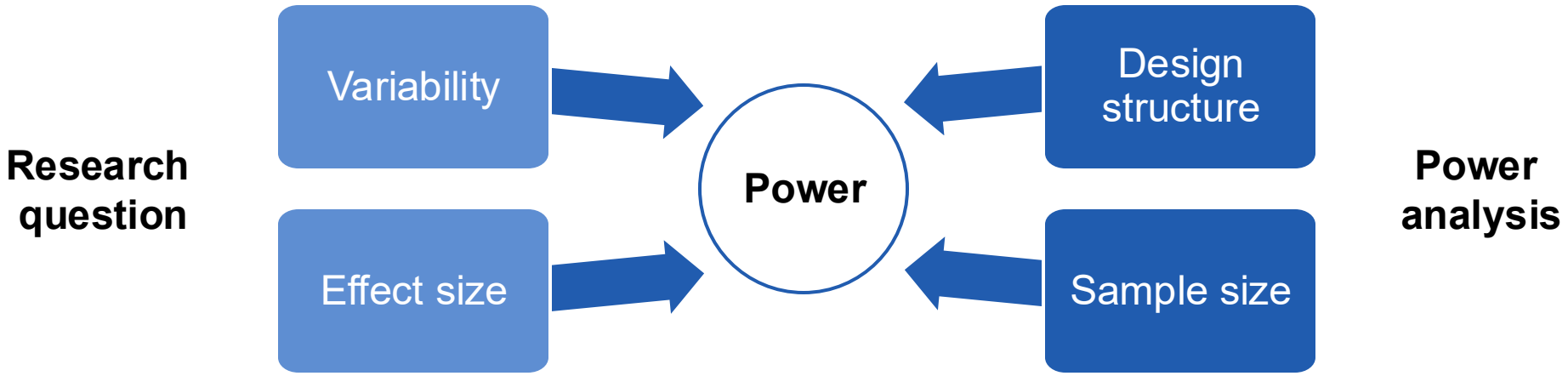
pwr4exp: An R package for robust statistical power analysis

