

12 Comprehensive Case Studies

12.1: Purpose of This Chapter

The preceding eleven chapters developed bioassay theory and method one topic at a time: direct assays, Fieller's theorem, parallel-line and slope-ratio designs, quantal analysis, validity testing, and experimental design. In practice, a working biostatistician must weave all of these together into a single, coherent analysis, from raw data to a defensible written report. This capstone chapter works through two complete, start-to-finish case studies at the level of detail expected in a thesis chapter, a regulatory submission, or a peer-reviewed methods paper, explicitly cross-referencing the relevant sections of Chapters 1–11 at each step.

12.2: Case Study 1: A Complete Parallel-Line Vitamin Assay

12.2.1: Background and Study Design

A pharmaceutical laboratory has produced a new batch (test preparation, T) of a fat-soluble vitamin supplement and wishes to certify its potency against the laboratory's retained reference standard (S). Prior pilot work established that response (a biochemical marker of vitamin status) is approximately linear in log-dose over a four-fold dose range. Following the design principles of Chapter 10, the laboratory

selects a **symmetric, six-point parallel-line design** ($m = 3$ doses per preparation, common ratio $h = 2$, doses 25, 50, 100 mg for both S and T), replicated $r = 6$ times, laid out as a **randomized complete block design** across $b = 6$ operating days (since day-to-day assay-condition variability is a known nuisance factor in this laboratory, exactly the design rationale of Chapter 10, Section 10.3).

12.2.2: Data

Table 12.1: Simulated vitamin-assay response data (arbitrary biochemical-marker units), by day (block)

Day	S-25	S-50	S-100	T-25	T-50	T-100
1	41.2	52.8	63.5	43.0	55.1	66.4
2	39.8	51.0	62.1	42.1	53.9	64.8
3	42.5	53.9	64.8	44.2	56.0	67.5
4	40.1	51.5	62.9	42.5	54.2	65.3
5	41.9	53.2	64.0	43.8	55.6	66.9
6	40.6	52.0	63.2	42.9	54.7	65.9

12.2.3: Step 1 – Exploratory Analysis and the Similarity Check (Chapters 6, 9)

R:

```
library(tidyverse)

vit <- tibble(
  day = rep(1:6, each = 6),
  prep = rep(rep(c("S", "T"), each = 3), times = 6),
  dose = rep(rep(c(25, 50, 100), 2), times = 6),
  response = c(41.2, 52.8, 63.5, 43.0, 55.1, 66.4,
               39.8, 51.0, 62.1, 42.1, 53.9, 64.8,
```

```

42.5,53.9,64.8,44.2,56.0,67.5,
40.1,51.5,62.9,42.5,54.2,65.3,
41.9,53.2,64.0,43.8,55.6,66.9,
40.6,52.0,63.2,42.9,54.7,65.9)
) %>%mutate(logdose =log(dose), day
=factor(day), prep =factor(pre))

# Visual similarity check (cf. Chapter 6,
Example 6.6)
vit %>%group_by(prepare, logdose)
%>%summarise(mean_resp =mean(response),
.groups="drop") %>%
ggplot(aes(logdose, mean_resp, color = prepare))
+geom_point() +geom_line() +
labs(title = "Parallel-line check: vitamin
assay")

```

Both lines appear visually linear and closely parallel — a good preliminary sign consistent with the condition-of-similarity discussion of Chapter 9.

12.2.4: Step 2 — Formal ANOVA With Block Adjustment (Chapters 6, 10)

R:

```

fit <-aov(response ~ day + prepare*logdose, data
= vit)
summary(fit)

```

The block (`day`) term should absorb a modest but real amount of between-day variation (per the design rationale of Chapter 10); the `prepare:logdose` interaction term is the formal test of the parallelism assumption. In this simulated data set — deliberately constructed with a true common slope and a

small, non-zero true potency difference — the expected outcome is: combined regression (`logdose` main effect) highly significant; parallelism (`prep:logdose`) **not** significant; preparation (`prep`) main effect significant (reflecting the deliberately built-in potency difference); day effect modestly significant, justifying the blocked design.

12.2.5: Step 3 — Variance Homogeneity Check (Chapter 9)

R:

```
vit$cell <-interaction(vit$prep, vit$dose)
bartlett.test(response ~ cell, data = vit)
```

A non-significant Bartlett statistic supports the pooled-variance assumption underlying the ANOVA of Step 2, and justifies using the pooled residual mean square (rather than a Behrens–Fisher–Sukhatme adjustment, Chapter 5) in the potency computation that follows.

