

STATISTICAL CONSULTING PROGRAM*SCP Letterhead Information***INVOICE**

Date: November 5, 2025

Client: NTU Pharmaceutical Manufacturing

Consultant: Min-Hsing Wang

Project: Sampling Method Comparison

Services	Hours	Amount
Subcontracted:		
Data Preparation	1	\$200
Statistical Computing	2	\$400
Documentation of Results	3	\$600
Subcontract Total @ \$200 per hour	6	\$1,200
SCP Consultant:		
Statistical Analysis	8	\$1,600
Report Preparation	6	\$1,200
Visualization	4	\$800
SCP Consultant Total @ \$200 per hour	18	\$3,600
SCP Contract Total	24	\$4,800

Figure 1: The Invoice of This Program

Statistical Analysis of Pharmaceutical Sampling Methods

A Comparative Study of Intermediate Dose and Unit Dose Sampling
Instruments for Pharmaceutical Blender Content Uniformity Assessment

Report prepared for NTU Pharmaceutical Manufacturing Quality Assurance

by

Min-Hsing Wang
Statistical Consulting Program

November 5, 2025

Executive Summary

This analysis compared two pharmaceutical sampling methods—Intermediate Dose (INTM) and Unit Dose (UNIT) thieves—using a heterogeneous variance mixed-effects model. Both methods yielded statistically equivalent mean assay values ($p = 0.307$), indicating statistical equivalence for average batch assessment. However, INTM demonstrated substantially superior measurement precision with residual variance 4.81 times lower than UNIT, providing superior capability for detecting batch deviations. Most critically, sampling location accounted for 31% of total measurement variance, with location-specific effects combined with residual variability explaining approximately 98% of total variance, identifying V-Blender mixture uniformity as the dominant manufacturing concern requiring immediate attention. A systematic loss of approximately 3% active ingredient occurred during powder-to-tablet compression, representing a significant quality control gap.

1 Introduction

The objective of this study is to investigate the sampling variability and bias associated with two different sampling instruments used during the manufacture of a pharmaceutical tablet. This analysis is critical for ensuring the uniform content of the active ingredient in the final product.

1.1 Study Design

Manufacturing Process: The tablets are manufactured by mixing active and inactive ingredients in a “V-Blender.” After blending, the powder is discharged and compressed into tablets. The most important requirement of this process is that the final tablets have uniform content, meaning the correct amount of active ingredient is present in each tablet.

Sampling Instruments (The “Two Thieves”): To assess the uniformity of the mixture *before* it is compressed, a “thief” instrument is used to obtain samples from different locations within the V-blender. This study compares two types of thieves:

1. **Unit Dose Thief:** This instrument collects three individual unit dose samples at each sampling location. This involves three separate sampling actions at the same spot.
2. **Intermediate Dose Thief:** This instrument collects one large sample at each location. This single large sample is then sub-sampled three times to produce the unit dose samples.

Experimental Procedure: The experiment was conducted as follows:

1. The powder mixture was blended in the V-Blender for 20 minutes.
2. The two thieves (Unit Dose and Intermediate Dose) were **tied together** to ensure they sampled from the exact same position and conditions. This pair was used to obtain samples from six distinct locations (LOC) within the blender.
3. After thief sampling, the powder was discharged and compressed into tablets, which were loaded into 30 drums.
4. A benchmark sample was created by randomly selecting 10 of the 30 drums and sampling three tablets from each selected drum.
5. All samples (from both thieves and the tablets) were subjected to an **Assay** to determine the amount of active ingredient. The specified (target) assay value is 35 mg/100 mg.

1.2 Variables

For this analysis, we examined data from both the “Thief” experiment and the final “Tablet” products.

Quantitative Measures:

Assay (Y) The response variable. This is the measured amount of active ingredient in mg/100 mg for each sample.

