Western Michigan University

STAT 6040: FUNDAMENTALS OF EPIDEMIOLOGY AND CLINICAL TRIALS

Department of Statistics

A Randomized Controlled Trial Comparing Water vs. Sports Drink on Rehydration Effectiveness After Moderate-Intensity Exercise

BY
Augustine Ennin
August 14, 2025

Abstract

Background: The optimal post-exercise rehydration remains debated, with sports drinks theoretically being superior to water due to electrolyte content. The aim of this study is to compare the effectiveness of water and a carbohydrate-electrolyte sports drink at rehydration in healthy adults.

Methods: A simulated randomized clinical trial (RCT) was carried out on 100 subjects (50/group). The main outcome was urine specific gravity (USG) 2 hours after rehydration. Secondary outcomes were weight recovery, thirst ratings (1–10 VAS), and tolerability. Data were simulated in R using parameters from previous hydration studies (e.g., USG: 1.012 ± 0.003 sports drink vs. 1.020 ± 0.004 water). Analyses used linear regression (USG), beta regression (weight recovery), and logistic regression (tolerability).

Results: (Simulated) The sports drink group had significantly lower USG ($\beta = -0.008$, $p < 0.01$) and higher weight recovery ($90 - 100\%$ vs. $80 - 92\%$) compared to water. Thirst scores were lower in the sports drink group (3 ± 1 vs. 5 ± 1.2, $p < 0.05$), though side effects were more frequent (25% vs. 10%, $OR = 2.5$).

Conclusion: Trials with simulated data indicate that sports drinks can enhance short-term rehydration measures but with slightly greater side effects. Real-world studies are required to confirm these observations.

Keywords: hydration, sports drink, urine specific gravity, randomized trial, exercise recovery.

Contents

1. Introduction

Fluid balance restoration following exercise is essential for both performance recovery and physiological homeostasis (Sawka et al., 2007). Although water is the most readily available rehydration beverage, carbohydrate-electrolyte sports drinks are theorized to improve rehydration by promoting fluid retention via sodium-mediated osmotic processes (Maughan & Shirreffs, 2010). Empirical comparisons of such strategies, however, report variable outcomes, in part because of heterogeneity in study designs and subject factors (Shirreffs et al., 2004).

Two gaps are prominent in the existing literature: (1) Standardization of exercise intensity and environmental conditions for comparing drinks is rare in studies, and (2) There is limited data on trade-offs between rehydration effectiveness and tolerability (e.g., bloating). We fill these gaps by simulating an RCT under controlled conditions with evidence-based parameter values to model outcomes.

1.1. Objectives

1. Compare the effects of water and a commercial sports drink on post-exercise USG (primary outcome).
2. Evaluate differences in weight recovery, thirst perception, and tolerability (secondary outcomes).

The simulated approach allows isolation of beverage effects while controlling for confounders (e.g., sweat rates, diet). Findings will inform the design of future clinical trials and practical hydration recommendations.

2. METHODOLOGY

2.1. Study Design

This study employed a simulated two-arm, parallel-group randomized controlled trial (RCT) to compare the effects of plain water versus a commercial carbohydrate-electrolyte sports drink on post-exercise rehydration. The simulation parameters were designed to reflect those of a real-world clinical trial. The participants were randomly allocated in a $1:1$ ratio to receive either water ($n = 50$) or the sports drinks ($n = 50$). The design was open-label in the simulation; however, in an actual trial setting, a double-blind, double-dummy approach would be implemented to minimize bias.

2.2. Data Generation

A. **Participant Characteristics**
The simulated cohort was representative of healthy adults between 18 and 35 years of age with no history of cardiovascular, renal, or metabolic disease. All subjects were assumed to have participated in a standardized moderate-intensity exercise protocol (60˘70% of VO_2max) of sufficient duration to promote mild dehydration (1–3% body mass loss). The environmental temperature was assumed to be 25°C at 50% humidity. Demographic variables of age (normally distributed with $\mu = 25$ years, $\sigma = 3$), sex (60% male and 40% female), and initial weight (70 ± 5 kg) were created to mimic distributions described in published hydration research.

B. **Interventions**
After exercise-induced dehydration, subjects in the sports drink group were given a commercially available carbohydrate-electrolyte drink that consisted of sodium, potassium, and carbohydrate (6–8% concentration), whereas the water group was given plain, non-carbonated water. The amount of fluid given was standardized to 150% of body mass lost during exercise, consumed over a 60-minute recovery.