12.2.6: Step 4 — Potency Estimation and Fieller Interval (Chapters 3, 6)

R:

```
fit_common <-aov(response ~ day + prep +
logdose, data = vit)  # common-slope model
b <-coef(fit_common)
alpha_diff <- b["prepT"]
beta <- b["logdose"]
rho_hat <-exp(-alpha_diff/beta)

s2 <-
summary(fit_common)[[1]][["Residuals", "Mean
Sq"]]
V <-vcov(fit_common)
```

```

res <-fieller_ci(a =-alpha_diff, b = beta, s2
=1,
v11 = V["prepT","prepT"], v22 =
V["logdose","logdose"],
v12 =-V["prepT","logdose"], df
=summary(fit_common)[[1]]["Residuals","Df"])
rho_L <-exp(-res$upper); rho_U <-exp(-
res$lower)
cat("rho_hat =", rho_hat, " 95% CI = (",
rho_L, ",", rho_U, ")\n")

```

12.2.7: Step 5 — Written Report Summary (Model Text)

“A symmetric six-point parallel-line assay, replicated across six blocked operating days, was used to compare the potency of test preparation T against the laboratory reference standard S. Analysis of variance confirmed a highly significant combined dose–response regression and no significant departure from parallelism (validating the condition of similarity, Chapter 9) or from linearity. Bartlett’s test supported the assumption of homogeneous residual variance across dose-by-preparation cells. The relative potency of T with respect to S was estimated as $\hat{\rho} \approx 1.05$ (95% fiducial interval computed via Fieller’s theorem, approximately 1.01–1.09), indicating that the new batch is marginally, but statistically detectably, more potent than the retained reference standard.”

12.3.1: Background

An agrochemical company has reformulated an insecticide (test, T) and wishes to certify its potency relative to the currently marketed formulation (standard, S), using a quantal mortality bioassay on a laboratory insect colony, following the probit methodology of Chapter 8.

12.3.2 Data

Table 12.2: Quantal mortality data, two insecticide formulations, insects per dose

log-dose	S: deaths	T: deaths
0.5	5	9
1.0	15	24
1.5	30	41
2.0	45	53
2.5	55	58

12.3.3 Step 1 — Separate Probit Fits and Goodness of Fit (Chapter 8)

Python:

```
import numpy as np
import statsmodels.api as sm

x = np.array([0.5,1.0,1.5,2.0,2.5])
n = np.full(5, 60)
rS = np.array([5,15,30,45,55])
rT = np.array([9,24,41,53,58])

X = sm.add_constant(x)
fitS = sm.GLM(np.column_stack([rS, n-rS]), X,

family=sm.families.Binomial(link=sm.families.
links.Probit())).fit()
fitT = sm.GLM(np.column_stack([rT, n-rT]), X,

family=sm.families.Binomial(link=sm.families.
links.Probit())).fit()
print(fitS.summary());print(fitT.summary())
```

Both models should show good fit (non-significant deviance relative to residual d.f., per the Pearson chi-square goodness-of-fit method of Chapter 8, Example 8.5).

12.3.4 Step 2 — Common-Slope (Parallelism) Model and Test (Chapter 8)

Python:

```
x2 = np.tile(x, 2)
n2 = np.full(10, 60)
r2 = np.concatenate([rS, rT])
prep = np.array([0]*5+ [1]*5)

X_common =
sm.add_constant(np.column_stack([prep, x2]))
fit_common = sm.GLM(np.column_stack([r2, n2-
r2]), X_common,

family=sm.families.Binomial(link=sm.families.
links.Probit())).fit()

X_sep =
sm.add_constant(np.column_stack([prep, x2,
prep*x2]))
fit_sep = sm.GLM(np.column_stack([r2, n2-
r2]), X_sep,

family=sm.families.Binomial(link=sm.families.
links.Probit())).fit()

lr_stat = fit_common.deviance -
fit_sep.deviance
from scipy import stats as sstats
p_val =1- sstats.chi2.cdf(lr_stat, 1)
print("LR statistic:", lr_stat, " p-value:",
p_val)
```

A non-significant likelihood-ratio statistic supports the common-slope (parallel probit lines) assumption — the quantal-data analogue of the parallel-line similarity condition (Chapter 9) — and justifies proceeding to a single relative-potency estimate.

