Estimating concentrations and purities across several GS-441524 brands sourced from China

(Last Updated: July 2020)

Summary

We tested 15 different samples of GS including several popular brands in current use, several older brands suspected of having issues, and several very new brands that do not have substantial clinical testing. We compared each sample's concentration and relative purity against a reference sample of known concentration. Results suggest that some brands are likely to have higher concentrations than advertised, while other brands have concentrations that may not meet the marketed concentration. Without more rigorous testing across numerous batches for all brands, it is not possible at this time to completely evaluate how brands may change in concentration across batches or over time.

The following graph provides estimates for the average active compound (GS-441524) concentrations for each brand. The percentage above each bar shows how well each sample's chemical composition matches the chemical composition of the reference sample of GS we tested (relative purity).

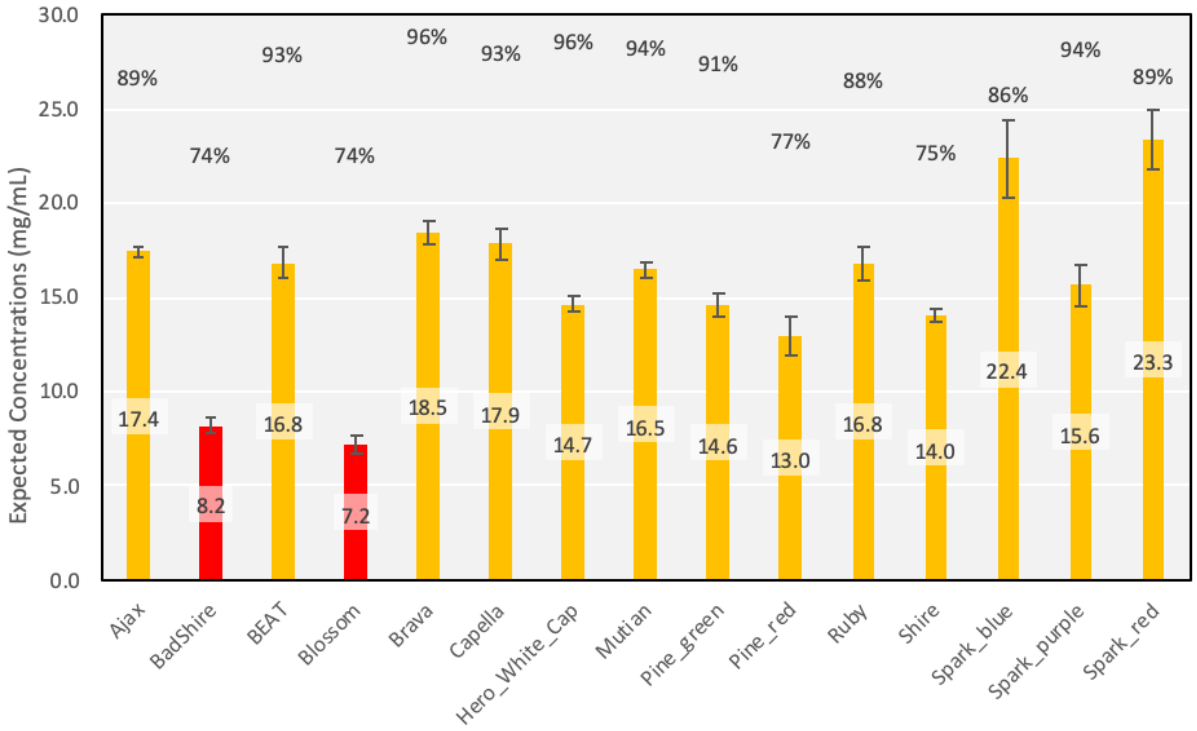

Figure 1. Estimated concentration for each sample. Error bars are 1 standard deviation. Purity estimates, relative to the reference sample, are listed above each bar as a percent estimate.

