GemmaCert

Cannabis Potency Variability and Cost-Effective Testing

A statistical analysis of potency variance and its potential solutions

Executive summary

Accurate drug labeling is critically important for patient and consumer health, and medical cannabis is no exception. While many of the jurisdictions regulating cannabis sales have begun to label the medicinally active ingredients of products, few have established methods for accurate and representative labeling. The problem of representative labeling is complicated by cannabis' wide-ranging potency. An individual batch of commercial cannabis will contain hundreds or thousands of flower specimens of varying potencies. One solution for a representative estimate is to test multiple samples and average the results.

However, current laboratory potency tests are lengthy and costly, making multiple tests impracticable. By contrast, spectral examination methods, though less accurate than the current industry standard, offer a quick and cost-effective way to leverage multiple tests. This white paper documents an extensive research study on the potency variance of cannabis batches legally acquired in Israel. It then deduces the number of spectral examination results required to achieve the accuracy of the industry standard method. In the end, spectral examination proves to be a far quicker technology capable of the same representativeness at a lower cost.

As cannabis and cannabis-based medicines gain international acceptance, myriad questions arise concerning their efficacy, dosing, delivery methods, and potency. Unfortunately, many of these questions have gone unanswered due to cannabis prohibition. Good Manufacturing Practices, as they apply to other pharmaceuticals, have not been uniformly enforced, if they have been enforced at all. The lack of pharmaceutical potency standardization, and the fledgling status of commercial cannabis analysis, should be of concern to prescribing doctors and medically fragile end users. In the foreseeable future, supply chain stakeholders may be liable for potency labeling errors.

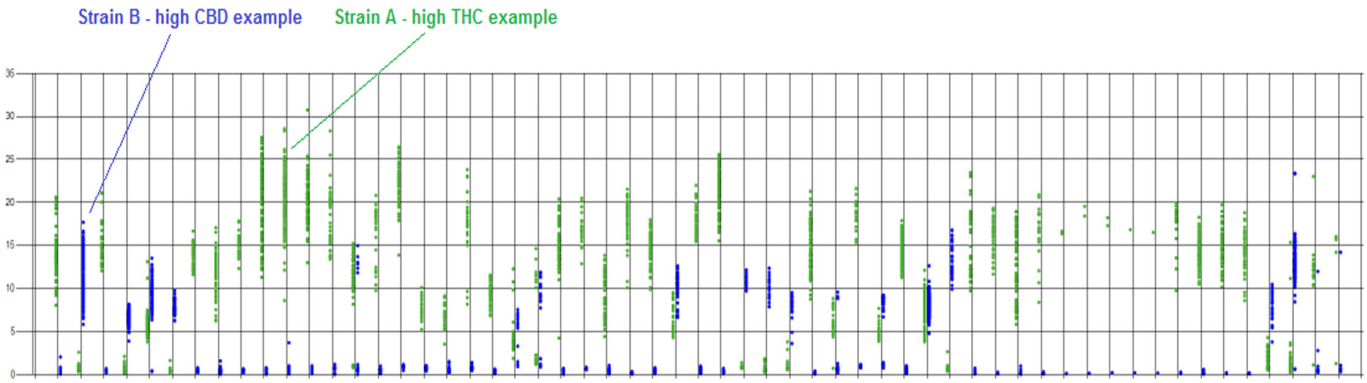
As cannabis moves from the black market into clinical acceptance, labeling will naturally be held to pharmaceutical standards. If pharmaceutical corporations were to sell products with active ingredient variance similar to cannabis crops, class-action legal suits would soon arise. Accurate potency labeling will be a priority for retailers, if only for the reason of self-preservation.

Cannabis labeling accuracy stems from the plant itself; cannabinoid content varies widely, even within a single commercial crop. Research has documented this variability,[1,2,3] but a casual survey of products for legal sale shows that cannabis doesn't lend itself to potency standardization. It's well known that potency differs between cannabis strains but, in fact, potency varies between plants of the same crop and even between flowers taken from the same plant.[4] So, the concentrations of the active ingredients of tetrahydrocannabinol (THC) and cannabidiol (CBD) in an individual commercial batch range widely.