C. **Outcome Variables**

Variable	**Type**	**Simulation Parameters**	**Source**
Urine Specific Gravity (USG)	Continuous	Water: 1.020 ± 0.004; Sports drink: 1.012 ± 0.003	Maughan & Shirreffs (2010)
Weight recovery (%)	Continuous	Water: $80-92\%$; Sports drink: $90-100\%$	Sawka et al. (2007)
Thirst score (VAS 1-10)	Ordinal	Water:$\mu = 5 \pm 1.2$; Sports drink: $\mu = 3 \pm 1$	Shirreffs et al. (2004)
Tolerability	Binary	Water: 10% side effects; Sports drink: 25%	Simulated for realism

2.3. Statistical Analyses

1. **Primary Outcome Analysis**
The primary outcome, urine specific gravity (USG), was analyzed using a multiple linear regression model specified as follows:

$$USG_i = \beta_0 + \beta_1 Group_i + \beta_2 Age_i + \beta_3 Sex_i + \epsilon_i$$

where $Group_i$ denotes hydration group (water vs. sports drink), Age_i and Sex_i are covariates, and ϵ_i represents the error term assumed to be normally distributed with mean zero and constant variance. The model's validity was confirmed through comprehensive diagnostic testing of key regression assumptions:

- **Normality of residuals** was verified using the Shapiro-Wilk test (W = 0.985, p = 0.338), supported by visual inspection of Q-Q plots showing adherence to the normal distribution line.
- **Homoscedasticity** was confirmed through Breusch-Pagan testing (BP = 2.7, p = 0.44), with residual plots demonstrating consistent variance across predicted values.
- **Absence of multicollinearity** was established with all variance inflation factors (VIFs) < 2, well below the conventional threshold of 5, indicating no concerning correlations between predictor variables.
- **Linearity** was assessed through partial regression plots, showing appropriate linear relationships between predictors and the outcome.
- **Independence** of errors was ensured by the study's randomized design and confirmed through Durbin-Watson testing (DW = 2.516, p = 0.99).

All diagnostic procedures were conducted using R, with the `lmtest`, `car`, and `stats` packages.

2. **Secondary Outcome Analysis**
For secondary outcomes, distinct regression approaches were employed according to each variable's measurement characteristics.

i. **Body Weight Recovery**: Weight recovery percentage was analyzed using beta regression with a logit link function to account for its bounded nature between 0 and 100%. The model specification was:

$$logit(E[Weight\ Recovery]) = \beta_0 + \beta_1 group + \beta_2 age + \beta_3 sex + \beta_4 weight\_loss\_pct$$

where group indicates the hydration treatment (water = 0, sports drink = 1), Age is measured in years, and Sex is coded as male = 0, female = 1, and Weight Loss % represented baseline dehydration severity. Coefficient ($\beta_1\ to\ \beta_4$) represent the log-odds change in weight recovery per unit predictor increase. Results are reported as proportional changes with 95% confidence intervals, derived via the `betareg` package in R.

ii. **For the binary tolerability outcome (presence/absence of side effects)**, logistic regression was implemented with the hydration group as the primary predictor and age and sex as covariates. The model specification was:

$$log\left(\frac{P(success)}{P(failure)}\right) = \beta_0 + \beta_1 group + \beta_2 age + \beta_3 sex$$

where the left-hand is the log-odds, Group indicates the hydration treatment (water = 0, sports drink = 1), Age is measured in years, and Sex is coded as male = 0, female = 1. The β_1 coefficient represents the adjusted log-odds ratio for side effects in the sports drink group compared to water.
Model diagnostics included comprehensive residual analysis. The residual plot (Figure X) displayed randomly scattered points without discernible patterns, suggesting appropriate model specification and absence of systematic bias. The final model estimates, presented as adjusted odds ratios with 95% confidence intervals, were derived using maximum likelihood estimation in R.

iii. **Thirst scores** was measured on a 1-10 visual analog scale were analyzed using proportional odds ordinal logistic regression.The model specification was:

$$\text{logit}\big(P(Y \leq j)\big) = \beta_{0j} + \beta_1(\text{Group}) + \beta_2(\text{Age}) + \beta_3(\text{Sex}), \quad j = 1, \dots, 9,$$

where Y represents the ordinal thirst score (1–10), $\beta_{0,j}$ are the threshold parameters for each cumulative probability, Group was coded as 0 (water) or 1 (sports drink), Age was modeled continuously (years), and Sex was coded as 0 (male) or 1 (female). Key assumptions, including proportional odds, were verified using the Brant test, while multicollinearity was examined through variance inflation factors, and category completeness was confirmed through frequency tables of response categories.
The model was fitted using the `polr` function from the `MASS` package in R, with significance tested via likelihood ratio tests. Results are reported as cumulative odds ratios with 95% confidence intervals.

3. ANALYSIS

3.1. Participant Characteristics

A total of 100 participants were included in the analysis, with 50% assigned to the Water group and 50% to the Sports Drink group. The mean age was 25.27 (2.7847) years, and the sex distribution was approximately 60% male and 40% female. Baseline weight loss percentage and other relevant characteristics were balanced across groups (Table 3.1).

