powder sample was measured by ring shear cell at a range of consolidation stresses σ_1, measured at three normal stresses during pre-shear σ_{pre} (at or near 1.0, 3.0 and 7.0 KPa). In particular, some of the measurements were concentrated around the two σ_1 values 2.3 (low σ_1, 2.3 ± 0.3KPa, measured at σ_{pre} = 1.0 KPa) and 6.0 KPa (mid σ_1, 6.0 ± 0.7KPa, measured at σ_{pre} = 3.0 KPa). From the measured flowability, an angle θ is calculated, such that:

$\theta = \arctan(1/FFC)$.

Predicted *FFC* and θ values at these two consolidation stresses and three pre-shear stresses are reported here, giving a total of five summary flowability measures per powder sample. GNU parallel[51] software was used to parallelise the computations across the summary flowability measures and the different number of sphere equivalent volume bins n. The hypothesis that the distributions of *FFC* and θ values across powder samples for each summary flowability measure differ from a uniform distribution is tested using a one sample Kolmogorov-Smirnov test, which is widely used for this purpose[52].

6.4 Powder Surface Area Measurements

The specific surface area (SSA) of 50 of the powder samples was measured by nitrogen adsorption using the isotherm from the Brunauer–Emmett–Teller (BET) theory.

6.5 Powder Surface Energy Measurements

The same 50 samples were also analysed using a surface energy analyser. The parameters measured were dispersive surface energy at 0, 3, 5 and 10% coverage, specific surface energy at 5% coverage and total energy, as published in previous work[36,37,45,53].

6.6 Modelling

A Gaussian radial basis function was used to convert the distance metrics obtained from particle size (and optionally shape) descriptors into similarity (kernel) metrics, with hyper-parameter α determining the length scale.

For models that use particle size and aspect ratio information, a scaling hyper-parameter *ds* is used that determines the ratio of the range in the particle size dimension to the range in the particle aspect ratio dimension.

6.6.1 Regression Models

Support vector regression (SVR) is used with the data sets' particle size (or particle size and shape, along with their respective distributions) similarity matrix as input[50] and hyper-parameters ν, and *C*.

6.6.2 Classification Models

Classification models for each summary flowability measure were built using particle size information only and particle size and shape information, as described for regression models in the previous section. The models were built using support vector classification to predict which of the two classes, *FFC* <= 3 or *FFC* > 3, each powder sample belongs to. The performance measure used to train the model is the area under the ROC curve.

7 Results

7.1 Distribution of Particle Volumes

The wide range of samples in the study is illustrated in figure 2. This figure illustrates that, across the sample population of 16 million particle, a wide range of particle properties was

observed. Of course for bulk properties such as flow it is the mixture of particle sizes in an individual sample which determines the flow for that material. Between the 106 samples more than 16 million particles with reported size and shape information were available for analysis. The bulk of most samples is made up of particles with volumes in the range 20 to 10^7 µm^3 but this encompasses the range of particle size of materials envisaged for use in CDC. A few extreme partial volumes are seen outside of that range, particularly at the high end of particle volume.

This covers the range of likely materials in pharmaceutical development for solid oral dosage forms. Figure 3 also shows that a wide range of morphologies and size and shape distributions was captured by the study.

7.2 Distributions of particle size and shape

Figure 4 shows particle volumes V_i as a function of particle size bins S_i and particle mean aspect ratio AR_i for 16 powder samples. The samples shown are arranged in decreasing order of flowability measured at low pre-shear stress, from the best to the least flowing sample (from top left to bottom right). Each sample is coloured according to its flow class. The area of each circle is proportional to the partial volume of powder in the particular particle size class. In general, it can be seen that, as flowability decreases, the bulk of the powder volume shifts from being made up of rather large particles (say 1.5e+04 µm^3 and larger) to being made up of small to medium particles (say less than 1.5e+04 and as low as 1e-02 µm^3).

7.3 Distribution of flowability and theta values

Table 2 shows the distributions of available *FFC* values and angles θ for the five shear cell experimental settings, as well as the outcome of the test of whether each distribution is significantly different from a uniform distribution. The distributions of *FFC* values and angles θ measured at the three pre-shear stresses σ_{pre} = 1, 3 and 7KPa is shown in figure 5. The dashed vertical lines show the median values. For all pre-shear stress values, the distribution of *FFC* values is skewed towards low

values. In contrast, none of the distributions of θ is significantly different from the uniform distribution as shown in the last column of table 2. The median and mean *FFC* observed is higher in the low σ_{pre} (σ_1) experiment than in both mid and high σ_{pre} (mid σ_1) experiments, showing that powder flowability tends to be lower when measured at low σ_{pre} (σ_1) in our datasets.

7.4 SVR modelling

The model learnt by the SVM is a linear model in a high dimensional space, so is equivalent to a linear regression with a custom set of descriptors. However, these descriptors need never be calculated to use the SVM. Given this hidden linearity, a transformation of the value to be predicted sometimes helps build a better model, e.g. predicting the logarithm if the process is fundamentally multiplicative. In this case, better models were achieved by predicting the arc tangent of the FFC, which we have referred to as θ. Without this transformation, the training process would be dominated by trying to fit the numerically large but less relevant FFC values (as the performance of a free flowing excipient is less vital than that of an API): i.e. predicting accurately whether the flow function coefficient is 2 or 4 is more important than predicting whether it is 5 or 10.

When a shape parameter was included in the model, the predictive accuracy always improved, as can be seen in Table *3*

A support vector regression model was built using flowability measured at mid σ_{pre}, and particle surface area in μm^2 (as reported by the Morphologi G3 instrument and instead of sphere equivalent volume) and aspect ratio summarised in $n = 40$ bins. This model yields $R^2 = 0.82$, as shown figure 5. This demonstrates that the approach is compatible with building a robust model across a range of particle properties.