Kai Wang*, Usman Arshad, Mutian Niu
Animal Nutrition, ETH Zurich
August 26, 2025

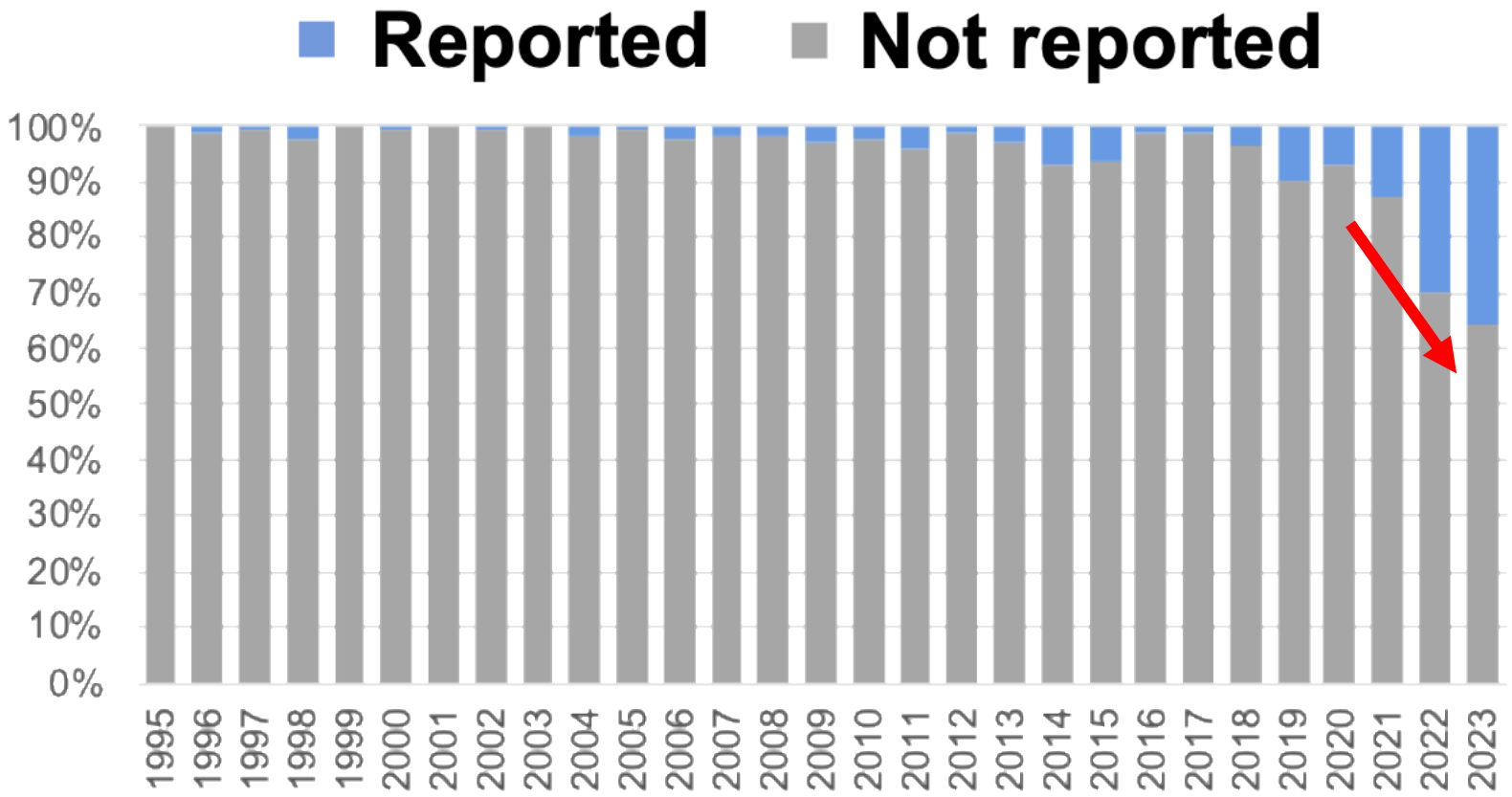
Power analysis is essential for hypothesis testing

	H_0 true	H_A true
Fail to reject H_0		Type II error ***False negative***
Reject H_0	Type I error ***False positive***	**Power**

- **In controlled experiment**
  - H_0: The treatment has no effect
  - **H_A: The treatment does have an effect**
- **Power: P[reject H_0 | H_A holds]**

