Inputs for Power Analysis: Literature Review

Course developed by
Deborah H. Glueck and Keith E. Muller

Slides developed by Jessica R. Shaw, Keith E. Muller, Albert D. Ritzhaupt and Deborah H. Glueck

 1

Learning Objectives

Identify the inputs for power or sample size analysis.

Describe how to search the literature for inputs.

2

UNDERSTAND KEY INPUTS

3

To design a study, specify four types of information

1. Design
2. Statistical test
3. Criterion
4. Five key inputs for power or sample size analysis

Kreidler et al., 2013
Munjal et al, 2013

4

Some inputs for power analysis are derived from the experimental setup

1. Design
2. Statistical test
3. Criterion
4. Five key inputs for power or sample size analysis

Kreidler et al., 2013
Munjal et al, 2013

5

Other inputs for power analysis are collected from external sources

1. Design
2. Statistical test
3. Criterion
4. Five key inputs for power or sample size analysis

Kreidler et al., 2013
Munjal et al, 2013

6

To perform calculations, specify the design, test, criterion, and inputs

Design describes the experiment.

Power, and thus sample size, depends on the choice of statistical test.

Criterion can be power or sample size.

Inputs provide information about the expected results of the experiment.

Kreidler et. al., 2013
Munjal et. al, 2013

7

Study design involves five inputs

1. Clustering
2. Predictors
3. Covariates
4. Repeated measures
5. Response

Kreidler et al., 2013
Munjal et al, 2013

8

Depending on the planned analysis, there are multiple statistical tests

Multivariate approach to repeated measures tests

1. Hotelling-Lawley Trace
2. Pillai-Bartlett Trace
3. Wilks' Lambda
4. Roy's largest root

Univariate approach to repeated measures tests

5. Uncorrected
6. Box
7. Geisser-Greenhouse
8. Huynh-Feldt

9

An analyst needs to specify a computational target

GLIMMPSE supports a computational target of power or sample size.

Choosing a specific sample size allows computing power.

Choosing a specific power allows computing a sample size.

10

Specifying the smallest group size implies the total sample size

Examples:

In a 1:1 randomization, if the smallest group size is 10, then the total sample size is 20.

In a 2:1 randomization design, if the smallest group size is 10, then the total sample size is 30.

11

Five key inputs are needed for power or sample size analysis

1. Predictors in model
2. Hypothesis
3. Hypothesized means and slopes (may suffice to specify only the differences)
4. Hypothesized standard deviations and correlations among measurements
5. Desired Type I error

12

The last three inputs are numeric

1. Predictors in model Hypothesis
2. Hypothesis
3. Hypothesized means and slopes (may suffice to specify only the differences)
4. Standard deviations and correlation between measurements
5. Desired type I error

13

In this lecture we will focus on reviewing the literature review to find values for the key numeric inputs of means, standard deviations, and correlations

Literature review is the systematic collection and examination of past research studies relevant to the present scientific goal.

14

UNDERSTAND HOW TO SEARCH LITERATURE FOR NUMERIC INPUTS

15

Literature review can help estimate unknown key numeric inputs for power and sample size calculation

Frequently unknown numeric inputs:

- Standard deviations
- Correlations between measurements
- Scientifically meaningful detectable difference

16

Literature review can also reveal standards for desired power

Standards may vary by field of study.

Power values of 0.80 are common.

We usually recommend 0.90 or 0.95.

The choice depends on ethical and scientific considerations.

17

Literature review can also reveal standards for Type I error

Standards may vary by field of study.

Type I error is often set to 0.05.

Sometimes, Type I error is set to 0.01.

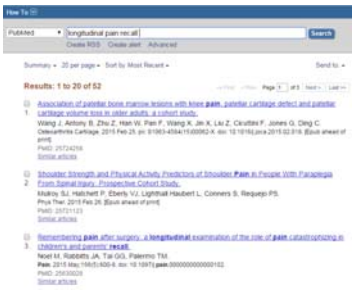
Bonferroni corrected Type I errors may be much smaller.

18

TOOLS FOR LITERATURE REVIEW

19

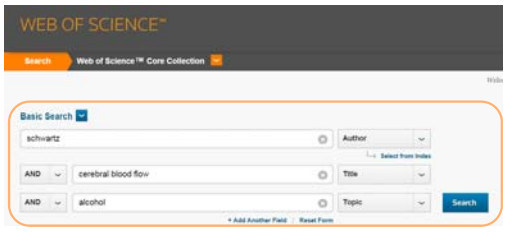
PubMed is a commonly used database interface for literature review



20

Web of Science is another useful tool for literature review

If you find a useful paper, you can search for it, and then find other papers that cited it.