7.5 Correlation with other measured bulk powder properties

For a subset (n=50) of the data surface area was carried out in addition to the particle size and shape data. The R^2 values obtained from fitting the residuals of the SVR models to bulk powder properties measured by BET and SEA are shown in Table 4. There is no correlation between the residuals and the specific surface area measured. Surface area does correlate with flow throughout the model, but in this case, it does not add to the predictive nature of the model, as the information provided by the surface area is captured in the detailed descriptors of size and shape. For the subset of data tested, the surface area data alone did not provide a model with similar predictive capabilities as the size and shape data.

No parameters from the SEA, for this data set, correlated with flow or provided any predictive additions to the models. The strongest correlation between the residuals and properties measured by surface energy analysis is obtained for the dispersive surface energy at 5% coverage, for the flowability model at low σ_{pre} (R^2=0.130). The SEA values were only tested on 50 samples, around half of the dataset. It is possible that inclusion of the other 50 samples could have improved these correlations better but this additional analysis was not conducted, as the initial testing did not indicate an effect.

In general, these parameters do not add significantly to the power of the model. Whilst it is possible that there are subsets of the data that would benefit from greater granularity (e.g. the surface area of finer systems), overall the targets for applicability of the model were met without including this extra data. Further examination of these parameters may yield benefits in later iterations of the model but they are not examined further in this paper.

Surface energy may from a key part in determining the flow of powders so it is not clear from this data whether this assumption is not appropriate or whether an alternative measure of surface energy is required.

7.6 Effects of experimental design

The outcome of the analysis of variance are summarised below, together with Tukey's honest significant differences where experimental factors are significant:

1. *n* (the number of particle size bins) does not significantly affect model quality, whether all models are considered, or only those that make use of the particle size and shape information (p = 0.777 and p = 0.840 respectively).
2. The choice of particle descriptors used to train the models (particle size only or particle size and shape information) significantly affects model quality (p = 5.3e-04), difference in R^2 = 0.0346, 95% confidence interval [0.018; 0.055].
3. The ring shear cell settings used to measure flowability significantly affects model quality (p = 3.92e-06).

7.7 Additional Modelling

In addition to the SVR model, a support vector classification (SVC) model was trained alongside the SVR prediction to produce a binary classifier for predicting FFC < 3, potentially facilitating the application of our approach in rapid identification and screening of powders with potentially problematic manufacturing properties. The SVC model was built using particle volume and aspect ratio information with n = 40 particle size bins and flowability measured at mid σ_{pre}, using leave-one-out cross-validation.

Table *5* shows the outcome of the classification models obtained for the various flow measurement settings. Once again, the worst models are those corresponding to low pre-shear stress.

8 Summary and Conclusions

The models developed and utilised here expand the predictive capabilities of flow, relevant to LIW feeders, across the range of materials likely to be processed by such systems. Inclusion of the model in the early stage of assessment of the suitability of a material for development by continuous direct compression. It also allows the input of criteria, in the form of potential (but feasible) particle distributions likely to achieve such flow.

As can be seen the predictions have considerable value, giving prediction of likely flow properties across a wide range of materials (in a pharmaceutical context), with reasonable precision. This was achieved despite the work being done across multiple laboratories with some conditions (e.g. the humidity of the flow experiments) not being controlled to an extent that may be possible in future. It is, of course, possible that future predictions, and more controlled models, could introduce some greater accuracy. However, this work suggests that particle size and shape, along with their respective distributions, are sufficient as a first pass to predict flow to a degree not achieved previously.

It can be seen that different particle morphologies make a key contribution to observed flow, and that large needles (for instance) are detrimental to flow in a way that might be predicted but has not been systematically observed. The data also captures the fact that materials have a particle shape distribution, e.g. low aspect ratio at small particle sizes, but more needle like at larger sizes. Such a distribution may occur in an incompletely attrited system.

Similarly, a particle size *distribution* around a nominal D_{50} or similar value can make a key contribution to whether that material flows or not. A bimodal distribution which has a nominally similar D_{50} to another sample with a unimodal distribution, and have very different flow properties, by measurement and thus as predicted by the modelling. The presence, or otherwise, of

agglomerates which facilitate the improved flow for bimodal distributions, may not have been captured by previous studies, as they may have been removed during processes such as dry dispersion for laser light scattering. It is already known that, for some excipients, that the presence of a small proportion of a larger particles in amongst finer particles is a key enabler of flow[43], and this appears to hold for agglomerates in this case.

Previous work indicated that particle size distribution and particle shape were key enablers of flow[19], allowing reasonable prediction based on particle properties, but mainly for relatively free flowing systems[54]. This work confirms that hypothesis with a wider range of materials, across a wide range of flow conditions. A predictive power of around 0.8 is similar to that seen for previous models, and may be inherent in a test such as shear testing, but here we have expanded that capability to relatively poor flowing materials. The presence of agglomerates or other flow enabling constructs (which may not be detected in point measurements such as D_{50}) may be key to understanding flow. It has previously been demonstrated that the presence of a small number or larger, disruptor, particles may explain the improved flow performance of grades of microcrystalline cellulose.[43] One of the merits of static based image analysis is that an appropriate method can be developed which does not break agglomerates or primary particles (which can occur in inadequately controlled dispersion techniques or be utilised specifically to break particles[55]), and that the post processing of data can include or exclude agglomerates if necessary[31]. The issue of the presence, or otherwise, of agglomerates in samples merits significant further investigation. Agglomerates can be induced in formulations, sometimes deliberately to enhance flow (e.g. in inhaled systems such as the Turbohaler), and it is possible that manipulation of agglomeration of poorly flowing materials could help induce flow for other pharmaceutical powders.