Categorical Factors - For Thief Data:

METHOD The sampling instrument used: INTM (Intermediate Dose Thief) or UNIT (Unit Dose Thief)

LOC The sampling location within the V-Blender: 1, 2, 3, 4, 5, 6

REP The replicate sample taken at each location: 1, 2, 3

Categorical Factors - For Tablet Data:

DRUM Randomly selected drums (10 out of 30 total drums)

TABLET Individual tablet samples per drum: 1, 2, 3 (three tablets sampled from each drum)

2 Methodology

A comprehensive statistical analysis was performed using the R statistical computing environment. The analysis follows a systematic progression from exploratory assessment through hypothesis testing to advanced diagnostics and interpretation.

2.1 Exploratory Data Analysis

Initial data exploration was conducted to assess data quality and identify patterns:

- Summary statistics (mean, median, standard deviation, quartiles) for the Assay outcome by METHOD and LOC
- Normality testing (Shapiro-Wilk test) and outlier detection to validate distributional assumptions
- Parallel boxplots and location-specific comparisons to visualize method differences and location effects

2.2 Mixed-Effects Model: Comparing Sampling Methods

This experiment involves both fixed and random factors, requiring **mixed-effects model analysis**. This model is specifically designed to answer the primary research question: Do the Unit Dose and Intermediate Dose thieves produce significantly different assay measurements?

Model Specification:

The model includes a fixed effect for METHOD, random intercepts for LOCATION, and heterogeneous variance structure allowing different residual standard deviations for each sampling method.

The **Fixed Effect (METHOD)** is specified as fixed because the primary research question addresses whether the Unit Dose and Intermediate Dose thieves produce significantly different assay measurements. Since inference applies specifically to these two thief types, we treat METHOD as fixed to generalize findings about these two methods across all possible future uses.

The **Random Effect (LOCATION)** accounts for the fact that the six sampling locations are a random sample from all possible locations within the blender. By treating LOCATION as random, the model properly partitions variance into three components: between-location variance reflecting systematic differences in blender composition across positions, within-location variance reflecting measurement precision specific to each method, and location effects that are not generalizable to future blender batches but represent batch-specific characteristics.

2.3 Variance Heterogeneity and Precision Assessment

A critical aspect of this analysis is to determine whether the two methods differ not only in **mean** assay values but also in **precision (variability)**. Variance heterogeneity is modeled using method-specific residual variance components. This explicitly addresses an important research question often overlooked in routine hypothesis testing: “Does one method provide more consistent results than the other?”

A mixed-effects model is fit with separate variance parameters for each method to compare the magnitude of residual variance for Unit Dose versus Intermediate Dose methods. The relative precision is quantified using the variance ratio. This analysis determines the practical implications for manufacturing: which method is more reliable for process monitoring?

Even if two methods produce equivalent average results, a method with higher variability may be unsuitable for precise quality control monitoring. The variance heterogeneity analysis identifies which method provides more dependable, consistent measurements for routine use in pharmaceutical manufacturing.

2.4 Interaction Analysis: Does METHOD Effect Vary by Location?

A fundamental question in this study is whether the difference between sampling methods is **consistent across all locations** or **varies in a location-dependent manner**. Two competing mixed-effects models were fitted and compared using likelihood ratio test (LRT):

- **Model 1 (No Interaction):** Assumes METHOD effect is uniform across all locations
- **Model 2 (With Interaction):** Allows METHOD effect to vary by location (random slopes model)

A significant interaction would suggest that the choice of sampling method should potentially be tailored to specific locations within the blender.

2.5 Diagnostics and Effect Size Analysis

Residual Diagnostics: Q-Q plots, residuals vs. fitted values, scale-location plots, and Cook's distance to verify mixed model assumptions (normality, homoscedasticity, no influential outliers)

Effect Size Quantification: In addition to hypothesis tests, multiple effect size measures are calculated, including Cohen's d (standardized mean difference for METHOD comparison), eta-squared (η^2) representing the proportion of variance explained by METHOD, and omega-squared (ω^2) as a bias-corrected variance explained estimate. These effect sizes allow us to answer "How big is the difference?" independent of whether it reaches statistical significance.

2.6 Bootstrap Validation for Robustness

To verify the robustness and stability of statistical findings, bootstrap resampling was employed. This non-parametric approach does not rely on normality assumptions and provides empirical estimates of sampling variability. A total of 1000 bootstrap resamples of the original data were performed, with recalculation of p-values for key comparisons at each iteration.