21

Web of Science allows discovery of manuscripts similar to the planned study

Click on the Times Cited number to the right of the screen.

22

Examine review papers that have cited the original paper of interest

Review papers may be useful in your search for inputs for power analysis.

23

Plan your literature review and document your progress

Identify questions of interest, such as unknown correlations or standards of design in a given field.

Be careful to not reuse an inappropriate design or analysis, such as ignoring clustering.

Identify search terms.

Example: field of study + study design

Search for selected terms.

24

Carefully document the results of your search

You may choose to estimate inputs based on a group of studies in your field of research or one highly representative study.

Study	Author	Title	Field of study	Study design	Standard deviation	Correlation between measurements		
						3 months apart	6 months apart	12 months apart
1	Logan *et al.*	...	Pain recall	Longitudinal	0.98		Slightly higher than 0.4	0.4
2	...	...	Pain recall	Longitudinal	0.87	...	...	...
3	...	...	Pain recall	Longitudinal	1.02	...	...	...

25

We show an example literature review summary for numeric inputs

Vignette

Researchers propose a randomized controlled trial to compare pattern of weight gain over pregnancy for two treatment groups. The trial will enroll pregnant women at risk of gestational diabetes.

26

Vignette

The women will be randomized to either a dietary intervention designed to prevent excess weight gain, or to the standard of care. Weight will be measured at 6 time points during pregnancy.

27

Vignette

Analysis plan: repeated measures analysis of variance

Null: No difference in pattern of weight gain for two groups.

Planned test: Hotelling-Lawley test

Type I error rate: 0.05

Covariate: pre-pregnant BMI

28

After finding relevant studies, summarize key inputs found in each

Study	Trt. Group ($\bar{X}$, SE in kg)	Control Group ($\bar{X}$, SE in kg)	N (treat/ control)	Treatment
Quinlivan *et al.* (2007)	7 (0.65)	13.8 (0.67)	63/61	Continuity of care + weight tracking + 5 min diet counseling + mental health assessment
Wolff *et al.* (2008)	6.6 (1.1)	13.3 (1.4)	23/27	10 (1hr) diet consults throughout pregnancy
Thornton *et al.* (2007)	5.0 (0.6)	14.06 (0.7)	116/116	1 visit with RD + daily food log throughout pregnancy

Note: Standard error, not standard deviation. Always determine which you have.

29

Summarize unpublished data from one of your own similar studies

The unpublished study provided an estimate of correlations between measurements.

Correlations are rarely published, which makes estimating correlations for power analysis difficult.

30

Summarize unpublished data from one of your own similar studies

	HS obese women Mean (SD)
Pre-pregnant weight (kg)	95.23 (16.76)
Last weight (kg)	105.05 (17.65)
Correlation	0.87
Predicted weight for 39 weeks (kg)	106.22 (17.50)
Correlation	0.89

31

Notice that there are multiple means and standard deviations in the table

	HS obese women Mean (SD)
Pre-pregnant weight (kg)	95.23 (16.76)
Last weight (kg)	105.05 (17.65)
Correlation	0.87
Predicted weight for 39 weeks (kg)	106.22 (17.50)
Correlation	0.89

32

To account for uncertainty in the power and sample size calculation, a researcher may:

Use a power curve.

Consider many different values of mean differences close to the expected difference.

Consider different values of standard deviation.

Use the experimental situation to guide the choice of a conservative or liberal estimate.

33

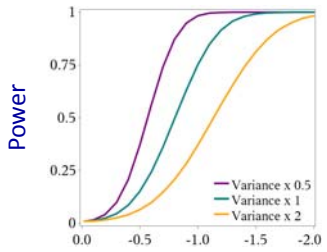
The next several images show how to incorporate uncertainty from estimates in power curves

One can draw several curves, one for each variance estimate.

For a given power, one can look at the range of mean differences one could detect for difference variance estimates.