The ability of the model to predict whether a material is likely to meet the flow criteria for a material with acceptable performance in LIW could be a key step in rationally choosing materials that are suitable for this form of development. It is important to note that the test of the system for

"true/false" answers set a nominal target of ±3, so that a measured value of 2.9 which was predicted at 3.1 would not be regarded as an accurate prediction. In practice, as there is not an absolute cut-off, the model predicts that the system would be on the cusp of poor performance. In any such case, a material would benefit from intervention and possible modification.

As it stands the model has only been developed for single materials, rather than blends. Whilst it may be suitable for some blends and granules, this has not yet been examined. It is also necessary to note that if a flow agent was added to a material (e.g. 0.1% colloidal silicon dioxide) it is unlikely that this would be captured by the particle size and shape data (although surface area would be perturbed in this case), even by techniques such as Morphologi G3, as the silicon dioxide is likely to adhere to the surface of particles. Thus, the model is meant to identify single materials, which may require enablement rather than be universal for flow prediction, at this stage.

Regarding model accuracy, the final R^2 value of ~0.8 obtained from this work was considered fit for purpose and meets the aims of the study, which was to demonstrate the applicability of the approach of using advanced analytical techniques and data analysis to predict flow. While higher R^2 values could likely be achieved by using data from an individual pharmaceutical company, the final model incorporates a range of particle sizes generated independently across 4 pharmaceutical companies, thereby reducing issues of bias and over-fitting to one particular class of compound, and presenting a robust model with the potential for use across the pharmaceutical industry and their portfolio of compounds. The reason for the differences between sites could also be investigated further. It is possible that a more rigorous and exhaustive protocol on testing could be generated and followed, as small differences between sites can be observed in the protocols.

The work presented here demonstrates that one form of data analysis allows a prediction to be made from static image analysis data, if appropriately sampled and analysed. It is, of course, possible that other data techniques, for instance multivariate analysis techniques, could also build

successful models from this data. The value of this technique in elucidating previously difficult to predict parameters is demonstrated, and will no doubt improve as alternate and improved physical testing techniques and modelling approaches are applied to this and other data sets.

Once a team knows which sizes, shapes and shape distributions are required to facilitate flow they can work with colleagues and partners to collaborate to develop suitable materials. This facilitates a discussion about whether particle engineering could be used to develop a suitable material. All of this can occur early in a development programme, to allow rational formulation choices.

The system here demonstrates the value of the overall approach. It would difficult to replicate this precise system, as many of the specific materials here cannot be shared, as they are proprietary.

9 Conflicts of interest

The authors declare no conflicts of interest.

10 Figures and figure Captions

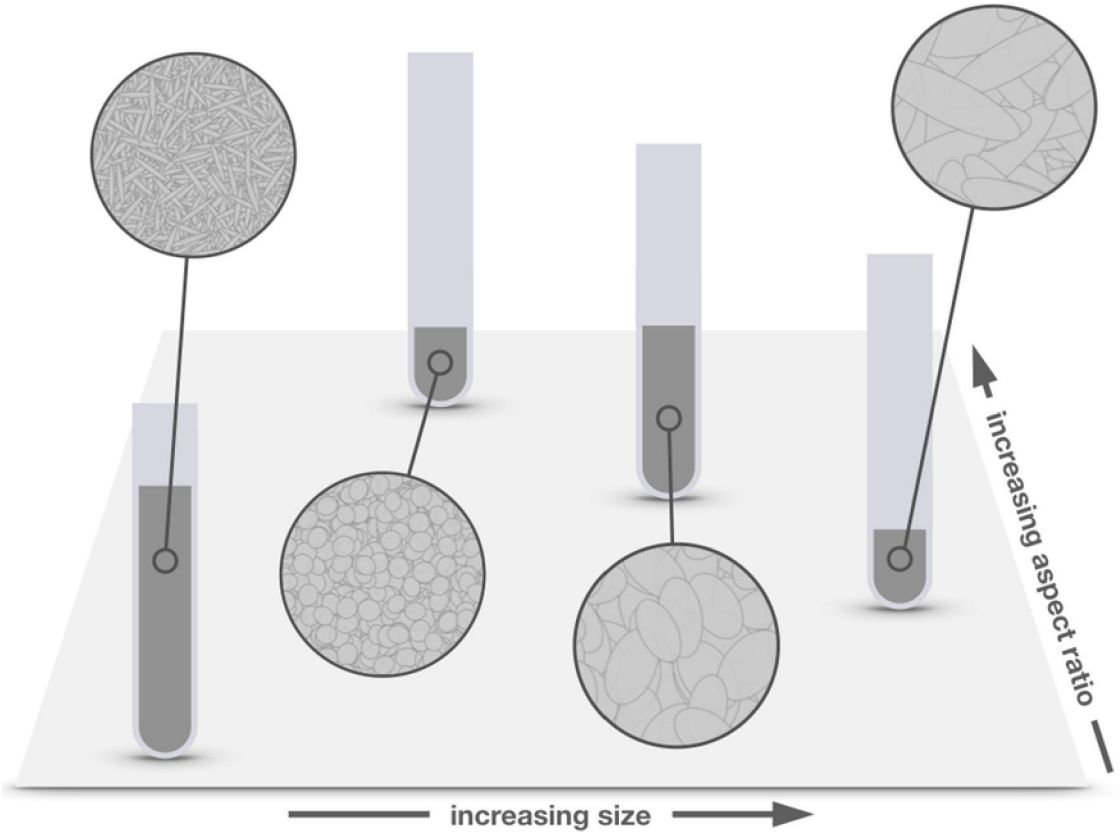

Figure 1: Principles of binning data for size and shape to generate powder fingerprint. Schematic rather than specific description, so scale bars not included.