The following summary table is provided to help guide brand selection and dose estimation for individuals seeking to treat their cats using GS. Brands where multiple batches were tested appear to show higher variability than brands where only one batch was tested. This variability is consistent with concentrations differing slightly across vials. Samples that we analyzed using multiple vials are likely to

provide better estimates for concentrations that are representative of the average vial being used, compared to samples that were taken only from one vial for any given batch. Concentration estimates provided below show the 95% confidence interval for estimating the true concentration of a given sample. Brands with concentration estimates that fall outside of the marketed concentration at the low-end should be used with caution (increasing suggested doses), or avoided until additional testing can be completed. These test results should not be used as the sole determinant for choosing a GS brand. It is important to also consider the treatment outcomes for brands based on well-established clinical successes for a large number of cats both through the 84-day treatment and observation periods. There may be other factors, including sample contamination, that cannot be detected by the UV-vis methodology employed in this study.

Brand	***Measured Concentration (95% CI)***	***Marketed Concentration (mg/mL)***	***Likely to be in Spec?***	***# of Samples Tested***
Ajax	16.9 - 17.9	15	Yes	1
BadShire	7.4 - 8.9	15	No	1
BEAT	15.3 - 18.4	15	Yes	1
Blossom	6.3 - 8	15	No	1
Brava	17.3 - 19.6	15	Yes	1
Capella	16.3 - 19.4	15	Yes	2
Hero_White_Cap	13.8 - 15.5	16.5	No	1
Mutian	15.6 - 17.3	16.7	Maybe	1
Pine_green	13.4 - 15.8	15	Maybe	1
Pine_red	11 - 14.9	15	No	1
Ruby	15.1 - 18.5	15	Yes	1
Shire	13.4 - 14.6	15	No	1
Spark_blue	18.3 - 26.4	20	Maybe	1
Spark_purple	13.6 - 17.7	15	Maybe	3
Spark_red	20.3 - 26.4	17	Yes	1

Methodologies

We used a UV-Visible spectrophotometer to analyze each GS sample. This device utilizes ultraviolet light at various wavelengths and measures how much of the light is absorbed by the sample. We compared absorbance for each sample with absorbance for samples of known concentration that we prepared from GS powder. This allowed us to estimate the GS concentration and relative purity for each sample.

At the time of this writing, we found no previously published, publicly available standard UV-vis curves that we could use to assess the purity or concentration of GS-441524 samples. As such, it was necessary for us to create our own standards and then use these standards to evaluate the other samples. For this standard, we used GS powder from one of the most popular GS brands with a long term, proven clinical track record. It is clear, however, that this biases the results toward solutions coming from that brand.

We created our standard, stock solution by dissolving 1 mg of lyophilized GS-441524 in a 1 mL solution of 10% ethanol in water by volume at a pH of 1.5. We apportioned this stock solution to make 6 different known dilutions using the same 10% ethanol/water solution at a pH of 1.5. We made three replicates of each concentration to account for variability in samples and in measurements. Table 1 shows the resulting concentrations. We measured the absorption profile for each concentration and used the results to check instrument linearity near each concentration and create reference spectra for comparison with the various GS samples.

Table 1. Known concentrations of GS used to make reference curves

GS-441524 Concentration (mg/mL)
0.30
0.25
0.20
0.15
0.10
0.05

We tested 1 μL of each brand's GS-441524 sample at two different dilution levels: 1-to-60 and 1-to-120. We repeated the measurement for each dilution three times resulting in six measurements for each sample.

Additionally, we tested three "blanks". These "blanks" (from two different brands) contained only the solvent/diluent used by the respective brands to dissolve lyophilized GS-441524. We diluted the blanks in the same manner as the GS samples and ran them in the UV-Vis spectrophotometer to make sure that no inactive compounds contained in the diluent would impact measurements. We found that none of the blank samples had any significant, measurable absorbance. Table 2 shows a list of the samples we tested along with some notes describing the samples age or history:

Table 2. Data summary for the samples that were collected and tested for each brand

