THE EFFECT OF SAMPLE SIZE ON RESULT ACCURACY USING STATIC IMAGE ANALYSIS

Optical microscopy has been a useful tool for particle characterization for many years. This is considered the referee technique since it is the most direct method available. One caveat is that the number of particles inspected (counted) must be sufficient so that there is confidence in the results generated. Several approaches have been suggested for determining the minimum number of particles counted in order to achieve desired confidence levels. This study reviews two standards that set the number of particles counted and compares these to experimental results.

Introduction

The number of particles counted is directly related to the width (breadth) of the distribution. If all of the particles in a sample are a single size, then only a single particle needs to be measured. If the sample has a broad distribution, then it is intuitively obvious that many particles should be measured. How many particles to inspect has been the subject of debate and may still be an unresolved issue. The number of particles to count is influenced by several factors including:

- Width of the distribution
- Desired confidence level
- Basis of the distribution (number vs. volume)
- Results of interest (mean, D10, D90, etc.)

In practice, the number of particles inspected is also influenced by whether the measurements are made manually or using automated image analysis software. Countless papers are published every year containing one view of an SEM image followed by conclusions regarding the particle size distribution and morphology of the sample. On the other end of the spectrum lie results generated by automated image analysis systems where thousands or hundreds of thousands of particle have been inspected.

USP<776>

The United States Pharmacopeia (USP) test USP <788> OPTICAL MICROSCOPY[1] was published prior to wide scale use of automated image analysis and generally refers to manual microscopy.

These standard states: “The number of particles characterized must be sufficient to ensure an acceptable level of uncertainty in the measured parameter”… “For a sufficiently large population (n > 30), the uncertainty in the estimate for the mean particle diameter, d, is given in the formula:”

$$d \pm zc\left(\frac{s}{\sqrt{n}}\right)$$

Where n = number of particle to count
d = mean particle diameter (number distribution)
z_c = desired confidence coefficient
s = standard deviation of the test specimen

According to USP <776> z_c is about 1.96 at the 95% confidence level for n>30. Using the equation shown above, an example calculation would suggest that for a sample with a standard deviation of 20 µm, 61 particles could be counted in order to achieve at uncertainty of 5 µm for the number mean.

Several observations on this approach could include questioning if n>30 constitutes “a sufficiently large population,” pointing out that volume distributions are preferred over number[2], and noting that the equation provides no guidance on other parts of the distribution (D10 and D90). The FDA expects D10, D50, and D90 values for active ingredient specifications, not just the mean diameter.

ISO 13322

Another approach to determining the necessary number of particles to observe for creating statistically valid particle size distributions is given in ISO 13322[3]. The following quote is taken from Section 1. Scope:

“Even though automation of the analysis is possible, this technique is basically limited to narrow size distributions of less than an order of magnitude. A standard deviation of 1,6 of a log-normal distribution corresponds to a distribution of less than 10:1 in size. Such a narrow distribution requires that over 6,000 particles be measured in order to obtain a repeatable volume-mean diameter. If reliable values are required for percentiles, e.g. *D*90 or other percentiles, at least 61,000 particles must be measured.”

All of the statements in the previous paragraph can be questioned. Image analysis is currently being used successfully to determine particle size (and shape) distributions much broader than one order of magnitude. Although the guidance given in ISO 13322 for the number of particles to be counted provides a theoretically correct and cautious approach, the experimental data shown in this technical note indicates that accurate results can be achieved even when inspecting far fewer particles.

The basis of the approach described in ISO 13322 was first published by Masuda and. Gotoh[4]. This theoretical approach assumes the particles are log-normally distributed, and suggests that the required number (N) of particles with a given error (σ) and a given confidence limit is estimated in accordance with the following equation:

$$\log N = -2 \log \sigma + K$$

where K is numerically determined by the confidence limit, particle distribution and other parameters.

Following the approach given in ISO 13322 tables are generated displaying the number of particles to be counted depending on the distribution width, confidence limit, and desired mean value. One of these tables is shown in Table 1 below.

Number of particles required n*, δ = 0.05, p = 0.95

δ	*GSD*	*n*(MMD)*	*n*(Sauter)*	*n*(MVD)*
0.05	1.1	585	389	131
	1.15	1460	934	294
	1.2	2939	1808	528
	1.25	5223	3103	843
	1.3	8526	4920	1247
	1.35	13059	7355	1750
	1.4	19026	10504	2363
	1.45	26617	14457	3093
	1.5	36007	19295	3956
	1.55	47358	25093	4952
	1.6	60811	31919	6092

Table 1: From ISO 13322

Interpreting Table 1 can be challenging. The basis of this table is to show the number of particles that need to be inspected in order to determine a median or mean diameter within a 5% error with 95% probability. The geometric standard deviation (GSD) of the distribution of the sample must be known prior to using this table. Unfortunately, few particle sizing techniques use GSD to express distribution width. Most techniques express distribution width using either the standard (arithmetic) deviation or the span (D90-D10/D50). The column n*(MMD) refers to the mass median diameter, the column n*(Sauter) refers to the surface median diameter, and the column n*(MVD) refers to the mean volume diameter.

If the three columns referred to volume, surface, and number distributions, the data would make more sense. It is difficult to understand why more particles need to be inspected in order to calculate the surface median than the volume mean diameter.

Experimental

It is also possible to take an experimental approach to the subject of how many particles to inspect to achieve desired accuracy. For this study the PSA300 static image analysis system (Figure 1) was used to measure particle size distributions of several samples. All samples were prepared using the powder Disperser Unit in order to separate agglomerates and create an even distribution of particles on the slide. Each sample was measured on the PSA300 by varying the number of fields inspected on the slide in order to inspect more or less particles. Plots showing the number of particles inspected vs. D10, D50, and D90 then provide insight into the effect of sample size. All plots were converted from a number to a volume basis.