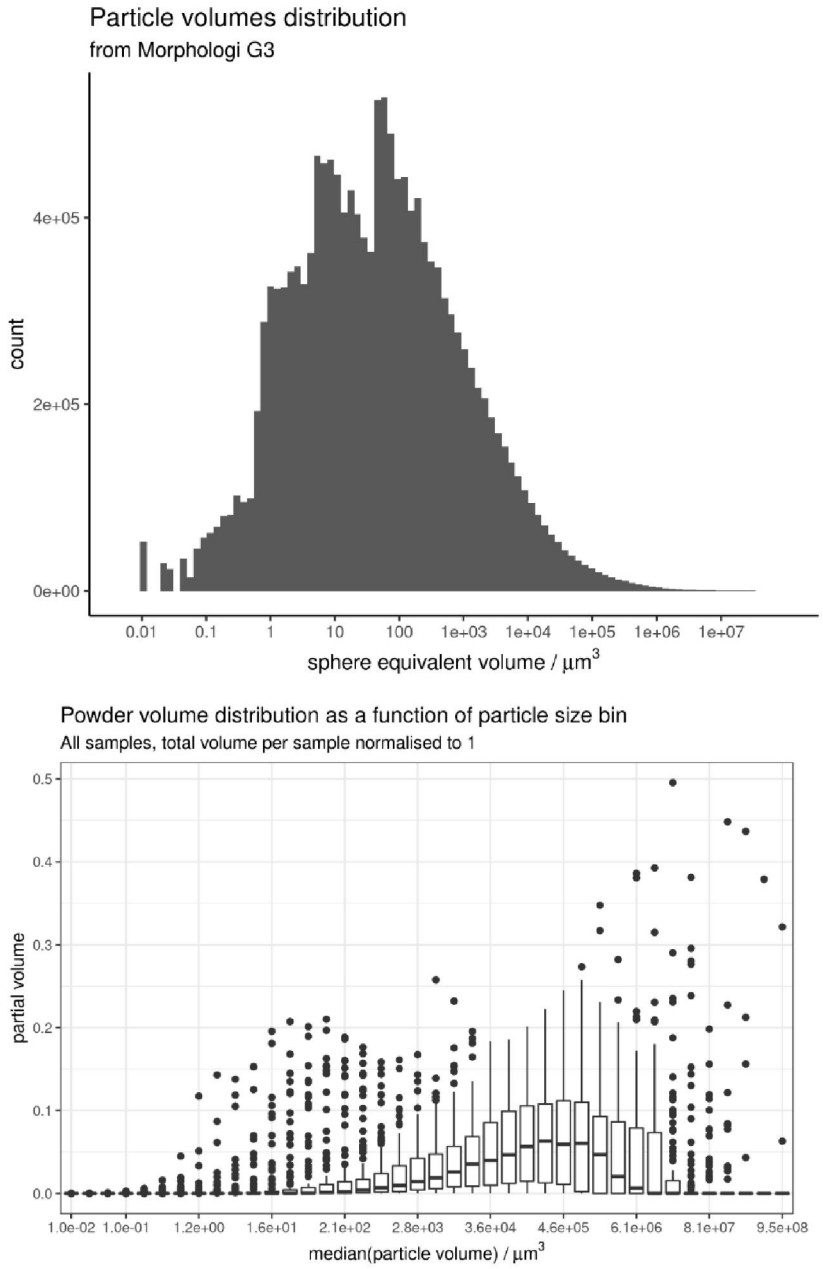

Figure 2: a) Distribution of measured estimated sphere-equivalent volume for all samples examined. b) Summary of partial volumes of powders for all samples as a function of particle volume range.

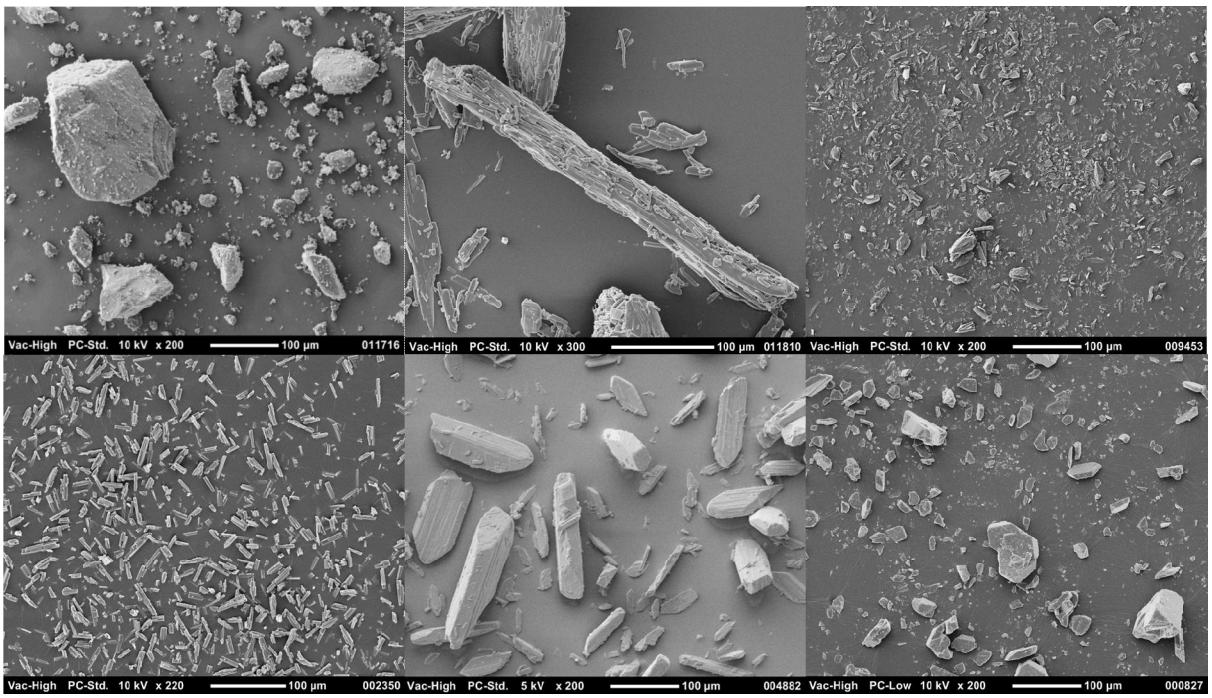

Figure 3: Sample Photomicrographs showing range of materials tested in this study, indicating the representative range of samples in the study.