For that reason, cannabis testing cannot be approached in the same way as a single-compound pharmaceutical, neatly contained in a tablet or intravenous compound. While traditional pharmaceuticals are carefully tested by standards established uniquely for each patented drug,[5] the ratios of THC, CBD and other cannabinoids continually shift expectations. Evidently, non-traditional and application-specific approaches to cannabis testing are needed.

Documenting the Scope of Strain Variation

Previous research has tracked the upward trend in the potency of cannabis confiscated during drug arrests but has not sufficiently documented the potency variance within each confiscated batch. Similarly, cannabis in legal markets manifests wide ranges in potency, even for the same strain as cultivated by different growers, but the scope of strain variation has not been scientifically documented. Without research, the issue of potency variation within crops is one that regulators and laboratories struggle to address. In response, the scientists at GemmaCert Ltd. set out to document potency variance more extensively than ever before. The goal was not only to show variance but to assess how alternate testing technologies can, and should, be applied to achieve more representative product labeling despite large population variance. Over 2,500 samples were legally acquired from medicinal producers in Israel. Most of the samples were cultivated in one season by one grower, which indicates a greater consistency of cultivation methods and product quality than might be expected on the market overall. Approximately 28 samples were selected from each of 54 batches, with each batch containing one strain from one crop. Each sample was homogenized by pulverization and potency tested for THC and CBD content per industry standard high-pressure liquid chromatography (HPLC). Figure 1 plots the results as cannabinoid content by percentage of sample weight. The horizontal axis presents strains with their names redacted to preserve grower confidentiality. THC is depicted in green; CBD is depicted in blue. The graph illustrates astounding variability, both between strains and between samples of the same strain.



Among these varieties, peak THC variance is 18 and the peak CBD variance is 11. That is, strain A exhibited a minimum value of just 9% THC and a maximum value of 27% THC. Likewise, strain B exhibited CBD content as low as 6% and as high as 17%. Yet these two strains are the extremes, with high variance caused perhaps by poor batch selection. Some strains, however, show potencies clustered within just a few percentage points.

If potency variation of cannabis is significant and can't be controlled, how can we more accurately label products?

Smaller batches could, in theory, be more representatively labeled. Batches collected from individual plants may show tighter variances than batches grouped from multiple plants. However, smaller batches naturally increase testing costs[6] because multiple tests are required to characterize the potency of any sized batch.

So, better potency characterization unavoidably requires more testing. An accepted way to characterize a highly variable population is to test repeatedly and average the results.[7] This "collate and average" approach would not eliminate the variance of the products on the market, but it would yield a number closer to the average value.

Why Test Repeatedly? An Example

To demonstrate the need for more tests per batch, we can explore a simplified example of a two-test HPLC technique for characterizing a 10-pound batch. One test returns a 14% THC result; the second reports 18% THC. Averaging the results might lead us to characterize the batch at 16% THC. Yet, if we performed more tests, the resulting mean value would change, possibly significantly. We will likely find that one of the original results deviates from the average more than the other. By chance, our results are badly skewed. If we continued our hypothetical testing, subsequent results could very well cluster around 18%, with many returning values above 18%. The average may even be greater than 18%. Our first assumption, based on an insufficient number of tests, proves inaccurate and fails to confidently characterize the batch.

Why Accuracy Fails

Traditional HPLC testing methods are exceptionally accurate when properly devised and executed. Any variation in HPLC results is more likely caused by improper sample preparation than the equipment. Yet, despite its accuracy, HPLC is an impractical tool for high-volume cannabis potency testing for the following reasons:

- HPLC testing requires 30-45 minutes per test
- HPLC must be operated by highly skilled and highly paid technicians
- HPLC requires large overhead expenses
- HPLC produces hazardous waste solvents and consumes disposable equipment
- HPLC destroys the sample

To summarize, HPLC is too resource-intensive to be acceptable for the seven, eight, or more tests necessary to ensure the representativeness of a single batch. Spectral examination, by contrast, is less accurate but also much less resource-intensive. Using the collate and average approach described above, spectral technologies like Near-Infrared Spectrometry (NIRS) can prove more representative. Despite its limited accuracy, NIRS offers the following benefits:

+ NIRS testing requires 60-120 seconds
+ NIRS equipment can be operated by moderately trained technicians
+ NIRS creates less overhead expense
+ NIRS produces no hazardous waste solvents and requires no ongoing expense budget for consumable lab supplies
+ NIRS leaves the high-value cannabis sample intact

However, the question remains: how many tests must be performed with the less-accurate NIRS equipment in order to eclipse the accuracy of one HPLC query?