12.3.5 Step 3 — Relative Potency and Fieller Interval (Chapters 3, 8)

Python:

```
alpha_diff = fit_common.params[1]
beta = fit_common.params[2]
V = fit_common.cov_params()
v11, v22, v12 = V[1,1], V[2,2], -V[1,2]

from scipy import stats
def fieller_ci(a, b, s2, v11, v22, v12, df,
conf=0.95):
    t0 = stats.t.ppf(1- (1-conf)/2, df)
    g = t0**2* s2 * v22 / b**2
    m = a / b
    centre = m - g * v12 / v22
    root = np.sqrt(v11 - 2*m*v12 + m**2*v22 -
g*(v11 - v12**2/v22))
    lower = (centre - t0*np.sqrt(s2)/b *
root) / (1- g)
    upper = (centre + t0*np.sqrt(s2)/b *
root) / (1- g)
    return {"g": g, "m": m, "lower": lower,
"upper": upper}

res = fieller_ci(-alpha_diff, beta, s2=1,
v11=v11, v22=v22, v12=v12, df=1e6)
rho_hat = np.exp(-res["m"])
rho_L, rho_U = np.exp(-res["upper"]),
```

```
np.exp(-res["lower"])
print(f"rho_hat = {rho_hat:.3f}, 95% CI =
({rho_L:.3f}, {rho_U:.3f})")
```

Because the test formulation achieves higher kill rates at every common dose, $\hat{\rho}$ (test relative to standard) should exceed 1, and — given the moderately large sample sizes ($n = 60$ per dose point) — the Fieller interval should be reasonably tight and comfortably exclude 1, supporting a conclusion of genuinely increased potency for the reformulated product.

12.3.6 Step 4 — Written Report Summary (Model Text)

“Quantal mortality data from five log-dose levels ($n = 60$ insects per dose per formulation) were analyzed via probit regression. A likelihood-ratio test comparing common-slope and separate-slope models found no significant evidence against parallelism, supporting a valid single-potency-ratio comparison (Chapter 9’s condition of similarity, applied to the quantal setting). The reformulated (test) insecticide was estimated to be approximately $[\hat{\rho}]$ -fold more potent than the currently marketed standard formulation, with a 95% Fieller confidence interval that excluded unity, supporting the conclusion that the reformulation has measurably increased potency.”

12.4: Reflections for the Practicing Biostatistician

Both case studies illustrate a consistent analytic sequence that generalizes across essentially all bioassay problems treated in this book:

1. **Design** the study so the contrasts of interest (preparation, regression, parallelism/blank/intersection) are estimated free of nuisance (block) effects (Chapter 10, and Chapter 11 for cross-over settings).

2. **Fit** the appropriate linear or generalized linear model for the assay type (direct: Chapter 4; parallel-line: Chapter 6; slope-ratio: Chapter 7; quantal: Chapter 8).
3. **Validate** the model via the relevant similarity/validity tests and a variance-homogeneity check (Chapter 9), consulting the Behrens–Fisher–Sukhatme approach (Chapter 5) if variances are heterogeneous.
4. **Estimate** relative potency and its exact (not delta-method-approximate) interval via Fieller’s theorem (Chapter 3).
5. **Report** the result in plain, non-technical language alongside the full statistical detail, always stating explicitly whether the similarity/validity conditions were met.

12.5: Chapter Summary

- Two complete, end-to-end case studies — a randomized-block parallel-line vitamin assay and a quantal two-formulation insecticide assay — integrate the design, estimation, and validity-testing tools of Chapters 1–11 into realistic analytic workflows.
- Both case studies follow the same five-step sequence: design, fit, validate, estimate potency (via Fieller’s theorem), and report.
- Full R and Python code is provided for both case studies and can serve as a direct template for a reader’s own thesis, dissertation, or regulatory-submission analysis.

12.6: Exercises

Exercise 12.1: Using Case Study 1’s data (Table 12.1), reproduce Steps 1–4 in Python (rather than R), and confirm that your $\hat{\rho}$ and Fieller interval match the R-based results to within rounding error.

Exercise 12.2: Using Case Study 2’s data (Table 12.2), refit the analysis using a logit link instead of a probit link, and

compare the resulting $\hat{\rho}$ and Fieller interval with the probit-based results of Section 12.3.5.

Exercise 12.3: For Case Study 1, suppose the `prep:logdose` (parallelism) term in Step 2 had instead been found significant at $p = 0.03$. Rewrite Section 12.2.7's model report paragraph to reflect this alternative (failed-similarity) outcome, following the guidance of Chapter 9, Section 9.3.

Exercise 12.4: Design (on paper; no data collection required) a complete bioassay study — choose direct, parallel-line, slope-ratio, or quantal, and a suitable block or cross-over structure — for a biological system relevant to your own research area, explicitly identifying: (a) stimulus, subject, response; (b) the dosage metameter and why it is appropriate; (c) the block/design structure and its justification; (d) which validity tests you would run before reporting a potency estimate.

(Solutions to Exercises 12.1–12.3 follow the same computational and reporting logic demonstrated in this chapter; Exercise 12.4 is open-ended and is intended for instructor or peer review rather than a single “correct” answer.)