Figure 4: Partitioning of total powder volume as a function of particle volume and mean aspect ratio for samples spanning the range of measured powder flows. These samples represent 16 samples from the 106 tested, the 16 samples being from across the spectrum of materials tested overall. Each graph corresponds to one powder sample, it shows 40 circles, one for each particle volume range (x-axis). The volume ranges correspond to increasing particle volumes from left to right. Each circle is located along y at the mean aspect ratio of all particles within the corresponding volume range and has an area that is proportional to the total volume of the particles in that range. The median particle volume is shown on the x-axis for a few regularly spaced volume ranges.

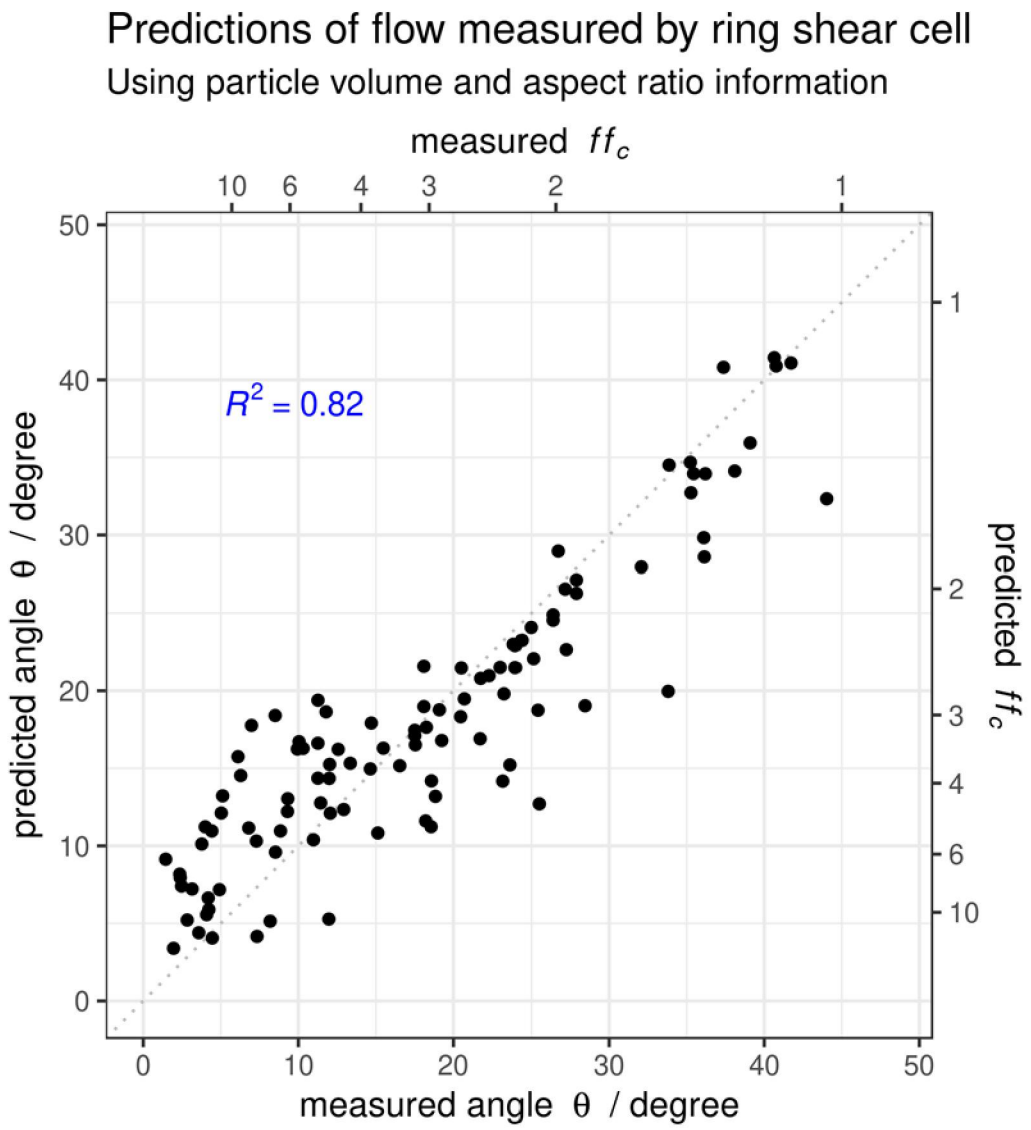

Figure 5: Outcome of support vector regression model to predict flow measured by shear cell using particles sphere-equivalent volumes and aspect ratios.