Proving the Equitability of NIRS and HPLC

To answer the question posed above, the GemmaCert scientists analyzed the data they collected and determined the number of NIRS tests necessary to attain the representativeness of a given number of HPLC tests. This question must be considered in light of two variables. The first is the expected variability of the batch and the variance caused by the inaccuracy of the testing method. Only extensive research, like that performed by GemmaCert to aid in this calculation, can reveal the expected cannabis batch variability as a starting point. Batches with extremely wide variability may not be characterized at all, while hypothetical batches with no variability could be characterized with one HPLC test. The accuracy limitation of NIRS presents additional variance that must be considered as well.

The second factor in equating NIRS to HPLC is desired accuracy. The estimation tolerance, how far an acceptable result strays from the actual mean, affects how many NIRS tests are necessary to equal the representativeness of an HPLC test.

Our Statistical Analysis

Before beginning our calculations, we must define the terms of our inquiry:

M – number of specimens in a batch

N – number of samples for spectral examination

K – number of samples for HPLC analysis

μ – target attribute mean in this population

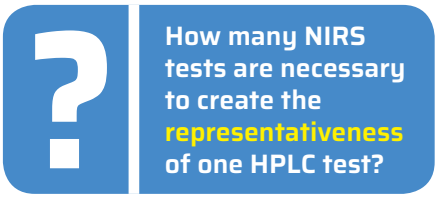
σ^2 – target attribute variance in this population

H_i – Individual HPLC analysis result for target attribute

$RMSE_H$ – HPLC result accuracy measure for target attribute

S_i – Individual Spectral examination result for target attribute

$RMSE_S$ – Spectral examination accuracy measure for target attribute



To equate spectral examination's cannabis batch testing ability to HPLC's, we must first define the results' distribution. That is, we must express the spread of the test results that can be expected given the variance of the batch and the accuracy of the test. For this exercise, we'll assume a normal distribution as represented by a classic bell curve.

The accuracy, or rather the inaccuracy, of our testing method creates variance in the results beyond the naturally occurring variance in the potency of the batch specimens. The batch specimens are the same for both testing methods, so our inquiry is focused on the variance caused by the testing method. These factors define our statistics:

HPLC: $H_i \sim D(\mu, \sigma^2 + RMSE_H^2)$ where "D" denotes distribution

Spectral: $S_i \sim D(\mu, \sigma^2 + RMSE_S^2)$

The root mean square error, or RMSE, will be higher for spectral examination than HPLC. It is the only factor distinguishing the equations above. In fact, we'll assume an RMSE of zero for HPLC because of its accuracy.

Integrating the number of test results into the equation above allows us to define the average variance for a single result. It sets the stage for direct comparison. The resulting modification shown below derives an average variance of an individual result for the given distribution.

HPLC: $1/K\ \Sigma H_i \sim D\ (\mu, (\sigma^2 + RMSE_H^2)/K)$

Spectral: $1/N\ \Sigma S_i \sim D\ (\mu, (\sigma^2 + RMSE_S^2)/N)$

With the average potency common to both distributions, and with the number of test results included, we can deduce the number of spectral test results necessary for accuracy equivalent to the given number of HPLC results.

$$(\sigma^2 + RMSE_H^2)\ /\ K \geq (\sigma^2 + RMSE_S^2)/N)$$

That is, if the above equation is true, the number of spectral results (N) has exceeded the accuracy of the number of HPLC results (K). To simplify, the equation can be restated as:

$$N \geq K\ (\sigma^2 + RMSE_S^2)/\ (\sigma^2 + RMSE_H^2)$$

We're going to assume HPLC is perfectly accurate. That means RMSEH2 is zero, and the equation is reduced to:

$$N \geq K\ (\sigma^2 + RMSE_S^2)/\ \sigma^2$$

The number of spectral results (N) necessary to keep the equation true depends on the population variance (σ^2) and the spectral examination error, ($RMSE_S^2$).