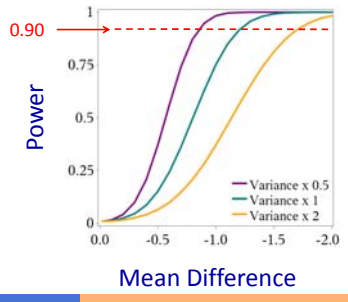
34

Accounting for uncertainty



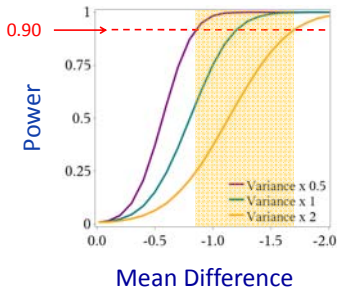
35

Accounting for uncertainty



36

Accounting for uncertainty



37

Be careful about using published numbers

One previous table included standard errors, and one included standard deviations.

Standard errors are smaller, standard deviations are larger.

Power analysis will be wrong if one mistakes standard errors for standard deviations, or vice versa.

38

PRACTICE

39

For practice, we obtain inputs for sample size analysis for the memory of pain study

Vignette

Researchers conducted a study to determine if dental patients who are instructed to use a sensory focus have a different pattern of long-term memory of pain than participants who did not.

Logan, Baron and Kohout, 1995

40

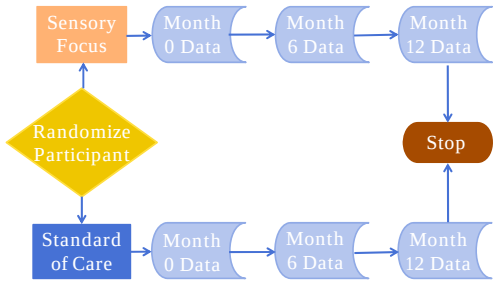
Vignette, continued

Participants were selected and randomly assigned to either intervention or no intervention. Those in the intervention group listened to automated audio instructions to pay close attention only to the physical sensations in their mouth. Participants in the no intervention group listened to automated audio instruction on a neutral topic to control for media and attention effects.

Logan, Baron and Kohout, 1995

41

Observed pain after root canal was measured at 0, 6, and 12 months



Logan, Baron and Kohout, 1995

42

Here, the interest is in an interaction hypothesis: a between-by-within hypothesis

No time by treatment interaction

43

Some of the information for the study is known already

Design: repeated measures analysis of variance

Analysis plan: Fit general linear multivariate model with three repeated measurements of memory of pain as outcome, and indicator variables for treatment as predictors

44

Some of the information for the study is known already

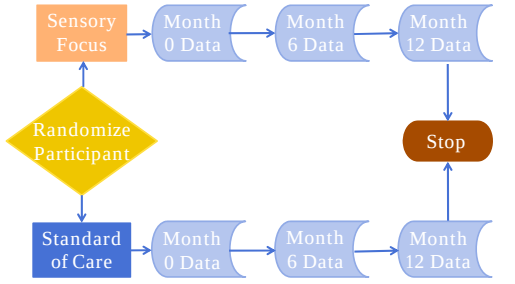
Statistical test: Hotelling-Lawley test

Criterion: sample size

Type I error: 0.01

45

It remains to summarize the numeric inputs for the sample size calculation



Logan, Baron and Kohout, 1995

46

We must search the literature to obtain the following numeric inputs for the sample size calculation

Hypothesized difference in means

Standard deviations and correlation between measurements

Kreidler et. al., 2013
Munjal et. al, 2013

47

Published articles contain hints for numeric inputs

Correlation Between Outcomes Over Time

Gedney, Logan, and Baron (2003) identified predictors of the amount of experienced pain recalled over time...One of the findings was that memory of pain intensity at 1 week and 18 months had a correlation of 0.4. ...assume that the correlation between measures 18 months apart will be similar to the correlation between measures 12 months apart. Likewise, the correlation between measures 6 months apart will be only slightly greater than the correlation between measures 18 months apart.

Logan et al, 2009

48

Published articles contain hints for numeric inputs

Correlation Between Outcomes Over Time

Gedney, Logan, and Baron (2003) identified predictors of the amount of experienced pain recalled over time...One of the findings was that memory of pain intensity at 1 week and 18 months had a correlation of 0.4... assume that the correlation between measures 18 months apart will be similar to the correlation between measures 12 months apart. Likewise, the correlation between measures 6 months apart will be only slightly greater than the correlation between measures 18 months apart.