11 Tables and Table Captions

Tables

Nature of tested material	Number	Number of milled materials
API	58	20
Excipients	48	1
Total	106	21

Table 1 API's and excipients tested for flow in this study. Includes a number of duplicate samples

flow measure setting	**flowability summary**	**minimum**	**first quartile**	**median**	**mean**	**third quartile**	**maximum**	***p* value**
high σ_{pre}	Angle	3.58	11.26	20.51	21.07	32.08	44.03	0.220
low σ_1	Angle	4.00	15.33	23.26	24.49	34.07	43.57	0.753
low σ_{pre}	Angle	3.58	15.33	23.85	24.51	35.28	42.24	0.628
mid σ_1	Angle	2.59	11.18	20.68	21.16	32.08	42.83	0.543
mid σ_{pre}	Angle	2.59	11.26	20.68	21.16	32.08	42.83	0.543
high σ_{pre}	FFC	1.03	1.60	2.67	4.40	5.02	15.98	1.8E-13
low σ_1	FFC	1.05	1.48	2.33	3.14	3.65	14.32	1.3E-15
low σ_{pre}	FFC	1.10	1.41	2.26	3.31	3.65	15.98	1.3E-15
mid σ_1	FFC	1.08	1.60	2.65	4.52	5.06	22.07	1.3E-15
mid σ_{pre}	FFC	1.08	1.60	2.65	4.58	5.02	22.07	1.3E-15
		lowest angle					*highest angle*	
		2.59					44.03	
		lowest FFC					*highest FFC*	
		1.03					22.07	

Table 2: Summary of the shear cell experiment measurements. A wide range of flowability is observed, from very cohesive powders at FFC ≈ 1 to free flowing powders with FFC > 10, depending on the shear cell experiment. The last column "p value" shows the outcome of the Kolmogorov-Smirnov test of whether the distribution of FFC (angle) for a given shear cell experiment is significantly different from a uniform distribution as is the case for all FFC distributions ($p < 2e\text{-}13$) and none of the angle distributions ($p > 0.2$).

Experiment condition for *FFC* summary	**Particle information**	***R*^{2}**
high σ_{pre}	Volume	0.78
high σ_{pre}	volume + AR	0.82
low σ_1	Volume	0.65
low σ_1	volume + AR	0.67
low σ_{pre}	Volume	0.65
low σ_{pre}	volume + AR	0.69
mid σ_1	Volume	0.76
mid σ_1	volume + AR	0.81
mid σ_{pre}	Volume	0.77
mid σ_{pre}	volume + AR	0.80

Table 3. Summary of the model outcome for predicting flowability given particle size (and shape) information for the various ring shear cell FFC measurements. In all cases, the addition of shape information (mean aspect ratio) during model training improves the model, especially when its performance is lower with particle volume information only (when $R^2 < 0.8$).

flow measure settings	BET	Surface energy Analysis					
	log(SSA)	DSE at 0%	DSE at 3%	DSE at 5%	DSE at 10%	SSE	TE
low σ_{pre}	0.000	0.010	0.098	***0.130***	0.056	0.000	0.074
low σ_1	0.000	0.030	***0.103***	0.128	0.069	0.008	***0.094***
mid σ_{pre}	0.000	0.001	0.022	0.017	0.000	***0.028***	0.000
mid σ_1	0.000	***0.064***	0.084	0.128	***0.078***	0.003	0.086
high σ_{pre}	***0.003***	0.046	0.077	0.058	0.008	0.000	0.033

Table 4. R^2 values obtained by fitting the residuals of the five SVR models obtained from using particle size and shape information to other bulk powder properties measured by BET or using a surface energy analyser. SSA: specific surface area in m^2/g, DSE: dispersive surface energy in mJ/m^2 (DSE at 3%: dispersive surface energy at 3% coverage), SSE: specific surface energy, TE: total energy. The highest R^2 value for each bulk property (table column) is highlighted. There is no correlation between the models residuals and SSA measured by B.E.T. The strongest correlation with a measurement obtained from surface energy analysis is obtained for DSE at 5% for the model obtained from flow at low σ_{pre}.

flow measure settings	**number of samples with *FFC* ≤ 3 (out of 106)**	**performance**	**true negative**	**false negative**	**false positive**	**true positive**
low σ_{pre}	61	0.80	34	4	10	57
low σ_1	59	0.79	53	9	5	38
mid σ_{pre}	48	0.84	37	6	49	53
mid σ_1	47	0.83	53	10	10	33
high σ_{pre}	47	0.83	53	11	4	37

Table 5. Summary of the model outcome for predicting the powder flow class given particle size (and shape) information for the various ring shear cell FFC measurements. The models were trained to return true if FFC ≤ 3, false otherwise. False negative: predicted FFC > 3 but measured FFC ≤ 3. False positive: predicted FFC ≤ 3 but measured FFC > 3.

13. References

1. Chattoraj S, Sun CC 2018. Crystal and Particle Engineering Strategies for Improving Powder Compression and Flow Properties to Enable Continuous Tablet Manufacturing by Direct Compression. Journal of Pharmaceutical Sciences 107(4):968-974.

2. Galbraith SC, Liu H, Cha B, Park SY, Huang Z, Yoon S 2018. Modeling and simulation of continuous powder blending applied to a continuous direct compression process. Pharmaceutical Development and Technology:1-11.

3. Gouveia FF, Felizardo PM, Menezes JC. 2018. Chapter 14 - Lifecycle Management of PAT Procedures: Applications to Batch and Continuous Processes. Multivariate Analysis in the Pharmaceutical Industry, ed.: Academic Press. p 323-345.

4. Lakio S, Ervasti T, Tajarobi P, Wikström H, Fransson M, Karttunen AP, Ketolainen J, Folestad S, Abrahmsén-Alami S, Korhonen O 2017. Provoking an end-to-end continuous direct compression line with raw materials prone to segregation. European Journal of Pharmaceutical Sciences 109:514-524.

5. Moghtadernejad S, Escotet-Espinoza MS, Oka S, Singh R, Liu Z, Román-Ospino AD, Li T, Razavi S, Panikar S, Scicolone J, Callegari G, Hausner D, Muzzio F 2018. A Training on: Continuous Manufacturing (Direct Compaction) of Solid Dose Pharmaceutical Products. Journal of Pharmaceutical Innovation 13(2):155-187.

6. Roth WJ, Almaya A, Kramer TT, Hofer JD 2017. A Demonstration of Mixing Robustness in a Direct Compression Continuous Manufacturing Process. Journal of Pharmaceutical Sciences 106(5):1339-1346.