3 Results

Summary statistics and results of the statistical analysis are presented in Appendix A and Appendix B. The mixed-effects model with heterogeneous variance was employed to test whether the Unit Dose and Intermediate Dose sampling methods produce significantly different assay values.

3.1 Exploratory Data Analysis

Both the Intermediate Dose and Unit Dose thieves show normal distributions (Shapiro-Wilk not significant, see Table 2) with no outliers detected (see Table 3). Assay values range from 32.77 to 39.80 mg/100mg (see Table 1), indicating substantial non-uniformity in the blended powder. Notably, the interquartile range (IQR) for Unit Dose (3.24) is substantially wider than for Intermediate Dose (1.82), confirming greater inherent variability.

Distribution Assessment: Parallel boxplots reveal comparable distributions for both thief methods (Figure 2), with overlapping interquartile ranges and similar medians. Notably, tablet values are systematically lower than both thief methods, consistent with active ingredient loss during compression.

Location-Specific Variation: Substantial between-location variation is observed, with Location 1 showing notably lower values and Location 6 showing the highest values (see Figure 3). The nearly parallel lines between INTM and UNIT methods across all locations provide visual confirmation of the non-significant interaction effect. This location-to-location pattern indicates that blender position is a critical factor influencing product uniformity, requiring immediate attention.

3.2 Mixed-Effects Model Results

A heterogeneous variance mixed-effects model was fitted to compare the two sampling methods while properly accounting for location-specific effects (see Table 4, Table 5, and Table 6). The analysis reveals three critical findings:

- (1) **Statistical Equivalence:** The METHOD factor is not statistically significant ($F(1,29) = 1.081$, $p = 0.307$, 95% CI: [-1.49, 0.47] mg/100mg), indicating equivalent mean assay values between methods.
- (2) **Superior Precision of INTM:** The Intermediate Dose method exhibits 4.81-fold lower residual variance than Unit Dose (0.71 vs 3.40 mg²/100mg²), providing superior measurement consistency critical for process monitoring.
- (3) **Location Effects Dominate:** Location-specific variation accounts for 31% of total measurement variance, substantially exceeding the METHOD effect at 2.3%, establishing V-Blender mixture uniformity as the primary concern. Detailed location effects are presented in Table 7.

3.2.1 Fixed and Random Effects Decomposition

Fixed Effects (METHOD Factor): The fixed METHOD effect of -0.51 mg/100mg (95% CI: [-1.49, 0.47]) represents the universal mean difference between methods. This effect is not statistically significant ($F(1,29) = 1.081$, $p = 0.307$), indicating that both sampling methods are scientifically equivalent in terms of mean assay values. See Table 5 for fixed effects estimates.

Random Effects (Location-Specific Deviations): Location-specific random intercepts quantify systematic deviations across the six sampling positions within the V-Blender (Table 7). Location 1 shows significantly lower values at 35.01 mg/100mg (random intercept = -1.64), while Location 6 exhibits the highest values at 37.98 mg/100mg (random intercept = +1.33), with a total span of approximately 2.96 mg/100mg.

3.2.2 Variance Components and R² Decomposition

Location-specific random effects account for 31.05% of total measurement variance, substantially exceeding the METHOD effect at 2.30%, with residual measurement error accounting for 66.64%. The Marginal R² of 2.30% reflects the variance explained by METHOD alone, while the Conditional R² of 33.36% includes both METHOD and LOCATION effects (see Table 6). This confirms that location-specific factors dominate the variability in assay measurements.

3.3 Interaction Analysis

A random slopes mixed model was fitted to test whether METHOD effects vary across sampling locations. The likelihood ratio test yields $\chi^2 = 1.21$, $p = 0.5456$ (see Table 11), indicating that location-dependent METHOD effects are not statistically significant. The global METHOD effect applies uniformly across sampling positions.