Logan et al, 2009

49

Published articles contain hints for numeric inputs

Correlation Between Outcomes Over Time

Gedney, Logan, and Baron (2003) identified predictors of the amount of experienced pain recalled over time...One of the findings was that memory of pain intensity at 1 week and 18 months had a correlation of 0.4... assume that the correlation between measures 18 months apart will be similar to the correlation between measures 12 months apart. Likewise, the correlation between measures 6 months apart will be only slightly greater than the correlation between measures 18 months apart.

Logan et al, 2009

50

Using the previous paragraph, we can identify correlation

Correlation at 6 months apart

Correlation at 12 months apart

Logan et al., 2009

51

Using the previous paragraph, we can identify our power inputs

Correlation at 6 months apart

Correlation at 12 months apart

0.4

Logan, Henrietta L., Aarti Munjal, Brandy M. Ringham, and Deborah H. Glueck

52

We must estimate certain inputs based on information available

Correlation Between Outcomes Over Time

Gedney, Logan, and Baron (2003) identified predictors of the amount of experienced pain recalled over time...One of the findings was that memory of pain intensity at 1 week and 18 months had a correlation of 0.4.. assume that the correlation between measures 18 months apart will be similar to the correlation between measures 12 months apart. Likewise, the correlation between measures 6 months apart will be only slightly greater than the correlation between measures 18 months apart.

Logan et al, 2009

53

We must estimate certain inputs based on information available

Correlation Between Outcomes Over Time

Gedney, Logan, and Baron (2003) identified predictors of the amount of experienced pain recalled over time...One of the findings was that memory of pain intensity at 1 week and 18 months had a correlation of 0.4.. assume that the correlation between measures 18 months apart will be similar to the correlation between measures 12 months apart. Likewise, the correlation between measures 6 months apart will be only slightly greater than the correlation between measures 18 months apart.

Logan et al., 2009

54

Interpret the phrase "slightly greater" to estimate the correlation at 6 months

Correlation at 6 months apart

[ ]

Correlation at 12 months apart

0.4

Logan et al., 2009

55

Interpret the phrase "slightly greater" to estimate the correlation at 6 months

Correlation at 6 months apart

0.5

Correlation at 12 months apart

0.4

Logan et al., 2009

56

Reading further, we identify the standard deviation of measurements

Standard Deviation of the Outcome

Logan, Baron, and Kohout (1995) examined whether sensory focus therapy during a root canal procedure could reduce a patient's experienced pain. The investigators assessed experienced pain on a 5 point scale both immediately and at one week following the procedure. The standard deviation of the measurements was 0.98.

Logan et al., 2009

57

Reading further, we identify the standard deviation of measurements

Standard Deviation of the Outcome

Logan, Baron, and Kohout (1995) examined whether sensory focus therapy during a root canal procedure could reduce a patient's experienced pain. The investigators assessed experienced pain on a 5 point scale both immediately and at one week following the procedure. The standard deviation of the measurements was 0.98.

Logan et al., 2009

58

Information from published articles provides numeric inputs

Standard deviation of memory of pain

0.98

Logan et al., 2009

59

Published means may give suggestions for the treatment difference

Treatment	Baseline	6 Months	12 Months
Sensory Focus (SF)	3.6	2.8	0.9
Standard of Care (SOC)	4.5	4.3	3.0
Treatment Difference (SF - SOC)	-0.9	-1.5	-2.1
Net Difference Over Time (12 Months - Baseline)			-1.2

Logan et al., 2009

60

Scientifically important differences usually provide the best choice

Health scientists typically have a good sense of what size difference is important, and what size difference is not important. Extended conversation may be needed to probe the knowledge of the investigators and stimulate them to choose. Using published values can create problems discussed in later lectures.