7. Alam MA, Shi Z, Drennen JK, Anderson CA 2017. In-line monitoring and optimization of powder flow in a simulated continuous process using transmission near infrared spectroscopy. International Journal of Pharmaceutics 526(1-2):199-208.

8. Hanson J 2018. Control of a system of loss-in-weight feeders for drug product continuous manufacturing. Powder Technology 331:236-243.

9. Burcham CL, Florence AJ, Johnson MD 2018. Continuous manufacturing in pharmaceutical process development and manufacturing. Annual Review of Chemical and Biomolecular Engineering 9:253-281.

10. Yadav IK, Holman J, Meehan E, Tahir F, Khoo J, Taylor J, Benedetti A, Aderinto O, Bajwa G 2019. Influence of material properties and equipment configuration on loss-in-weight feeder performance for drug product continuous manufacture. Powder Technology 348:126-137.

11. Shah UV, Wang Z, Olusanmi D, Narang AS, Hussain MA, Tobyn MJ, Heng JYY 2015. Effect of milling temperatures on surface area, surface energy and cohesion of pharmaceutical powders. International Journal of Pharmaceutics 495(1):234-240.

12. Leung LY, Mao C, Srivastava I, Du P, Yang CY 2017. Flow Function of Pharmaceutical Powders Is Predominantly Governed by Cohesion, Not by Friction Coefficients. Journal of Pharmaceutical Sciences 106(7):1865-1873.

13. Nakamura S, Otsuka N, Yoshino Y, Sakamoto T, Yuasa H 2016. Predicting the occurrence of sticking during tablet production by shear testing of a pharmaceutical powder. Chemical and Pharmaceutical Bulletin 64(5):512-516.

14. Swize T, Osei-Yeboah F, Peterson ML, Boulas P Impact of Shear History on Powder Flow Characterization Using a Ring Shear Tester. Journal of Pharmaceutical Sciences.

15. Wang Y, Snee RD, Meng W, Muzzio FJ 2016. Predicting flow behavior of pharmaceutical blends using shear cell methodology: A quality by design approach. Powder Technology 294(Supplement C):22-29.

16. Koynov S, Glasser B, Muzzio F 2015. Comparison of three rotational shear cell testers: Powder flowability and bulk density. Powder Technology 283:103-112.

17. Karttunen A-P, Wikström H, Tajarobi P, Fransson M, Sparén A, Marucci M, Ketolainen J, Folestad S, Korhonen O, Abrahmsén-Alami S 2019. Comparison between integrated continuous direct compression line and batch processing – The effect of raw material properties. European Journal of Pharmaceutical Sciences 133:40-53.

18. Fu X, Huck D, Makein L, Armstrong B, Willen U, Freeman T 2012. Effect of particle shape and size on flow properties of lactose powders. Particuology 10(2):203-208.

19. Yu W, Muteki K, Zhang L, Kim G 2011. Prediction of Bulk Powder Flow Performance Using Comprehensive Particle Size and Particle Shape Distributions. Journal of Pharmaceutical Sciences 100(1):284-293.

20. Samiei L, Kelly K, Taylor L, Forbes B, Collins E, Rowland M 2017. The influence of electrostatic properties on the punch sticking propensity of pharmaceutical blends. Powder Technology 305:509-517.

21. Engisch WE, Muzzio FJ 2015. Feedrate deviations caused by hopper refill of loss-in-weight feeders. Powder Technology 283:389-400.

22. Engisch WE, Muzzio FJ 2014. Loss-in-Weight Feeding Trials Case Study: Pharmaceutical Formulation. Journal of Pharmaceutical Innovation 10(1):56-75.

23. Ferreira AP, Olusanmi D, Sprockel O, Abebe A, Nikfar F, Tobyn M 2016. Use of similarity scoring in the development of oral solid dosage forms. International Journal of Pharmaceutics 514(2):335-340.

24. Escotet-Espinoza MS, Moghtadernejad S, Scicolone J, Wang Y, Pereira G, Schäfer E, Vigh T, Klingeleers D, Ierapetritou M, Muzzio FJ 2018. Using a material property library to find surrogate materials for pharmaceutical process development. Powder Technology 339:659-676.

25. Leane M, Pitt K, Reynolds GK, Dawson N, Ziegler I, Szepes A, Crean AM, Dall Agnol R 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. Pharmaceutical Development and Technology 23(10):964-977.

26. Sandler N, Wilson D 2010. Prediction of granule packing and flow behavior based on particle size and shape analysis. Journal of Pharmaceutical Sciences 99(2):958-968.

27. Wilson D, Bunker M, Milne D, Jawor-Baczynska A, Powell A, Blyth J, Streather D 2018. Particle engineering of needle shaped crystals by wet milling and temperature cycling: Optimisation for roller compaction. Powder Technology 339:641-650.

28. Hildebrandt C, Gopireddy SR, Fritsch AK, Profitlich T, Scherließ R, Urbanetz NA 2019. Evaluation and prediction of powder flowability in pharmaceutical tableting. Pharmaceutical Development and Technology 24(1):35-47.

29. Ferreira AP, Gamble JF, Leane MM, Park H, Olusanmi D, Tobyn M 2018. Enhanced Understanding of Pharmaceutical Materials Through Advanced Characterisation and Analysis. AAPS PharmSciTech 19(8):3462-3480.

30. Stauffer F, Vanhoorne V, Pilcer G, Chavez PF, Rome S, Schubert MA, Aerts L, Beer TD 2018. Raw material variability of an active pharmaceutical ingredient and its relevance for processability in secondary continuous pharmaceutical manufacturing. European Journal of Pharmaceutics and Biopharmaceutics 127:92-103.