Brand	Comment
Ajax	Brand X from 2019. Became Ruby.
Bad Shire	Batch from January 2020
BEAT	Yellow-green cap, new in June 2020
Blossom	pH of 0.9
Brava	Batch from late April 2020
Brava blank	No acid
Brava blank	pH 1.5
Capella	Batch from June 2020
Capella	Batch from early May 2020
Hero	Batch from February 2020
Mutian II	New, full bottle from June 2020
Pine	With green polypeptide label
Pine	Red cap with loose retainer
Ruby	With yellow tinge, age unknown
Shire	Original with GS label from August 2019
Spark	Batch used in March 2020
Spark blank	No GS
Spark Blue	New blue lid version
Spark purple	Batch from February 2020
Spark purple	Batch from March 2020
Spark purple	New purple lid sent with blank
Spark Red	New red lid version

Results and Analysis

Figure 2 shows the absorption spectra for each standard concentration listed in Table 1. We found that each concentration has a well-defined absorption spectrum and good precision across the UV wavelength range commonly used to measure RNA.

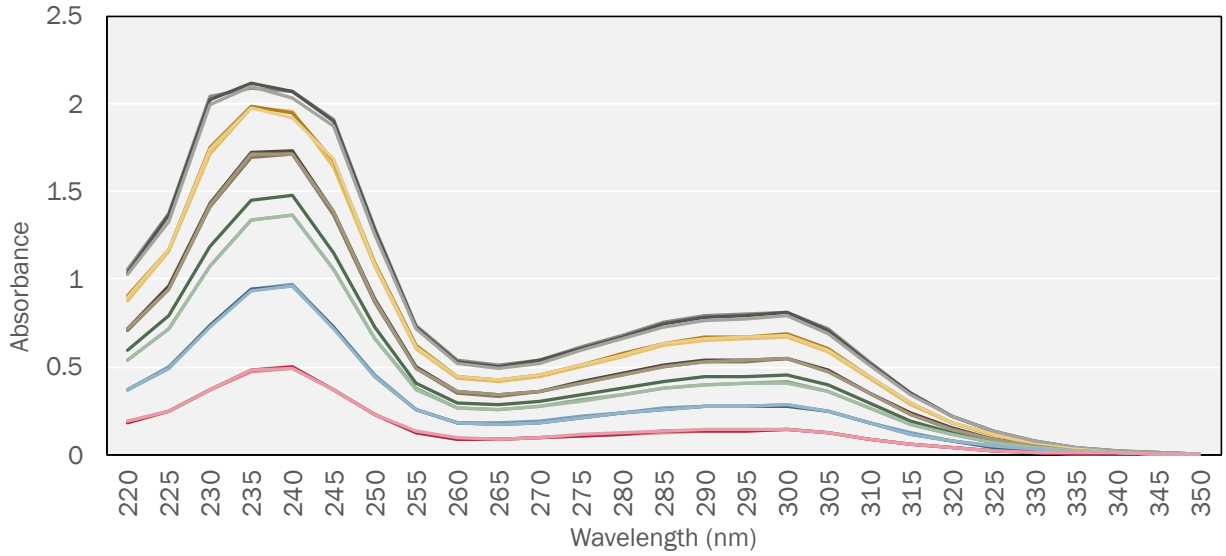

Figure 2. UV Absorption Spectra for GS-441524 standards from low to high: 0.05 mg/mL, 0.10 mg/mL, 0.15 mg/mL, 0.20 mg/mL, 0.25 mg/mL, and 0.30 mg/mL

Statistical Analysis of GS-441524 brands

Of the brands sampled, two brands had a history of suspected treatment failures. This includes the samples labeled "Bad Shire" and "Blossom" in our analysis. In 2019, numerous cat owners had been treating their cats successfully with "Shire." Shire was the lowest-cost brand on the market at one point, and therefore had high utilization rates relative to other brands. Starting in January of 2020, numerous cat owners that were using Shire began to report relapses and problems with their cats not showing improvements. The rate of relapse and symptom recurrence for these cats resulted in the FIP Warrior group no longer advising the use of Shire. A similar incident occurred for the brand known as "Blossom" earlier in 2019, where only a small number of cats using Blossom showed signs of recovery. It is possible that these brands either did not have sufficient GS in solution, or that the lyophilized GS dissolved in solution had a high level of impurity, resulting in ineffective treatment amounts and possible toxicity to cats.