Logan et al., 2009

61

Conduct simple subtraction to calculate the treatment difference

Treatment	Baseline	6 Months	12 Months
Sensory Focus (SF)	3.6	2.8	0.9
Standard of Care (SOC)	4.5	4.3	3.0
Treatment Difference (SF - SOC)	-0.9	-1.5	-2.1
Net Difference Over Time (12 Months - Baseline)			-1.2

Logan et al., 2009

62

Conduct simple subtraction to calculate the treatment difference

Treatment	Baseline	6 Months	12 Months
Sensory Focus (SF)	3.6	2.8	0.9
Standard of Care (SOC)	4.5	4.3	3.0
Treatment Difference (SF - SOC)	-0.9	-1.5	-2.1
Net Difference Over Time (12 Months - Baseline)			-1.2

Logan et al., 2009

63

Conduct simple subtraction to calculate the treatment difference

Treatment	Baseline	6 Months	12 Months
Sensory Focus (SF)	3.6	2.8	0.9
Standard of Care (SOC)	4.5	4.3	3.0
Treatment Difference (SF - SOC)	-0.9	-1.5	-2.1
Net Difference Over Time (12 Months - Baseline)			-1.2

Logan et al., 2009

64

Net difference is the difference between treatment differences over 12 months; interaction effect

Treatment	Baseline	6 Months	12 Months
Sensory Focus (SF)	3.6	2.8	0.9
Standard of Care (SOC)	4.5	4.3	3.0
Treatment Difference (SF - SOC)	-0.9	-1.5	-2.1
Net Difference Over Time (12 Months - Baseline)			-1.2

Logan et al., 2009

65

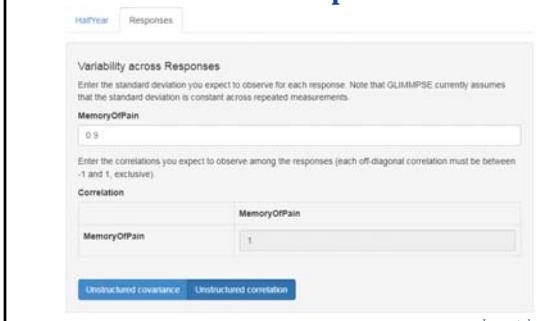
Having identified all inputs, specify each within GLIMMPSE

HalfYear	1	2	3
Response	MemoryOfPain	MemoryOfPain	MemoryOfPain
Treatment			
SensoryFocus	3.6	2.8	0.9
StandardOfCare	4.5	4.3	3

Logan et al., 2009

66

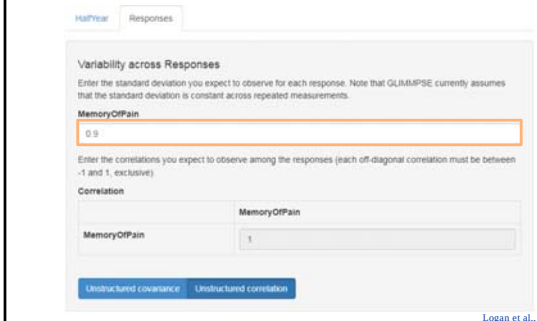
Input your expected standard deviation for each response



Logan et al., 2009

67

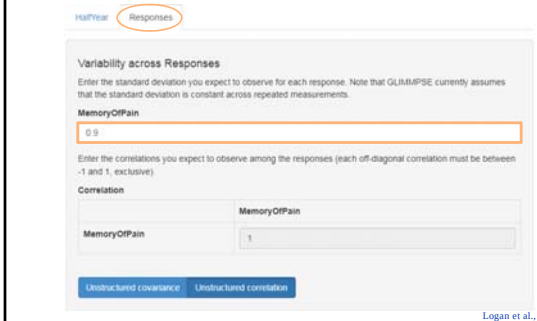
Input your expected standard deviation for each response



Logan et al., 2009

68

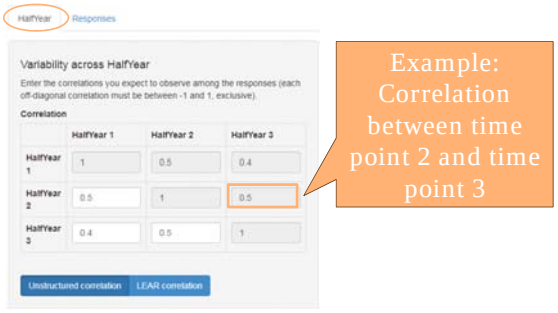
Input your expected standard deviation for each response



Logan et al., 2009

69

Specify correlations between repeated measures



Example: Correlation between time point 2 and time point 3

Logan et al., 2009

70

REVIEW OF LEARNING OBJECTIVES

71

Briefly discuss why a literature review is a key step in power and sample size calculation

Literature review is an important step in power and sample size calculation because it provides insight into unknown parameters such as standard deviations, correlations between measurements, and differences of interest.

72


Questions?
73