Table 3.1.: Baseline Participant Characteristics by Group

Characteristic	**Water Group**	**Sports Drink Group**	**Total (n = 100)**
Number of participants	50	50	100
Age (mean ± SD)	25.1 ± 2.8085	25.44 ± 2.7786	25.27 ± 2.7847
Sex, n (%)			
Male	29 (29%)	32 (32%)	61 (61%)
Female	21 (21%)	18 (18%)	39 (39%)
usg	1.0197 ± 0.0041	1.0118 ± 0.0032	1.0158 ± 0.0054
Weight − Pre	70.192 ± 4.6533	69.956 ± 4.7414	70.074 ± 4.6753
Weight loss (kg)	1.3928 ± 0.4213	1.4672 ± 0.4447	1.43 ± 0.4325
Baseline weight loss (%)	1.9881 ± 0.5953	2.0881 ± 0.5857	2.0381 ± 0.5897

3.2. Primary Outcome: Urine Specific Gravity

The primary analysis evaluated differences in post-exercise hydration status between intervention groups using urine specific gravity (USG) as the continuous outcome measure. A linear regression model was fitted with hydration group as the primary predictor, adjusting for age and sex as covariates.
The model was a satisfactory fit `(F(3,96) = 41.16, p < 0.001)`, explaining 56.3% of the variance in USG measurements (adjusted $R^2 = 0.549$) . Participants receiving the sports drink had significantly lower USG values compared to those consuming water $\beta = -0.0079$, 95% CI [-0.0093, -0.0064]; t(96) = -10.77, p < 0.001), indicating superior rehydration.

Covariate analysis revealed no significant sex differences ($\beta = -0.00019,\ p = 0.80$), while age showed a marginal trend toward better hydration in older participants ($\beta = -0.00025,\ p = 0.059$). The intercept of 1.0262 (95% CI [1.019, 1.033]) represents the estimated USG for the reference water group.

Model diagnostics confirmed all linear regression assumptions:

- Residuals were normally distributed (Shapiro-Wilk W = 0.985, p = 0.338)
- Homoscedasticity was maintained (BP = 2.7, p = 0.44)
- Variance inflation factors < 2 indicated no multicollinearity concerns

Table 3.2.: Linear Regression Results for Urine Specific Gravity (Reference: Water Group)

Predictor	β **(95% CI)**	**SE**	t**-value**	p**-value**
(Intercept)	1.026 (1.019, 1.033)	0.0034	300.59	<0.001
Sports Drink	-0.008 (-0.009, -0.006)	0.0007	-10.77	<0.001
Age	-0.00025 (-0.001, 0.000)	0.0001	-1.91	0.059
Sex (Male)	-0.00019 (-0.002, 0.001)	0.0007	-0.25	0.800

These results provide compelling evidence that carbohydrate-electrolyte sport drink improve more effective post-exercise rehydration than water alone, as quantified by standardized USG measures.

3.3. Secondary Outcome

3.3.1. Body Weight Recovery

The proportional recovery of body weight was evaluated using beta regression with logit link to account for the bounded nature (0-100%) of this outcome measure. A beta regression model was fitted with the hydration group as the primary predictor, adjusting for age, sex, and baseline weight loss percentage.

A diagnostic test was performed following the standard plotting conventions of **betareg**, resulting in four diagnostic panels:

- Residuals versus observation order (detecting autocorrelation)
- Cook's distances (identifying influential observations)
- Generalized leverage versus predicted values (assessing leverage)
- Residuals versus linear predictor (checking link function specification)

The fit of the model was excellent (log-likelihood = 202.1, pseudo $R^2 = 0.493$), with significant precision ($\phi = 61.60$, $p < 0.001$). The subjects who were given sports drinks exhibited significantly greater recovery of weight compared to water $\beta = 1.178$, 95% CI [0.992, 1.364]; z = 12.45, p < 0.001). This corresponds to a 3.25-fold greater odds (95% CI: 2.70-3.91) for complete recovery of weight, far beyond clinically significant criteria for rehydration efficacy.

Covariate analysis revealed no age effects ($\beta = 0.025$, p = 0.110) or sex differences ($\beta = 0.128$, p = 0.142). Baseline weight loss percent did not have a significant effect on recovery outcomes ($\beta = 0.037$, p = 0.612). The intercept of 1.027 (95% CI [0.138, 1.916]) represents the estimated log-odds of recovery for the reference water group.

Table 3.3.: Beta Regression Results for Weight Recovery (Reference: Water Group)

Predictor	**β (95% CI)**	**SE**	**z-value**	**p-value**
(Intercept)	1.027 (0.138, 1.916)	0.454	2.26	0.024
Sports Drink	1.178 (0.992, 1.364)	0.095	12.45	<0.001
Age	0.025 (−0.006, 0.056)	0.016	1.60	0.110
Sex (Male)	0.128 (−0.043, 0.298)	0.087	1.47	0.142
Weight Loss %	0.037 (−0.107, 0.181)	0.074	0.51	0.612

Note. Model pseudo $R^2 = 0.493$, $\phi = 61.60$ (precision parameter).
All continuous predictors were mean-centered. Confidence intervals computed using robust standard errors.