Given this known negative result from these brands, we used a dummy variable to denote whether or not a sample was directly associated with problems (0), or had been effective at treating cats (1). We used linear regression to determine the light absorption wavelengths that were able to best predict negative or positive treatment outcomes for cats.

We employed Bayesian Information Criterion (BIC)[1] to identify the set of useful absorption wavelengths to make these predictions. We obtained absorption spectra for the samples at 5-nm intervals from 225 nm to 310 nm and added this data to an "over-fitted" linear regression model. Using the stepAIC package in R, we added or subtracted factors from the model, and we assessed the overall BIC for each new model formulation to determine if the improvement in fit made up for the loss in degrees of freedom for that model. [2]

The final results from this analysis indicate that absorption at the 225, 230, 240, 250, 255, 275, 300, and 305 nm wavelengths were best at predicting negative or positive treatment outcomes for cats given the samples that were tested and the defined "bad outcome" variable. The beta coefficients (expected effect values) from the regression results indicated that the 225, 230, 255, and 300 nm absorption profiles had a larger effect for negative outcomes, while the other wavelengths had a larger effect for positive treatment outcomes. This binary division served as a basis for creating two models to estimate the total concentration of GS in solution for each brand's sample.

Curve Fitting

We created two linear regression models to predict the concentration of GS found in each sample. Given the criteria discussed above, one model used wavelengths that were significantly associated with negative treatment outcomes, while another model used wavelengths significantly associated with positive treatment outcomes.

$$concentration \sim \beta_0 + \beta_1\lambda_{240} + \beta_2\lambda_{250} + \beta_3\lambda_{275} + \beta_4\lambda_{305} + \varepsilon_r \qquad \text{(model 1)}$$

$$concentration \sim \beta_0 + \beta_1\lambda_{225} + \beta_2\lambda_{230} + \beta_3\lambda_{255} + \beta_4\lambda_{300} + \varepsilon_r \qquad \text{(model 2)}$$

To avoid overfitting one of the models, we used Bayesian Information Criterion to reduce the set of possible wavelengths for predicting concentration for Model 1 – the positive treatment outcome model. StepAIC in R was used to improve model fit. The final version of model 1 is provided below:

$$concentration \sim \beta_0 + \beta_1\lambda_{240} + \beta_2\lambda_{250} + \varepsilon_r \quad \text{(model 3)}$$

We fit both models (model 3 and model 2) to the reference concentration curves developed (16 measurements). The models fit the experimental data well because the measured absorbance is fairly linear with respect to concentration for these wavelengths. Additionally, the unknown concentrations for each brand's sample is substantially diluted to levels that are likely to fall between the concentration estimates on the reference curve. As such, it is expected that these regression models should have good predictive capacity for samples with similar concentrations and similar purities (curve features) to what is measured in the reference curves.

By using a regression approach to estimate purity, samples that have absorption spectra that deviate less from the reference curve will have less residual error (ε_r), and therefore results from model 3 and

[1] https://projecteuclid.org/euclid.aos/1176344136

[2] An introduction to degrees of freedom in this context can be found here: https://statisticsbyjim.com/hypothesis-testing/degrees-freedom-statistics/

model 2 will be very similar. Samples that deviate more from the reference curve will have very different predictions for model 2 and model 3. As model 2 has fewer degrees of freedom (more fitted parameters), it is anticipated that samples that deviate more from the reference curve will have considerably different concentration estimates for model 2 predictions. Because model 2 makes concentration predictions using absorption wavelengths correlated with samples that are known to be suspect at treating cats, the difference between predictions can be used to approximate the purity of a sample relative to the reference solutions developed. [3]

Given that the wavelengths used for each model were grouped together based on their ability to predict negative outcomes for cats, it is expected that samples with large differences between the two curves will be less pure, or have lower active compound in solution. We calculated the percent difference between predicted concentrations for each model for all sample tests. This difference in model prediction serves to estimate a sample's impurity (curve divergence). We estimated Beta coefficients for each model, and results are provided below in table 3.