Figure 1: The PSA300

The first sample studied was a polydisperse standard from Whitehouse Scientific, HS202. This sample has a known particle size distribution as measured on several techniques including microscopy. The certified values (volume basis) for this material are:

- D10 = 9.6 µm
- D50 = 13.4 µm
- D90 = 20.8 µm

Results from measuring this sample on the PSA300 are shown in Figure 2 and Table 2.

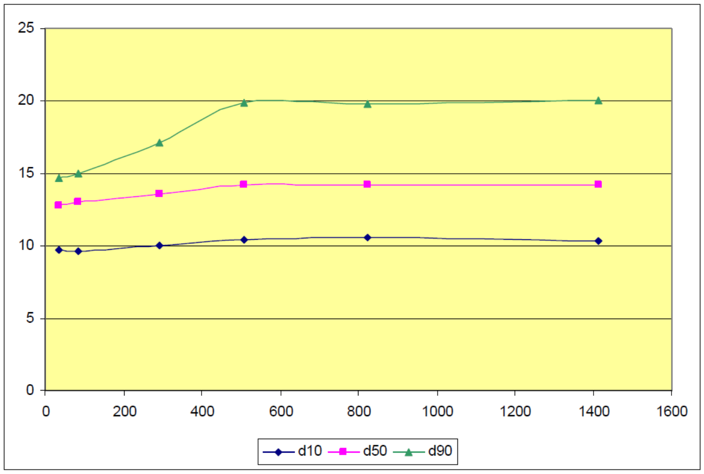

Figure 2: D10, D50, D90 (volume basis) vs. count for HS202

Count	D10	D50	D90
36	9.7	12.8	14.7
84	9.6	13	15
292	10	13.6	17.1
509	10.4	14.2	19.9
824	10.6	14.2	19.8
1412	10.3	14.2	20

Table 2: Data from Figure 2

The D90 plot is the easiest way to visualize the effect of sample size on results. For this sample, the D90 result reached the specification after ~500 particles were counted. The arithmetic standard deviation for this sample was determined to be ~ 4 µm. Using the GSD value of 1.6 from Table 1, ISO 13322 recommends counting at least 6092 particles to achieve a 5% at the "volume mean." The experimental data in Figure 3 and Table 3 indicated that only 500 particles need to be counted to achieve an error under 5% at the D90 from the volume distribution. The D50 value was within specification after inspecting only 300 particles.

An additional sample without a known distribution was measured to investigate the deviation between the theoretical requirements in ISO 13322 and observed experimental values. Figure 3 and Table 3 presents a similar plot of D10, D50, and D90 (volume basis) vs. number of particles inspected. The standard deviation for the sample was 27.

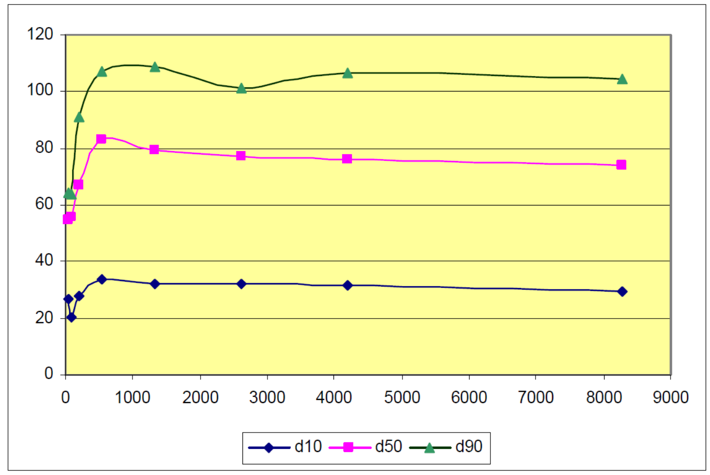

Figure 3: D10, D50, D90 (volume basis) vs. count for an unknown sample

Count	D10	D50	D90
47	26.7	54.8	64.5
82	20.3	55.8	64
207	27.8	66.8	91.3
534	33.5	82.8	107.3
1336	32.1	79.2	108.6
2627	32.3	77	101.4
4203	31.5	76.2	106.7
8270	29.7	73.8	104.2

Table 3: Data from Figure 3

The wavy nature of this plot suggests that the first data point is extremely inaccurate and that several large particles were inspected during the fourth and fifth measurement, creating the unusual appearance. Both of these points reinforce the need to inspect a large number of particles. Even with this broader distribution, the data for the D90 is acceptable after 4200 particles were inspected, far fewer than suggested by ISO 13322.

Conclusions

Image analysis is becoming a more accepted technique for particle size and shape analysis in many industries. The number of particles inspected can now be much higher with the implementation of automated image analysis. The equation given in USP <776> does not appear appropriate for most samples. A population of 30 to 60 particles is not supported either theoretically or experimentally. On the other hand, the theoretical values suggested by ISO 13322 appear significantly higher than data supported by experiments. The number of particles inspected for confident and accurate results may be closer to 6,000 for most samples rather than 60,000.

References

1. USP<788> Optical Microscopy, United States Pharmacopeia

2. Burgess, J., Duffy, E., Etzler, F., Hickey, A, Particle Size Analysis: AAPS Workshop Report, Cosponsored by the Food and Drug Administration and the United States Pharmacopeia, AAPS Journal 2004; 6 (3) Article 20 (https://doi.org/10.1208/aapsj060320)

3. ISO 13322-1, Particle Size Analysis – Image Analysis Methods – Part 1: Static Image Analysis Methods

4. Masuda, H. & Gotoh, K., Study on the sample size required for the estimation of mean particle diameter, Advanced Powder Technol., 10(2), 1999, pp. 159-173