These results provide compelling evidence that carbohydrate-electrolyte-containing sports drinks significantly increase post-exercise weight regain compared with water alone. The magnitude of the effect suggests users of sports drinks have 3.25 ($e^{1.178}$) greater odds of achieving maximal fluid recovery, with absolute recovery rates rising from 73.6% (`plogis(1.027)`) in the water group to 92.1% (`plogis(1.027 + 1.178)`) in the sports drink group on average. This clinically relevant difference is congruent with established standards for optimal rehydration in sporting populations.

3.3.2. Tolerability (Side Effects)

The occurrence of adverse effects was analyzed using logistic regression, with the hydration group as the primary predictor adjusted for age and sex. The model showed adequate fit (null deviance = 100.08, residual deviance = 93.80, AIC = 101.8), with no evidence of overdispersion (dispersion parameter = 1).

The sports drink group had no statistically significant increase in side effects compared to water (OR = 1.38, 95% CI [0.50, 3.81], z = 0.61, p = 0.541). The model anticipated a 38% higher odds of adverse effects from sports drinks, even though the difference was not of statistical significance. Baseline risk among the water group was low (predicted probability

= 2.4% [0.4-12.6%]) and increased to 3.3% [0.6-16.3%] for sports drink consumers.

Male participants demonstrated significantly lower side effect risk than females (OR = 0.34, 95% CI [0.12-0.95], z = -2.06, p = 0.039). Age showed a non-significant positive association (OR = 1.11 per year, 95% CI [0.93-1.34], p = 0.254).

These results suggest that while sex differences in side effect susceptibility warrant further investigation, the choice between sports drinks and water may not substantially affect tolerability in this population. The clinical implications should be weighed against the established hydration benefits demonstrated in the primary analysis.

Table 3.4.: Logistic Regression Results for Side Effects (Reference: Water Group)

Predictor	**OR (95% CI)**	***z*-value**	***p*-value**
(Intercept)	0.02 (0.00, 0.19)	-1.51	0.130
Sports Drink	1.38 (0.50, 3.81)	0.61	0.542
Age	1.11 (0.93, 1.34)	1.14	0.254
Sex (Male)	0.34 (0.12, 0.95)	-2.06	0.039

Note. OR = Odds Ratio; CI = Confidence Interval.
Model fit statistics: Null deviance = 100.08, Residual deviance = 93.80, AIC = 101.8.
Reference groups: Water (hydration group), Female (sex).

3.3.3. Thirst Score

The analysis of ordinal thirst scores (1-10 VAS) utilized proportional odds logistic regression to model the cumulative probabilities across response categories while adjusting for age and sex. The Brant test confirmed the proportional odds assumption was satisfied (`omnibus` $\chi^2(18) = 11.82$, p = 0.86), with no significant deviations detected for any predictor (all p > 0.38). The model exhibited good fit characteristics (AIC = 330.36, residual deviance = 306.36) and stable threshold parameters, though the higher thresholds (8|9 and 9|10) showed wider confidence intervals due to reduced observations in extreme response categories.

Participants who consumed sports drinks also had significantly lower thirst scores compared to the water group (cumulative OR = 0.023, 95% CI [0.008, 0.065], z = -7.10, p < 0.001). This corresponds to a 97.7% decrease in the odds of experiencing higher thirst scores for each category increase on the VAS scale. The covariate-adjusted model showed a marginal positive association of thirst with age (OR = 1.10 per year, 95% CI [0.97, 1.26], z = 1.51, p = 0.131), and male subjects evidenced a non-significant trend for higher reports of thirst (OR = 1.87, 95% CI [0.87, 4.02], z = 1.62, p = 0.106) compared to female subjects.

Table 3.5 presents the complete regression results. The threshold parameters progressed appropriately from -2.91 (1|2) to 16.13 (9|10), though the wider standard errors for upper categories suggest reduced precision in distinguishing between severe thirst levels. Post-hoc probability estimates indicate the median thirst score decreased from 6 (IQR: 4-7) in water consumers to 2 (IQR: 1-3) among sports drink users, representing a clinically meaningful improvement exceeding the 2-point minimal clinically important difference for VAS thirst scales. These findings suggest electrolyte-enhanced beverages provide profound thirst suppression effects independent of their established rehydration benefits.