31. Gamble JF, Tobyn M, Hamey R 2015. Application of image-based particle size and shape characterization systems in the development of small molecule pharmaceuticals. Journal of Pharmaceutical Sciences 104(5):1563-1574.

32. Shah UV, Olusanmi D, Narang AS, Hussain MA, Tobyn MJ, Hinder SJ, Heng JYY 2015. Decoupling the contribution of surface energy and surface area on the cohesion of pharmaceutical powders. Pharmaceutical Research 32(1):248-259.

33. Clarke J, Gamble JF, Jones JW, Tobyn M, Greenwood R, Ingram A 2019. Alternative approach for defining the particle population requirements for static image analysis based particle characterization methods. Advanced Powder Technology 30(5):920-929.

34. Gamble JF, Chiu W-S, Tobyn M 2011. Investigation into the impact of sub-populations of agglomerates on the particle size distribution and flow properties of conventional microcrystalline cellulose grades. Pharmaceutical development and technology 16(5):542-548.

35. Santos B, Carmo F, Schlindwein W, Muirhead G, Rodrigues C, Cabral L, Westrup J, Pitt K 2018. Pharmaceutical excipients properties and screw feeder performance in continuous processing lines: a Quality by Design (QbD) approach. Drug Development and Industrial Pharmacy:1-9.

36. Shah UV, Olusanmi D, Narang AS, Hussain MA, Gamble JF, Tobyn MJ, Heng JYY 2014. Effect of crystal habits on the surface energy and cohesion of crystalline powders. International Journal of Pharmaceutics 472(1-2):140-147.

37. Gamble JF, Leane M, Olusanmi D, Tobyn M, Šupuk E, Khoo J, Naderi M 2012. Surface energy analysis as a tool to probe the surface energy characteristics of micronized materials - A comparison with inverse gas chromatography. International Journal of Pharmaceutics 422(1-2):238-244.

38. Gamble JF, Dawson N, Murphy D, Theophilus A, Kippax P 2019. A Proposal for an Alternative Approach to Particle Size Method Development During Early-Stage Small Molecule Pharmaceutical Development. Journal of Pharmaceutical Sciences.

39. Gamble JF, Dennis AB, Hutchins P, Jones J, Musembi P, Tobyn M 2016. Determination of process variables affecting drug particle attrition withinmulti-component blends during powder feed transmission. Pharmaceutical Development and Technology:1-6.

40. Gamble JF, Tobyn M, Hamey R 2015. Application of Image-Based Particle Size and Shape Characterization Systems in the Development of Small Molecule Pharmaceuticals. Journal of Pharmaceutical Sciences.

41. Gamble JF, Ferreira AP, Tobyn M, DiMemmo L, Martin K, Mathias N, Schild R, Vig B, Baumann JM, Parks S, Ashton M 2014. Application of imaging based tools for the characterisation of hollow spray dried amorphous dispersion particles. International Journal of Pharmaceutics 465(1-2):210-217.

42. Gamble JF, Hoffmann M, Hughes H, Hutchins P, Tobyn M 2014. Monitoring process induced attrition of drug substance particles within formulated blends. International Journal of Pharmaceutics 470(1-2):77-87.

43. Gamble JF, Chiu WS, Tobyn M 2011. Investigation into the impact of sub-populations of agglomerates on the particle size distribution and flow properties of conventional microcrystalline cellulose grades. Pharmaceutical Development and Technology 16(5):542-548.

44. Hughes H, Leane MM, Tobyn M, Gamble JF, Munoz S, Musembi P 2014. Development of a Material Sparing Bulk Density Test Comparable to a Standard USP Method for Use in Early Development of API's. AAPS PharmSciTech 16(1):165-170.

45. Gamble JF, Davé RN, Kiang S, Leane MM, Tobyn M, Wang SSY 2013. Investigating the applicability of inverse gas chromatography to binary powdered systems: An application of surface heterogeneity profiles to understanding preferential probe-surface interactions. International Journal of Pharmaceutics 445(1-2):39-46.

46. Dorris GM, Gray DG 1980. Adsorption of n-alkanes at zero surface coverage on cellulose paper and wood fibers. Journal of Colloid and Interface Science 77(2):353-362.

47. Panaretos VM, Zemel Y 2019. Statistical aspects of wasserstein distances. Annual Review of Statistics and Its Application 6:405-431.

48. Rubner Y, Tomasi C, Guibas LJ 2000. The earth mover's distance as a metric for image retrieval. International journal of computer vision 40(2):99--121.

49. Vapnik V. 1998. Statistical learning theory. 1998. ed.: Wiley, New York.

50. Zeileis A, Hornik K, Smola A, Karatzoglou A 2004. kernlab-an S4 package for kernel methods in R. Journal of statistical software 11(9):1--20.

51. Tange O, others 2011. Gnu parallel-the command-line power tool. The USENIX Magazine 36(1):42--47.

52. Cyran KA, Kawulok J, Kawulok M, Stawarz M, Michalak M, Pietrowska M, Widłak P, Polańska J. 2013. Support vector machines in biomedical and biometrical applications. Smart Innovation, Systems and Technologies, ed. p 379-417.

53. Olusanmi D, Jayawickrama D, Bu D, McGeorge G, Sailes H, Kelleher J, Gamble JF, Shah UV, Tobyn M 2014. A control strategy for bioavailability enhancement by size reduction: Effect of micronization conditions on the bulk, surface and blending characteristics of an active pharmaceutical ingredient. Powder Technology 258:222-233.

54. Mullarney MP, Leyva N 2009. Modeling pharmaceutical powder-flow performance using particle-size distribution data. Pharmaceutical Technology 33(3):126-134.

55. Bonakdar T, Ali M, Dogbe S, Ghadiri M, Tinke A 2016. A method for grindability testing using the Scirocco disperser. International Journal of Pharmaceutics 501(1-2):65-74.