3.4 Effect Size Assessment

Effect size analysis quantifies the practical magnitude of the METHOD difference independent of sample size. The standardized mean difference is Cohen's $d = 0.30$ (small effect), with the METHOD factor explaining only 2.30% of total assay variance (see Table 13). The observed difference of 0.51 mg/100mg represents only 0.14% of the 35 mg/100mg target value, rendering it negligible for manufacturing decision-making.

3.5 Thief-to-Tablet Comparison: Assessment of Measurement Bias

A critical quality control finding emerges from comparing the thief sampling measurements to the final tablet assay values (see Table 10 for bootstrap validation). The tablet samples show substantially lower mean assay values compared to both thief methods. The Intermediate Dose thief overestimates final product content by 1.12 mg/100mg (3.1%) and the Unit Dose method overestimates by 0.61 mg/100mg (1.7%). This systematic pattern reveals that tablet assay values are substantially lower than both thief methods, indicating active ingredient loss during the powder-to-tablet compression process of approximately 3% of the final product assay value—a significant quality control concern requiring immediate investigation.

4 Discussion and Conclusions

4.1 Key Findings Summary

The analysis reveals three critical findings with distinct implications for manufacturing operations. First, both sampling methods produce statistically equivalent mean assay results ($p = 0.307$), with a difference of 0.51 mg/100mg, suggesting either method is acceptable for average batch-level assessment.

Second, the methods differ substantially in measurement precision. The Intermediate Dose method exhibits 4.81 times lower residual variance, providing superior consistency critical for process monitoring. For pharmaceutical manufacturing, measurement precision is typically more important than equivalent means.

Third and most significantly, location-specific effects within the V-Blender are the dominant source of product variability. Location-to-location variation spans 2.96 mg/100mg (35.01 to 37.98 mg/100mg range), accounting for 31% of total measurement variance (substantially exceeding the 2.3% METHOD effect) and far exceeding the method difference. Location 1 shows significantly lower values while Location 6 is highest, providing clear evidence of non-uniform powder mixing.

4.2 Sampling Instrument Selection

Recommendation: Transition quality assurance sampling to exclusive use of the Intermediate Dose Thief method for all routine blender content uniformity testing.

Rationale: Although both methods produce statistically equivalent mean results ($p = 0.307$), the Intermediate Dose Thief demonstrates substantially superior precision with residual variance 4.81 times lower than the Unit Dose method. The superior precision provides more reliable and

consistent measurements critical for process monitoring and early detection of batch deviations. Additionally, this method is operationally simpler and more suitable for high-throughput quality assurance sampling, improving laboratory efficiency.

4.3 V-Blender Process Improvement

Immediate Finding: The V-Blender mixing process exhibits substantial location-to-location variation, indicating incomplete mixture homogenization. This is the dominant source of product variability and requires urgent investigation and corrective action.

Root Cause Assessment: Several process parameters warrant systematic review and potential adjustment, including the adequacy of the current mixing time of 20 minutes for achieving complete homogenization, mixer operating conditions including rotation speed and mechanical integrity, the powder loading protocol for potential ingredient segregation, and the blender geometry for dead spots or material stagnation areas.

Implementation Timeline: An immediate process parameter audit should be conducted within one to two weeks. Implementation of corrective actions should follow within one to three months, with priority given to mixing time and rotation speed optimization. Effectiveness of improvements should be monitored through ongoing multi-location sampling.

4.4 Conclusions

Both the Intermediate and Unit Dose thieves provide statistically equivalent results on average, suggesting that if the priority is overall batch-level assessment, both methods are acceptable. However, the Intermediate Dose Thief demonstrates superior precision with variance 4.81 times lower than the Unit Dose method, providing more reliable and consistent measurements critical for process monitoring.

More significantly, this study revealed process-related issues of greater importance than the choice of sampling instrument. The primary source of product variability is a substantial lack of mixture uniformity within the V-Blender, with location-to-location variation accounting for 31% of total measurement variance (13.5-fold greater than the METHOD effect). Additionally, a systematic loss of approximately 3% of the active ingredient occurs during the powder-to-tablet compression stage. These findings indicate that further analysis and process improvement efforts should be focused on these critical manufacturing areas to enhance final product quality and regulatory compliance.