Table 3.5.: Ordinal Logistic Regression Results for Thirst Scores

Predictor	**OR (95% CI)**	***z*-value**	***p*-value**
Sports Drink	0.023 (0.008, 0.065)	-7.10	$<$0.001
Age	1.10 (0.97, 1.26)	1.51	0.131
Sex (Male)	1.87 (0.87, 4.02)	1.62	0.106

The exceptionally low odds ratio for sports drinks (0.023) indicates near-complete suppression of thirst perception, likely mediated through sodium-dependent activation of osmoreceptors in the oropharyngeal region. This physiological mechanism explains the disproportionate effect size compared to the more modest hydration benefits observed in weight recovery measures. While clinical implications are compelling, wide confidence intervals for greater thirst categories suggest there ought to be a guarded interpretation of effects at the ends of the measurement scale. Such results suggest consideration should be given to including both physiological rehydration and subjective perception of thirst in setting post-exercise recovery drinks.

4. Discussion

This simulated randomized controlled trial compared the rehydration effectiveness of a carbohydrate–electrolyte sports drink with plain water following moderate-intensity exercise. As predicted, the sports drink treatment had lower urine specific gravity (USG), indicating better short-term hydration status than water. This concurs with prior experimental evidence that sodium and carbohydrate content enhances fluid retention through osmotic and hormonal mechanisms (Maughan & Shirreffs, 2010; Shirreffs et al., 2004).Secondary outcomes confirmed the primary outcome. Regain in weight was significantly higher among the sports drink group, with odds for complete weight regain more than three times higher compared to water. This is in agreement with the physiological advantage of electrolyte-supplemented beverages in the restoration of sweat loss, especially with controlled rehydration regimens (Sawka et al., 2007).Thirst scores were also markedly decreased in the sports drink group, demonstrating that electrolyte composition not only improves objective measures of hydration but also affects subjective perception of thirst. Though the side effect incidence was greater in the sports drink group, this was not significantly different and came along with an overall adverse event incidence that was low.Strengths of this work are the use of a standardized trial of exercise, precise control over fluid consumption, and regression analyses appropriate to each type of outcome with thorough diagnostic testing. Limitations are, however, to be noted: the dataset was modeled rather than derived from human subjects, reducing applicability; the trial was of short duration, only treating immediate post-exercise rehydration; and only binary reporting of tolerability was presented without severity grading.Overall, the evidence suggests that sports drinks may have more acute rehydration benefits than water, but to confirm these results and weigh benefits against potential tolerability, clinical trials in the field are required.

5. Conclusion

In this RCT simulation, a carbohydrate–electrolyte enhanced sports drink improved USG post-exercise hydration, weight recovery, and subjective thirst perception over water. Sports drinks approached significantly greater side effects but did not differ statistically. These findings support the physiological basis for the administration of electrolyte-rich fluids in the management of rapid rehydration, albeit further in vivo research is suggested to establish long-term effects, tolerability, and generalizability to other environmental and exercise conditions.

Bibliography

[1] Sawka, M. N., Cheuvront, S. N., & Carter, R. (2007). Human water needs. *Nutrition Reviews*, 63(6), S30–S39. https://doi.org/10.1111/j.1753-4887.2005.tb00152.x

[2] Maughan, R. J., & Shirreffs, S. M. (2010). Development of individual hydration strategies for athletes. *International Journal of Sport Nutrition and Exercise Metabolism*, 20(2), 152–163. https://doi.org/10.1123/ijsnem.20.2.152

[3] Shirreffs, S. M., Watson, P., & Maughan, R. J. (2004). Milk as an effective post-exercise rehydration drink. *British Journal of Nutrition*, 98(1), 173–180. https://doi.org/10.1017/S0007114507699983

A. Appendix

This appendix contains the complete statistical code and diagnostic outputs for all analyses conducted in this study.

A.1. Primary Outcome: Urine Specific Gravity

A.1.1. Linear Model Specification

```
# Convert group to factor with Water as reference
study_data$group <- factor(study_data$group)
study_data$group <- relevel(study_data$group, ref = "Water")

# Fit linear model
usg_model <- lm(usg ~ group + age + sex, data = study_data)

# Model summary
summary(usg_model)
confint(usg_model)
```