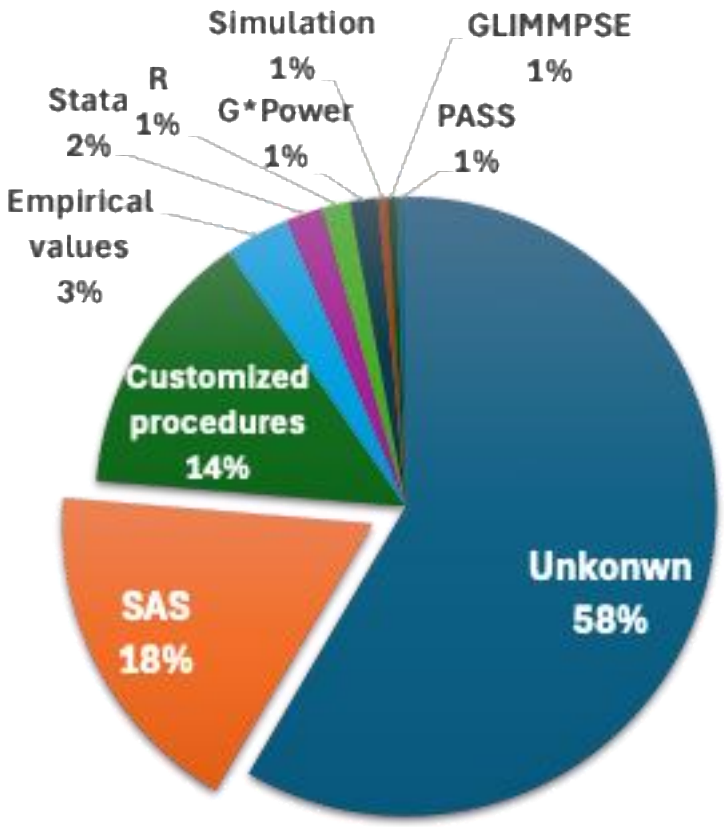


Power analysis was rarely reported in animal nutrition papers



- Literature review of 4,376 papers on JDS – Production: Animal Nutrition section
- **4.72%** reported power analysis

Challenges in conducting standardized power analysis



- **SAS** is most widely used
  - **The Stroup procedure** (1999, 2002) is versatile for many designs
  - No built-in **degrees-of-freedom approximation**

- **Most tools have limitations in:**
  - **Complex designs**: e.g., multiple blocking factors
  - **Treatment interactions**: e.g., higher-order effects
  - **Contrasts**: e.g., linear or polynomial contrasts
  - **Correlation structures**: e.g., repeated measures

Objective: to develop a comprehensive open-source power analysis tool

Theory background of the *pwr4exp* R Package

Model: $\boldsymbol{y = X\beta + Zu + \varepsilon} \qquad \boldsymbol{u} \sim N(\mathbf{0}, \boldsymbol{G}),\ \boldsymbol{\varepsilon} \sim N(\mathbf{0}, \boldsymbol{R})$

$$\boldsymbol{V} \coloneqq var(\boldsymbol{y}) = \boldsymbol{ZGZ}^T + \boldsymbol{R}$$

$$\boldsymbol{C} \coloneqq var(\boldsymbol{\beta}) = (\boldsymbol{X}^T\boldsymbol{V}^{-1}\boldsymbol{X})^{-1}$$

Test statistic

$$F = \frac{(\boldsymbol{K\hat{\beta}})^{\boldsymbol{T}}[\boldsymbol{K\hat{C}K}^{\boldsymbol{T}}]^{-\mathbf{1}}(\boldsymbol{K\hat{\beta}})}{ndf}$$

$ndf = \text{rank(K)}$

$\boldsymbol{ddf}$: **Satterthwaite approximation**

Distributions

$\boldsymbol{H_0}\colon\ \boldsymbol{F \sim F(ndf, ddf)}$

$\boldsymbol{H_A}\colon\ \boldsymbol{F \sim F(ndf, ddf, ncp)}$

$\boldsymbol{ncp = (K\beta)^T[KCK^T]^{-1}(K\beta)}$

Power calculation

$\boldsymbol{F_c}\colon\ \boldsymbol{P[F > F_c |\ H_0\colon F \sim F(ndf, ddf)] = \alpha}$

$\boldsymbol{power = P[F > F_c |\ H_A\colon F \sim F(ndf, ddf, ncp)]}$

Key steps in conducting power analysis using *pwr4exp*

Step 1: Design object

- Design structure
- Treatment structure
- Model parameters

- **Completely randomized design**
- **Randomized complete block design**
- **Latin Square design**
- **Split-plot design**
- **……**

Step 2: Power calculation

- **F test**
- **t test**

Start by creating a design object

What are the inputs?
- Know your data structure and model

$$\boldsymbol{y} = \boldsymbol{X\beta} + \boldsymbol{Zu} + \boldsymbol{\varepsilon}$$
$$\boldsymbol{u} \sim N(\mathbf{0}, \boldsymbol{G}),\ \boldsymbol{\varepsilon} \sim N(\mathbf{0}, \boldsymbol{R})$$

mkdesign()

```
mkdesign(
    formula,
    data,
    beta = NULL,
    means = NULL,
    vcomp = NULL,
    sigma2 = NULL,
    correlation = NULL,
    template = FALSE,
    REML = TRUE
)
```