Table 3. Beta coefficient estimates for the two regression models above. Standard errors are given in parenthesis

Model Parameter	Model 3	Model 2 (impurity estimates)
Intercept	-0.001123 (0.003527)	-0.001059 (0.004073)
225 nm		-0.303589 (0.226951)
230 nm	--	0.043856 (0.140264)
240 nm	-0.025177 (0.009880)	--
250 nm	0.277139 (0.015272)	--
255 nm	--	0.698105 (1.169328)
300 nm	--	0.145525 (1.019166)

[3] UV-Vis is typically used to assess the purity of DNA and RNA samples by comparing absorption ratios at different wavelengths in the spectra. We have used statistical analysis to recover the set of absorption wavelengths that are most likely to correspond to impurities affiliated with bad treatment outcomes. https://www.promega.com/resources/pubhub/enotes/how-do-i-determine-the-concentration-yield-and-purity-of-a-dna-sample/?fbclid=IwAR01jdxtBX7X5fMEq_RYNxAHAnnmfa8pWtN-y89JkAfijWaVKhE_gekmE5E

Using this approach, we obtained concentration estimates. These results are shown in Table 4:

Table 4. Concentration estimates based on regression models fit to a set of standardized (known) GS concentrations and their respective reference absorption spectra

Brand	Avg. Model 3 Estimate (mg/mL)	Avg. Model 2 Estimate (mg/mL)	Avg. % Difference
Ajax	19.6	21.8	11%
BadShire	11.0	8.2	26%
BEAT	18.1	19.3	7%
Blossom	9.6	7.2	26%
Brava	19.2	19.9	4%
Capella	19.2	20.5	7%
Hero_White_Cap	15.3	16.0	4%
Mutian	17.5	18.5	6%
Pine_green	15.9	17.3	9%
Pine_red	16.8	20.7	23%
Ruby	19.1	21.3	12%
Shire	18.7	23.4	25%
Spark_blue	26.3	30.2	15%
Spark_purple	16.6	17.6	6%
Spark_red	26.3	29.2	11%
Standard	0.2	0.2	0%

We assume that the purity (δ_{purity}) for each sample can be represented by subtracting 1 from the percent difference for the two model concentration estimates. To estimate absolute concentrations, the purity assumption is applied to the values estimated using Model 3:

$$concentration = concentration_{m.1} \times \delta_{purity}$$ (eq. 1)

Model 3 has more degrees of freedom, and is therefore more likely to provide an unbiased concentration estimation for the compounds that look like GS. Model 2, on the other hand, is more likely to show deviations from the reference curve we created using lyophilized GS.

Samples that contain larger concentrations of unknown compounds that are similar to GS are likely to absorb UV light at many of the same wavelengths as GS. However, different compounds will absorb at each wavelength at different ratios compared to pure GS. As such, Model 2 which is fit using several different wavelengths will create substantially different estimates compared to Model 3 if there are different compounds in solution relative to the reference curve. Using this approach, highly impure samples may have inaccurate estimates of concentration under Model 2, and therefore these samples will be calculated as having a lower active compound concentration in equation 1 above. Table 5 shows our estimate of active GS concentration for each sample:

Table 5. Concentration and purity estimates for all brands of GS that were tested