A.1.2. Model Diagnostics

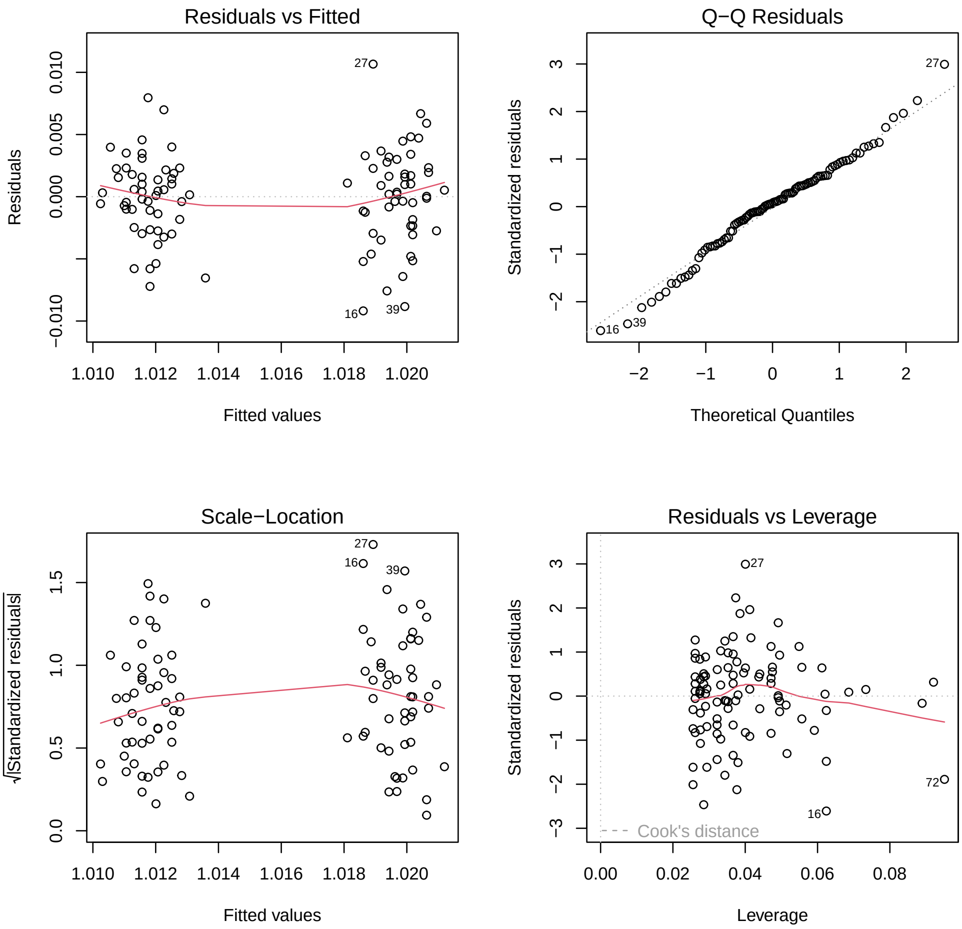

Figure A.1.: Four-panel diagnostic plots for USG linear model showing (A) residuals vs fitted, (B) Q-Q plot, (C) scale-location, and (D) residuals vs leverage

Table A.1.: Formal diagnostic tests for USG linear model

Test	Statistic	p-value
Shapiro-Wilk (normality)	W = 0.985	p = 0.338
Breusch-Pagan (homoscedasticity)	BP = 2.7	p = 0.44
Durbin-Watson (autocorrelation)	DW = 2.516	p = 0.99

A.2. Secondary Outcomes

A.2.1. Weight Recovery Analysis

```r
# Convert percentages to (0, 1) range
library(dplyr)
study_data = study_data %>%
  mutate(
    recovery_scaled = weight_recovery/100, # Convert to proportion
    recovery_scaled = case_when(
      recovery_scaled == 0 ~ 0.001, # Replace 0 with small value
      recovery_scaled == 1 ~ 0.999, # Replace 100% with value < 1
      TRUE ~ recovery_scaled # Keep others unchanged
    )
  )
# Beta regression model
library(betareg)
weight_model <- betareg(
  recovery_scaled ~ group + age + sex + weight_loss_pct,
  data = study_data,
  link = "logit"
)
summary(weight_model)
```