A Numerical Example

As an example, let's examine the results of the strain labeled "B" in the study above. It is a high-CBD strain with significant medicinal properties and an average CBD potency of 12.2% by weight. For the 166 samples of the strain that were analyzed, the variance (as documented by highly accurate HPLC testing) is as follows:

$\mu_{CBD} = 12.2$ (as expressed in percentage weight)

$\sigma_{CBD}^2 = 6.79$ (variance)

Having calculated the variance, we conservatively assume a spectral examination error of 1.5. We revisit the original formula with the numerical data to find that, for the given population, 1.34 spectral tests are necessary to match one HPLC result.

$$N_{CBD} \geq K\ (6.79 + 1.5^2)/\ 6.79 = 1.34\ K$$

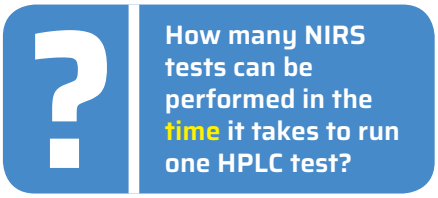
So, assuming a spectral examination error of 1.5, the accuracy of HPLC test can be attained with 1.34 times more spectral tests. In reality, one test is not sufficient to characterize a batch, neither by HPLC nor by spectral testing. As we'll see below, more than a dozen HPLC tests are necessary to characterize batch potency truly.

The quantity of tests needed to overcome cannabis' natural variance makes spectral examination all the more attractive.

Even assuming a much higher error for the spectral test, the method proves superior based on its resource efficiency and speed. Doubling the assumption of spectral RMSE still shows the value of NIRS, with 2.33 spectral tests equating to one HPLC test.

$$N_{CBD} \geq K\,(6.79 + 3^2)/\,6.79 = 2.33\,K$$

So, that is, 2.33 times more NIRS tests will match the potency testing accuracy of HPLC testing, even with an assumption of unrealistically poor accuracy.



Quantifying Confidence

How much estimation error are we willing to accept? Batch testing will always yield an estimate of the average. So, the satisfactory number of our tests will be determined by the degree to which we tolerate estimation error.

When we establish our estimation tolerance, we command a more relevant and thorough perspective on our method and its outcome. Accordingly, let us set an estimation tolerance of +/-10% and represent estimation tolerance as Δ.

Also, let us factor a 95% probability of our methods meeting the Δ standard above. It's worth noting that no testing protocol can achieve 100% certainty of meeting a given tolerance standard unless every sample in a batch is tested.

Following our earlier assumption of normal distribution, a 95% probability implies a result within two standard deviations. Consequently, stating Δ in terms of standard deviation of the average of the number of samples yields:

$$\text{Spectral:}\quad \Delta = 2\ ((\sigma^2 + RMSE_S^2)/N)$$
$$\text{HPLC:}\quad \Delta = 2\ ((\sigma^2 + RMSE_H^2)/K)$$

Otherwise shown as:

$$\text{Spectral:}\quad N = 4\,(\sigma^2 + RMSE_S^2)\,/\,\Delta^2$$
$$\text{HPLC:}\quad K = 4\,(\sigma^2 + RMSE_H^2)\,/\,\Delta^2$$

Our 10% estimation error tolerance, as it relates to the 12.2% CBD concentration, is 1.22:

$$\Delta_{CBD} = 12.2 \times 0.1 = 1.22$$

Finally, substituting with values stated above:

$$\text{Spectral:}\quad N_{CBD} = 4\,(6.79 + 1.5^2)\,/\,1.22^2 = 24 \text{ tests}$$
$$\text{HPLC:}\quad N_{CBD} = 4\,(6.79 + 0^2)\,/\,1.22^2 = 18 \text{ tests}$$

So, we've found that 24 NIRS tests will provide a result within a 10% margin of the actual average, as would 18 HPLC tests. The reason for the remarkable closeness in these numbers is batch variance. Batch variance represents most of the variance, far exceeding the variance created by the relative inaccuracy of NIRS.

In many instances batch variance will be much smaller, and that variance will likely decrease as protocols for batch selection improve. With these factors in mind, we recalculate the number of samples assuming a batch variance of 2. Accordingly:

Spectral: $N_{CBD} = 4\ (2 + 1.5^2) / 1.22^2 = 12$ tests

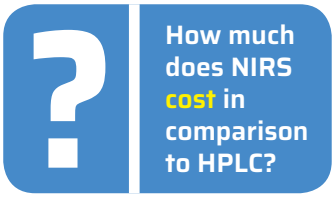
HPLC: $N_{CBD} = 4\ (2 + 0^2) / 1.22^2 = 6$ tests

Financial Considerations

Beyond the primary concern for patient safety, cost-effectiveness is paramount to cannabis potency testing. NIRS again proves advantageous. Though NIRS testing requires more tests to achieve representativeness, the cost of each test is far less for the reasons explained below.