- **The universal function for any design**
  - Model formula
  - Data frame

  ] **Design matrices: X, Z**
  - Fixed effects or means (β)
  - (Co-)variances components (G)
  - Residual variance (R),
    (optional) correlation structure (R)
    (imported from *nlme* package)

  ] **Model parameters under H_A**
  Educated assumptions
  - **Input templates for fixed and random effects parameters**

Built-in designs available in *pwr4exp*

- **Completely randomized design (CRD)**

  designCRD(treatments, label, replicates, formula, beta, means, sigma2, template = FALSE)

- **Randomized complete block design (RCBD)**

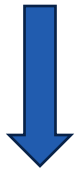
  designRCBD(treatments, label, blocks, formula, beta, means, vcomp, sigma2, template = FALSE)

- **Latin Square design (LSD)**

  designLSD(treatments, label, squares, reuse, formula, beta, means, vcomp, sigma2, template = FALSE)

- **Split-plot design (SPD)**

  designSPD(trt.main, trt.sub, label, replicates, formula, beta, means, vcomp, sigma2, template = FALSE)



A default model is set, but can be modified.

Power calculation with *F*- or *t*-test

Power of *F*-test

```
pwr.anova(
    object,
    sig.level,  # significance level
    type        # The type of SS
)
```

Power of *t*-test

```
pwr.contrast(
    object,
    which,       # The factor of interest
    by,          # The variable to condition on
    contrast,    # e.g., "pairwise", "poly", "trt.vs.ctrl", or a vector
    sig.level,   # significance level
    p.adj,       # whether or not using the Bonferroni method
    alternative, # "two.sided" or "one.sided"
    strict
)
```

Demo time!

Example – Latin Square design

- **4 × 4 Latin Square design with 2 × 2 factorial arranagement**
  - 16 cows (C), 4 periods (P), 4 squares
  - Temperature (Temp): T1 vs. T2
  - Dose: D1 vs. D2

	P1	P2	P3	P4
Cow 1	a	b	c	d
Cow 2	b	d	a	c
Cow 3	c	a	d	b
Cow 4	d	c	b	a

Expected means

Temp \ Dose	D1	D2
T1	35	38
T2	40	41

Treatments:

- a: T1D1
- b: T1D2
- c: T2D1
- d: T2D2

$$y_{ijkl} = \mu + Temp_i + Dose_j + (T \times D)_{ij} + C_k + P_l + e_{ijkl}$$

$$C_k \sim N(0, \sigma_c^2),\ P_l \sim N(0, \sigma_P^2),\ e_{ijkl} \sim N(0, \sigma_e^2)$$

Variance components

$\sigma_c^2 = 11, \sigma_P^2 = 2$

Residual variance

$\sigma_e^2 = 2$

Define the design and generate input templates

```
designLSD(
  treatments = c(2, 2),
  label = list(
    temp = c("T1", "T2"),
    dose = c("D1", "D2")),
  squares = 4,
  reuse = "col",
  template = TRUE
)
```

Two treatment factors each with 2 levels

(Optional) Labels of factor and factor levels

Four squares

Column blocks (periods) are identical acoss 4 squares

Generate the input templates for fixed and random effects

The default model: ~ *temp*dose + (1|row) + (1|col)*

Define the design and generate input templates

```
## $fixeff
## $fixeff$beta
##   (Intercept)         tempT2         doseD2 tempT2:doseD2
##             1              2              3              4
##
## $fixeff$means
## tempT1:doseD1 tempT2:doseD1 tempT1:doseD2 tempT2:doseD2
##             1              2              3              4
##
##
## $varcov
## $varcov$row
##             (Intercept)
## (Intercept)           1
##
## $varcov$col
##             (Intercept)
## (Intercept)           2
```

Fixed effects

Variance component

$\sigma_c^2 = 11$

$\sigma_P^2 = 2$

Input order of means

	D1	D2
T1	(1) 35	(3) 38
T2	(2) 40	(4) 14

Create the design object

```
lsd <- designLSD(
  treatments = c(2, 2),
  label = list(
    temp = c("T1", "T2"),
    dose = c("D1", "D2")),
  squares = 4,
  reuse = "col",
  means = c(35, 40, 38, 41),
  vcomp = c(11, 2),
  sigma2 = 2
)
```