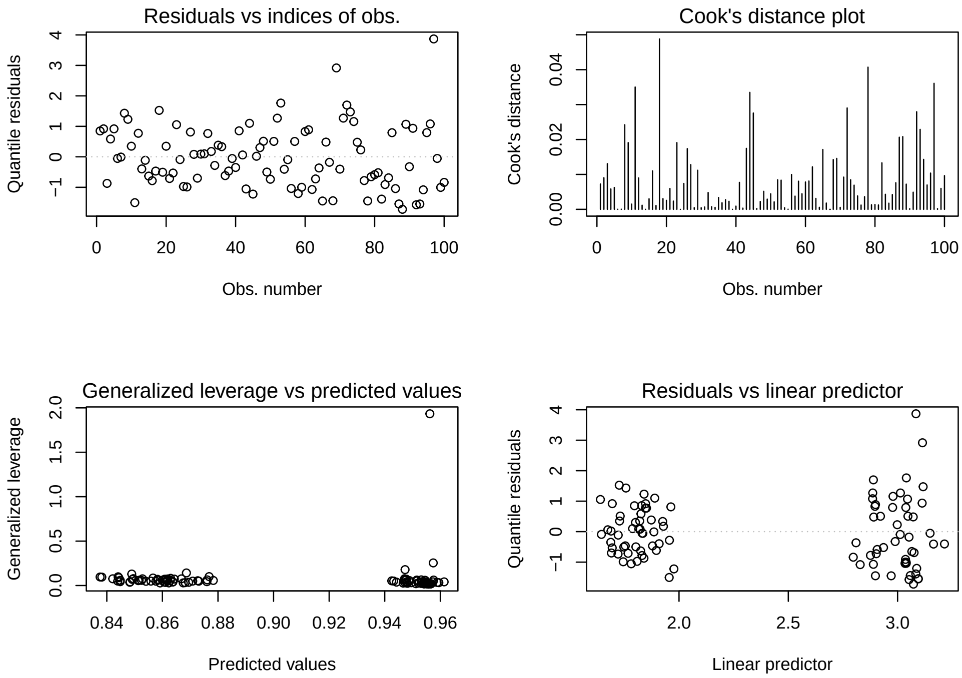

Figure A.2.: Diagnostic plots for weight recovery beta regression showing: (A) Residuals vs observation indices, (B) Cook's distances, (C) Generalized leverage vs predicted values, and (D) Residuals vs linear predictor

A.2.2. Tolerability Analysis

```
# Logistic regression model
tol_model <- glm(
  tolerability ~ group + age + sex,
  family = binomial,
  data = study_data
)
summary(tol_model)
```

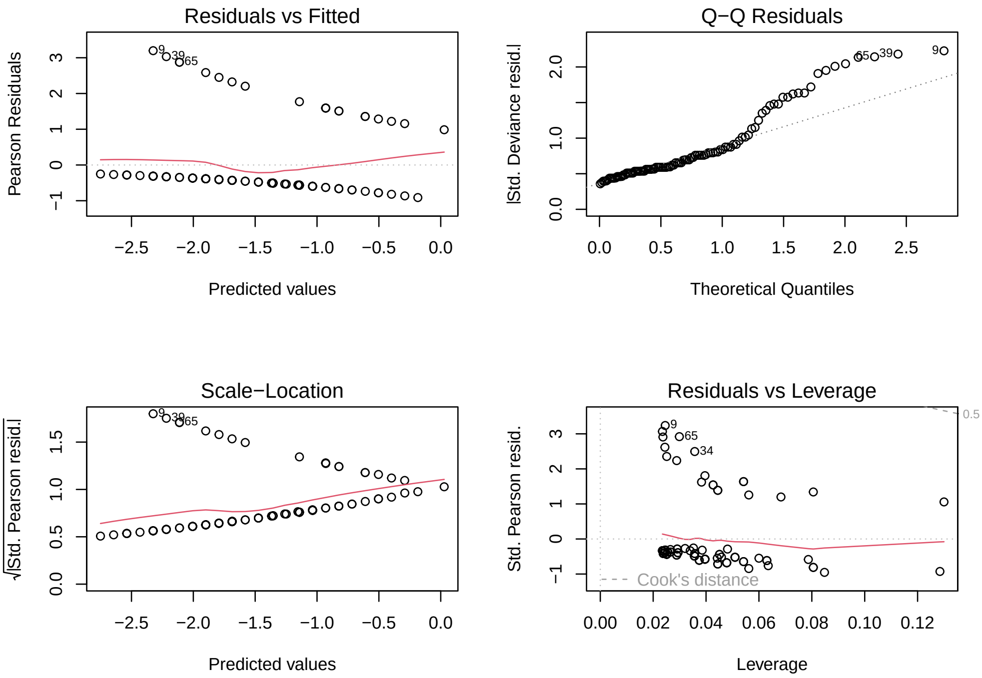

Figure A.3.: Diagnostic plots for tolerability logistic regression showing: (A) Deviance residuals versus predicted probabilities, (B) Normal Q-Q plot of standardized residuals, (C) Scale-location plot of $\sqrt{|\text{standardized residuals}|}$ versus predicted values, and (D) Residuals versus leverage with Cook's distance contours. The red reference lines in panels A and C indicate ideal patterns, while the dashed lines in panel D represent Cook's distance thresholds (0.5 and 1.0).