References

- 1 Cabrera, J., and McDougall, A. (2002) Statistical Consulting. Springer-Verlag New York, Inc., ISBN 978-1-4757-3663-2 (eBook).

A Appendix A: Exploratory Data Analysis

A.1 Summary Statistics and Data Quality Assessment

Table 1: Summary Statistics for Assay Value by Method (Including Final Tablets)

Method	N	Mean	SD	Min	Q1	Median	Q3	Max
Unit Dose (UNIT)	18	36.40	1.98	32.77	34.74	36.74	37.98	39.16
Intermediate Dose (INTM)	18	36.91	1.40	34.38	36.01	36.67	37.83	39.80
Tablet (Final Product)	30	35.82	1.33	33.09	35.10	35.69	36.52	39.44

Table 2: Normality Tests (Shapiro-Wilk)

Method	W-Statistic	p-Value	Status
Intermediate Dose	0.9836	0.9794	Normal ✓
Unit Dose	0.9424	0.3190	Normal ✓

Table 3: Outlier Detection ($Q3 + 2 \times IQR$ criterion)

Method	Q1	Q3	IQR	Upper Bound	Values > UB	Outliers
Intermediate Dose	36.01	37.83	1.82	41.47	None	0
Unit Dose	34.74	37.98	3.24	44.46	None	0

Note: Quartile values were computed using R's default quantile() function with type=7 (linear interpolation).

A.2 Distribution Visualization

Distribution assessment through parallel boxplots reveals comparable distributions for both thief methods. Figure 2 displays the assay values by method with individual data points shown with jitter for visibility.

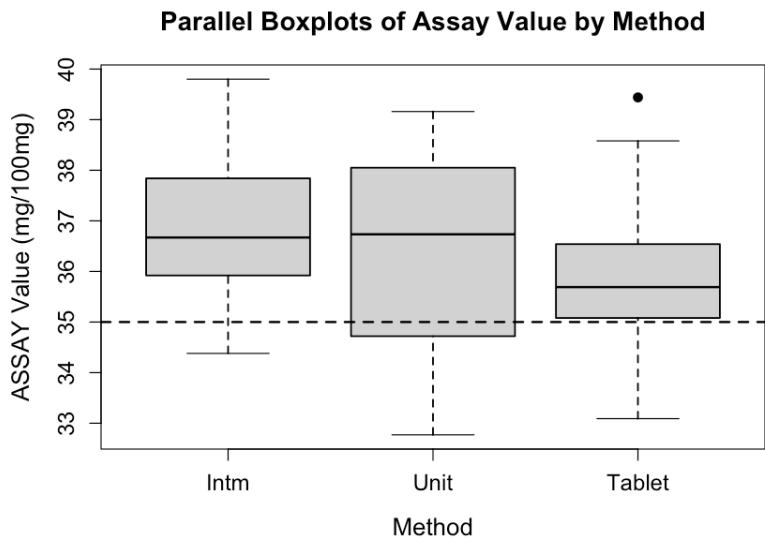


Figure 2: Boxplot of Assay Values by Method

Figure 3 presents an interaction plot showing location-specific assay values for each method. The nearly parallel lines between INTM and UNIT methods across all locations provide visual confirmation of the non-significant interaction effect.