NIRS tests may be performed by moderately skilled persons, while HPLC testing must be outsourced to a laboratory. In the chart below, a $20-per-hour wage is assumed for the NIRS operator who can perform each test in four minutes. The cost shown for an HPLC test reflects the national average of approximately $50. The difference is remarkable.

	Number of Tests	Price of Test	Total Price of Testing
NIRS	12	$1.33	$16
HPLC	6	$50	$300



In conclusion, spectral examination technologies can easily exceed the accuracy of high-pressure liquid chromatography for the potency testing of commercial cannabis. Though spectral examination methods like near-infrared spectrometry will never match the accuracy of a single HPLC test, the ability to quickly run multiple tests better addresses the high potency variance common to cannabis. NIRS may soon prove the industry standard for potency testing because of its lower labor costs, material costs, and higher speed. Furthermore, faster testing protocols can enable smaller batch sizes. The attentive selection of smaller batches based on the subjective qualities of the specimens will reduce variability and mitigate labeling inconsistencies.

Future research would do well to characterize the variance of cannabis further. With an accepted variance expectation, researchers will soon standardize the number of spectral tests necessary to consistently achieve representative results and acceptable labeling accuracy.

For more information on the cannabis testing applications of near-infrared spectrometry, contact GemmaCert Ltd.: info@gemmacert.com.

GemmaCert

Company Bio

GemmaCert is a biotechnology company, based in Israel since 2015, aiming to become a market leader in medicinal plant composition and potency analysis, starting with cannabis.
GemmaCert's skilled team of chemists, molecular biologists, biotechnologists, data scientists and programmers work tirelessly to advance cannabis analytical solutions.
In the long run, GemmaCert's breakthrough technology will enable patients and doctors to correlate cannabis composition with specific health conditions, significantly enhancing therapeutic treatment by cannabis and transforming the medical cannabis industry.

Address

Web: www.gemmacert.com

Trademarks and Copyright

GemmaCert is a Trademark of GemmaCert Ltd.

Disclaimer

Information in this document is subject to change without notice and does not represent a commitment on GemmaCert Ltd.
GemmaCert Ltd is not liable for errors contained in this document or for incidental or consequential damages in connection with furnishing or use of this material.
GemmaCert products are protected by U.S. and international copyright laws.

Sources

1. Zamengo, L.; Frison, G.; Bettin, C.; and Sciarrone, R. Variability of cannabis potency in the Venice area (Italy): a survey over the period 2010–2012. Drug Testing and Analysis 6, no. 1-2: 46-51. 2014.

2. Potter, D.; Clark, P.; and Brown, M. Potency of D9–THC and Other Cannabinoids in Cannabis in England in 2005: Implications for Psychoactivity and Pharmacology. J Forensic Sci. 2008.

3. Pijlman, F., Rigter, S., Hoek, J.; Goldschmidt, H.; and Niesink, R. Strong increase in total delta-THC in cannabis preparations sold in Dutch coffee shops. Addiction Biology 10.2: 171-180. 2005.

4. Namdar, D.; Mazuz, M.; Ion, A.; Koltai, H.; Variation in the compositions of cannabinoid and terpenoids in Cannabis sativa derived from inflorescence position along the stem and extraction methods. Industrial Crops and Products 113, 376-382. 2018.

5. United States Food and Drug Administration. Current Good Manufacturing Practices for Finished Pharmaceuticals. Title 21, Chapter I, Subchapter C, Part 211.

6. Sexton, M; Ziskind, J; Sampling Cannabis for Analytical Purposes. BOTEC Analysis Corp. I-502 Project #430-1e. 2013. https://lcb.wa.gov/publications/Marijuana/BOTEC%20reports/1e-Sampling-Lots-Final.pd

7. Gaines, P.; Accuracy, Precision, Mean and Standard Deviation. Inorganic Ventures ICP Operations Guide: Part 14. https://www.inorganicventures.com/accuracy-precision-mean-and-standard-deviation.

GemmaCert