Brand	Similarity to Reference Curve (Purity)	Active Compound Concentration - 95% CI (mg/mL)	Marketed Concentration (mg/mL)	As good as Spec	# of Measurements	# of Samples
Ajax	87% - 90%	16.9 - 17.9	15	Yes	6	1
BadShire	70% - 79%	7.4 - 8.9	15	No	5	1
BEAT	86% - 100%	15.3 - 18.4	15	Yes	6	1
Blossom	68% - 80%	6.3 - 8	15	No	6	1
Brava	93% - 99%	17.3 - 19.6	15	Yes	6	1
Capella	89% - 97%	16.3 - 19.4	15	Yes	12	2
Hero_White_Cap	93% - 98%	13.8 - 15.5	16.5	No	6	1
Mutian	91% - 97%	15.6 - 17.3	16.7	Maybe	6	1
Pine_green	83% - 100%	13.4 - 15.8	15	Maybe	6	1
Pine_red	67% - 87%	11 - 14.9	15	No	6	1
Ruby	83% - 93%	15.1 - 18.5	15	Yes	6	1
Shire	74% - 75%	13.4 - 14.6	15	No	6	1
Spark_blue	64% - 107%	18.3 - 26.4	20	Maybe	6	1
Spark_purple	87% - 101%	13.6 - 17.7	15	Maybe	18	3
Spark_red	73% - 106%	20.3 - 26.4	17	Yes	6	1

Conclusions

Overall, this method appears to provide decent estimates for concentration of GS in solution relative to a reference curve. Because the reference curve is made using lyophilized GS from only one brand, it was expected that the corresponding sample would have the highest purity estimates relative to the curve. Other samples with lower purity estimates will have larger deviations from the reference curve. These deviations may exist because the sample is more pure than the reference sample (fewer, similar impurities), or because it is less pure than the reference sample. As many of the samples have similar concentration estimates from both Model 3 and Model 2, it can be inferred that there are limited discrepancies between the reference curve and the sample curves that were measured for these brands. As such, concentration estimates for most brands are likely to be fairly accurate. However, there were several brands that substantially deviated from the reference curve – **Bad Shire**, **Shire**, **Blossom**, and **Pine_red**. These samples are unlikely to have treatment results similar to Brava (and other brands), and should therefore be avoided.

Finally, there is a set of samples with fairly decent alignment with the reference curve, yet the estimated concentration may fall outside of the marketed concentration. This includes: **Hero_White_Cap, Mutian, Pine_green, Spark_blue**, and **Spark_purple.** Without further testing, these brands should be used with caution. If used, these brands should only be assumed to have concentrations at the lower-end of the ranges given in Table 5, rather than at the marketed concentrations. That means, for instance, that White Cap Hero should be used as though it contains 13.7 mg/mL, rather than the marketed 16.5 mg/mL – as such, doses should be scaled upward by 17% when treating with White Cap Hero (e.g. 2.3 mL should be given rather than 2 mL).

Notes on Study Limitations

The statistical approaches we used in this paper to estimate concentrations and purity are limited by the data and measurements that were possible.

We estimate measurement error to be less than 5%. This manifests from the UV-Vis spectrophotometer error, which is accounted for with replicate measurements. There is also measurement error due to the limited quantity of lyophilized GS-441524 that was provided for testing. Due to the relatively small quantity, and that measurement precision was at the $1/10^{th}$ of a milligram level, there is some error associated with the absolute concentration estimates of the reference curve, although this is unlikely to fall outside of the sampling error of the machine, but may bias concentration estimates upwards or downwards systematically.

Additionally, for most samples, only one vial was available for testing. Given this, it is impossible to generalize results from this analysis to all vials being distributed by a given brand or sequential batches over time. More testing across batches and over time will be necessary to better understand the variability of samples from each brand. Concentration estimates and relative purity estimates may vary depending on the when the batches are tested, and how many samples from a given brand are tested. As such, results from this analysis should be interpreted cautiously.

Finally, this testing was done using a UV-Vis spectrophotometer, which is unable to estimate absolute purity of a sample. As such, we could only estimate the relative purity compared to the reference curve we created. All purity estimates in this study are relative to the one lyophilized GS sample from one GS brand. If this sample is at relatively low purity, then all samples with similar relative purity would also be similarly impure. Additionally, this analytical method is unable to detect any contaminants that may be

contained within the diluent, or that may fall outside of the absorption wavelengths tested. As such, some brands could be more or less contaminated, and this would not show up in this analysis. Alternative analytical chemistry techniques are necessary to assess absolute purity, and for better understanding potential risks or toxicity for cats.