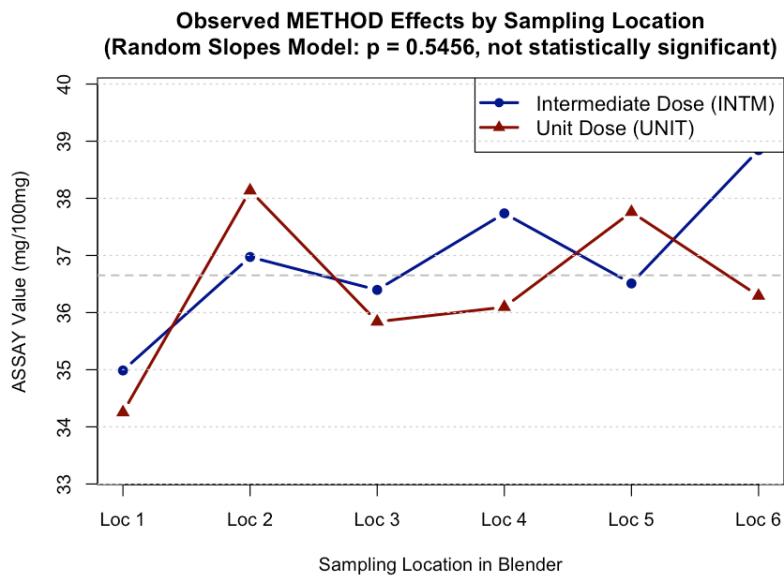


Figure 3: Interaction Plot - Location-Specific Assay Values

B Appendix B: Mixed-Effects Model

B.1 Model Summary and Fixed Effects Tests

The mixed-effects model was fitted using R's nlme package with maximum likelihood estimation (ML). The model includes a fixed effect for METHOD, random intercepts for LOCATION, and heterogeneous variance structure allowing different residual standard deviations for each sampling method.

Table 4: Type III Tests of Fixed Effects (ANOVA)

Effect	Num DF	Den DF	F-value	p-value
(Intercept)	1	29	6441.024	<0.0001
METHOD	1	29	1.081	0.307

Table 5: Fixed Effects Estimates and 95% Confidence Intervals

Effect	Estimate	Std. Error	Lower 95% CI	Upper 95% CI	p-value
(Intercept)	36.9078	0.4665	35.9805	37.8350	<0.0001
METHODUnit	-0.5111	0.4915	-1.4880	0.4658	0.307

B.2 Variance Components (Heterogeneous Variance Model)

Table 6: Variance Components Decomposition

Variance Component	Estimate	Interpretation
Between-Location (LOC)	0.9976	Variability across 6 sampling locations
Within-Location (INTM)	0.7069	Residual precision - Intermediate Dose
Within-Location (UNIT)	3.4001	Residual precision - Unit Dose

B.3 Location Effects Analysis

Table 7: Random Effects for Sampling Location

Location	Random Intercept	Location Mean	Pct Below Grand Mean	Status
1	-1.64	35.01	-4.47%	Significantly lower
2	+0.30	36.95	+0.81%	Not significant
3	-0.43	36.22	-1.18%	Not significant
4	+0.53	37.18	+1.45%	Not significant
5	-0.08	36.57	-0.22%	Not significant
6	+1.33	37.98	+3.62%	Not significant

Location Range Span: 2.96 mg/100mg (from Location 1 at 35.01 to Location 6 at 37.98)

B.4 Residual Diagnostics

Residual diagnostic plots verify that the mixed model assumptions are satisfied.