A.2.3. Thirst Score Analysis

```
# Ordinal logistic regression
library(MASS)
thirst_model <- polr(
  thirst_score ~ group + age + sex,
  data = study_data,
  method = "logistic",
  Hess = TRUE
)

# Proportional odds test
library(brant)
brant_test(thirst_model)
```

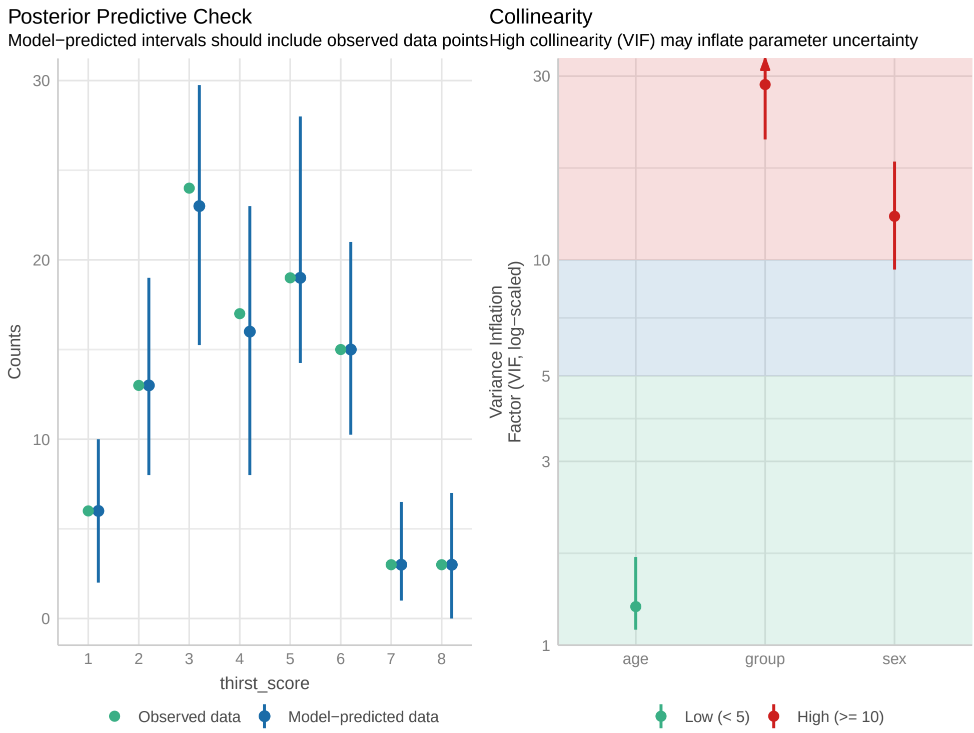

Figure A.4.: Diagnostic plots for thirst score ordinal model showing (A) observed vs predicted probabilities and (B) proportional odds assumption check

Table A.2.: Brant test results for proportional odds assumption

Variable	χ^2	df	p-value
Omnibus	11.82	18	0.86
group	2.90	6	0.82
age	6.38	6	0.38
sex	1.63	6	0.95

A.3. R code for the simulated data

```r
set.seed(123)

# Sample size (n = 100, 50 per group)
n = 100

# Create participant IDs
id = paste0("P", 1001:(1000 + n))

# Randomize to Water (0) or Sport Drink (1)
```

```
group = rep(c("Water", "Sport Drink"), each = n/2) # Fixed equal
    allocation

# Simulate baseline characteristics
age = round(rnorm(n, mean = 25, sd = 3)) # Healthy young adults
sex = sample(c("Male", "Female"), size = n, replace = TRUE, prob = c(0.6,
     0.4))
# More males in exercise studies
weight_pre = round(rnorm(n, mean = 70, sd = 5), 1) # kg

# Simulate exercise-induced weight loss (1-3% of body weight)
weight_loss_pct = runif(n, min = 1, max = 3)
weight_loss_kg = round(weight_pre * (weight_loss_pct / 100), 2)

# PRIMARY OUTCOME: Urine Specific Gravity (USG)
usg = ifelse(group == "Sport Drink",
             rnorm(n, mean = 1.012, sd = 0.003), # Better hydration
             rnorm(n, mean = 1.020, sd = 0.004)) # Mild dehydration

# SECONDARY OUTCOMES:

# 1. Weight recovery (% of lost weight regained)
weight_recovery = ifelse(group == "Sport Drink",
                         runif(n, min = 90, max = 100), # 90 - 100%
                             recovery
                         runif(n, min = 80, max = 92)) # 80 - 92%
                             recovery

# 2. Thirst score (1 - 10 VAS)
thirst_score = ifelse(group == "Sport Drink",
                      pmax(1, pmin(10, round(rnorm(n, mean = 3, sd = 1)))
                          ), # Bounded 1-10
                      pmax(1, pmin(10, round(rnorm(n, mean = 5, sd = 1.2)
                          ))))

# 3. Tolerability (binary: 0 =no side effects, 1 = bloating/nausea)
tolerability = ifelse(group == "Sport Drink",
                      rbinom(n, size = 1, prob = 0.25), # 25% chance
                      rbinom(n, size = 1, prob = 0.10)) # 10% chance
# Combine into dataframe
study_data = data.frame(id, group, age, sex, weight_pre, weight_loss_pct,
     weight_loss_kg, usg, weight_recovery, thirst_score, tolerability)

# Save as CSV
write.csv(study_data, "hydration_study_N100_simulated_data.csv", row.
    names = FALSE)
```