Calculate power of *F*-test

```
pwr.anova(lsd)
```

```
## Power of type III F-test
##            NumDF DenDF sig.level   power
## temp           1    42      0.05 1.00000
## dose           1    42      0.05 0.99982
## temp:dose      1    42      0.05 0.78909
```

Calculate power of *t*-test

- **Compare T1 vs. T2 at each dose level**

```
pwr.contrast(lsd, which = "temp", by = "dose")
```

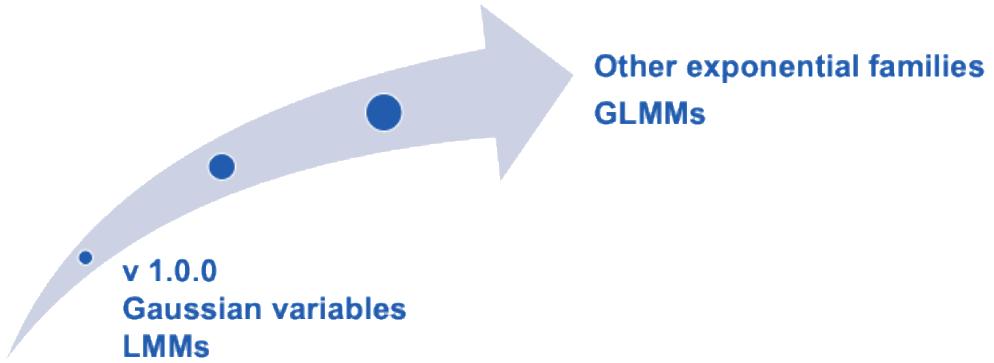
```
## $`dose = D1`
##                  effect df sig.level power alternative
## tempT1 - tempT2     -5 42      0.05     1   two.sided
##
## $`dose = D2`
##                  effect df sig.level     power alternative
## tempT1 - tempT2     -3 42      0.05 0.9999516   two.sided
```

- **Compare each treatment with control (T1D1)**

```
pwr.contrast(lsd, which = "temp:dose", contrast = "trt.vs.ctrl")
```

```
##                                 effect df sig.level     power alternative
## tempT2:doseD1 - tempT1:doseD1      5 42      0.05 1.0000000   two.sided
## tempT1:doseD2 - tempT1:doseD1      3 42      0.05 0.9999516   two.sided
## tempT2:doseD2 - tempT1:doseD1      6 42      0.05 1.0000000   two.sided
```

Some practical aspects



- **Plan the analysis up front**
  - Define the hypothesis and study objective
  - Select the appropriate statistical model
- **Collect prior information or conduct a pilot study**
  - Estimate treatment effects and variance components for the key response
- **Perform a sensitivity analysis**
  - Use confidence intervals instead of point estimates
  - Identify the parameter that most influences power (e.g., effect, variances, sample size)
- **Choose an adequate sample size to avoid under- and over-powering**
  - e.g., use smallest effect size or largest variance within the plausible regions (confidence interval)

Find out more

pwr4exp package page

Documentation & Bug report

J. Dairy Sci. 108:8981–9003
https://doi.org/10.3168/jds.2024-25793

Invited review: Enhancing quality of dairy cattle research through adequate power analysis

Kai Wang, Usman Arshad, and Mutian Niu*
ETH Zürich, Department of Environmental Systems Science, Animal Nutrition, Institute of Agricultural Sciences, Zürich 8092, Switzerland

ETH zürich

ETH Zürich
Animal Nutrition Group
Institute of Agricultural Sciences

Thank you

Sponsors:

Agroscope

SCHWEIZERISCHE VEREINIGUNG FÜR TIERWISSENSCHAFTEN
Association Suisse pour les Sciences Animales
Swiss Association for Animal Sciences