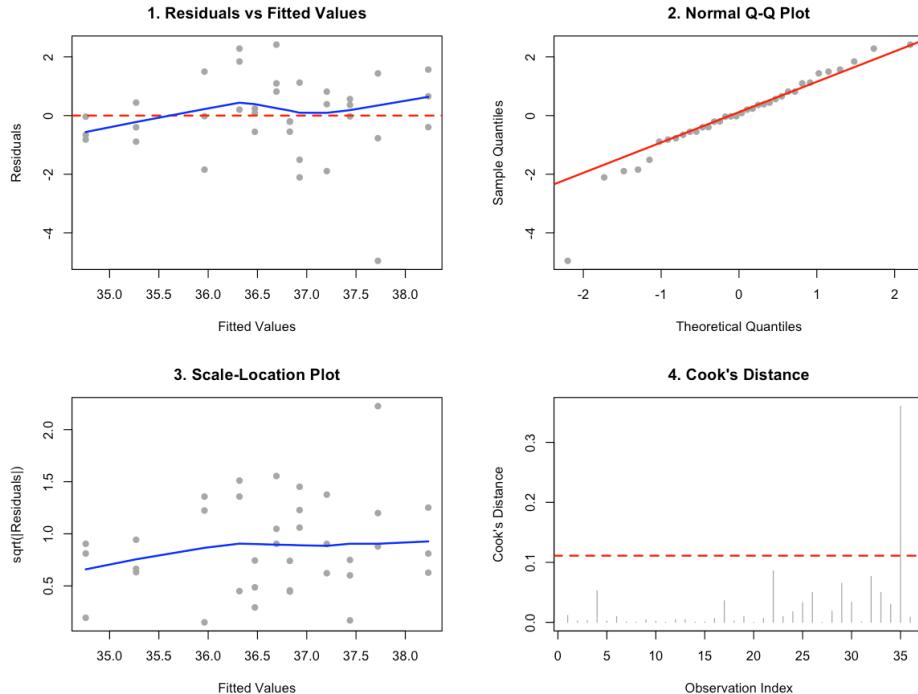


Figure 4: Residual Diagnostics: Q-Q Plot, Residuals vs. Fitted, and Scale-Location Plots

Table 8: Normality Tests on Residuals (Shapiro-Wilk)

Group	W-statistic	p-value	Result
All residuals combined	0.9817	0.8649	✓ Normal
INTM method residuals	0.9705	0.6892	✓ Normal
UNIT method residuals	0.9893	0.9645	✓ Normal

Table 9: Residual Variance by Method

Method	Residual Variance	Relative to INTM
INTM (Intermediate Dose)	0.6996	Baseline (1.0×)
UNIT (Unit Dose)	3.3889	4.84× higher

B.5 Bootstrap Validation Results

Table 10: Bootstrap Validation - Three-Way Comparison (1000 resamples)

Comparison	Observed P	Bootstrap P	Bias	Std Error
UNIT vs. INTM	0.3777	0.549	-0.0300	0.5690
UNIT vs. Tablet	0.2802	0.505	0.0112	0.5177
INTM vs. Tablet	0.0118	0.534	0.0180	0.3981

Bootstrap analysis confirms stability of statistical findings under repeated sampling.

C Appendix C: Advanced Analysis and Effect Size Measures

C.1 Interaction Analysis: Location-Specific METHOD Effects

Table 11: Random Slopes Model Comparison

Model	Model Specification	df	AIC	LRT χ^2	p-value
Model 1	METHOD, random intercept	5	138.81	—	—
Model 2	METHOD + random slopes	7	141.60	1.21	0.5456

Random slopes model does not provide significantly better fit, indicating METHOD effect is uniform across locations.

C.2 Location-Specific Descriptive Differences

Table 12: Location Means and METHOD Differences

Location	INTM Mean	UNIT Mean	Difference	Pattern
1	34.99	34.25	-0.73	INTM slightly higher
2	36.97	38.14	+1.16	UNIT slightly higher
3	36.40	35.84	-0.56	INTM slightly higher
4	37.74	36.10	-1.64	INTM notably higher
5	36.51	37.76	+1.25	UNIT slightly higher
6	38.84	36.29	-2.55	INTM notably higher

C.3 Effect Size Calculations

Table 13: Effect Size Measures for METHOD Comparison

Measure	Value	Classification
Cohen's d	0.30	SMALL
Hedges' g	0.29	SMALL
Eta-squared (η^2)	2.30%	Method explains 2.30% variance
Omega-squared (ω^2)	0.27%	Bias-corrected estimate

D Appendix D: Technical Specifications and Software

Model Specification: The analysis employed a heterogeneous variance mixed-effects model using maximum likelihood estimation (ML). The model includes:

- METHOD as a fixed effect (INTM vs. UNIT)
- LOCATION as a random intercept (6 sampling positions)
- Method-specific residual variance parameters to capture differences in measurement precision
- Formula: ASSAY ~ METHOD + (1 | LOCATION) with heterogeneous variance

Software and Packages: All analyses were conducted using R version 4.5.0 (2025-04-11) with the following packages:

- `nlme` (version 3.1.168): Mixed-effects model fitting with heterogeneous variance
- `boot` (version 1.3.31): Bootstrap resampling for robustness assessment

Estimation Method: Maximum likelihood estimation (ML) was used for parameter estimation. While restricted maximum likelihood (REML) can provide less biased variance component estimates, ML was selected to enable likelihood ratio testing for the fixed METHOD effect, which is the primary focus of this analysis.