Improving experimental design:

Ethical implications and how the Experimental Design Assistant (EDA) can help



September 14, 2023

Esther Pearl, PhD Programme Manager for Experimental Design, National Centre for the Replacement, Refinement, and Reduction of Animals in Research (the NC3Rs).



National Centre for the Replacement Refinement & Reduction of Animals in Research

Improving experimental design: Ethical implications and how the Experimental Design Assistant (EDA) can help

Dr Esther Pearl

14 September 2023

National Institutes of Health Office of Laboratory Animal Welfare Online Seminar

Pioneering Better Science

www.nc3rs.org.uk



National Centre for the Replacement Refinement & Reduction of Animals in Research

Who we are \checkmark 0

Our portfolio Our

Our funding schemes \lor 3Rs

3Rs resource library V

Q

3Rs for the public

News

Events

Sign in (2)

Ø

PIONEERING BETTER SCIENCE

Replacement, Reduction and Refinement

A UK-based scientific organisation dedicated to helping the research community worldwide to identify, develop and use 3Rs technologies and approaches.

Our mission

Learning objectives



By the end of this talk you will:

Understand why we need to improve experimental design

Identify key ways to improve biomedical research

 Know what the Experimental Design Assistant (EDA) is and its benefits to researchers

 Understand how the EDA can be used as part of the ethical review process



Poll



Introductory poll

- Are you involved in designing experiments?
- Have you had any formal training in experimental design?
- Do you have access to statistical support?
- Do you sit on an IACUC?
- Do you conduct your own research?



Why we need to improve experimental design



Research waste



- 33 meta-studies in ecology
- 10,464 studies

THE LANCET

VIEWPOINT | VOLUME 374, ISSUE 9683, P86-89, JULY 04, 2009

Avoidable waste in the production and reporting of research evidence

Iain Chalmers, DSc 🛛 Rac Prof Paul Glasziou, RACGP

Published: June 15, 2009 • DOI: https://doi.org/10.10.

85% of investment in

biomedical research wasted



82% wasted

Purgar M, Klanjscek T and Culina A (2022). Quantifying research waste in ecology. *Nat Ecol Evol.* <u>http://dx.doi.org/10.1038/s41559-022-01820-0</u>

7

Research waste in *in vivo* studies Ethical implications

If research is not reported in enough detail, or if findings are not reliable, benefits cannot be realised

\rightarrow Research is unethical



The Economist World politics Business & finance Economics Science & technology Culture Unreliable research Trouble at the lab

rouble at the lab

Scientists like to think of science as self-correcting. To an alarming degree, it is not Dct 19th 2013 | From the print edition
© Trimekreper





Likely benefits to science and society

- New scientific knowledge
- Improvements in human (or animal) health or safety



Likely harms to the animals involved

- Scientific procedures and their effects
- Contingent suffering due to housing, transport, etc.



Contemporary definitions of the 3Rs

	Standard	Contemporary
Replacement	Methods which avoid the use of animals (or species 'protected' under law) in areas where they otherwise would have been used	Accelerating the development and use of models and tools, based on the latest science and technologies, to address important scientific questions without the use of animals (i.e. <i>in vitro</i> or <i>in silico</i>)
Reduction	Methods which minimise the number of animals used per experiment (or maximise the information gained from a given number of animals)	Appropriately designed and analysed animal experiments that are robust and reproducible, and truly add to the knowledge base
Refinement	Improvements to scientific procedures and husbandry which minimise pain, suffering, distress or lasting harm and/or improve animal welfare	Advancing research animal welfare by exploiting the latest <i>in vivo</i> technologies and by improving understanding of the impact of welfare on scientific outcomes



Resources from the NC3Rs



The National Academies of SCIENCES · ENGINEERING · MEDICIN

CONSENSUS STUDY REPO

PROCEEDINGS OF A WORKSHOP

Reproducibility and Replicability in Science

Ne

0

Enhancing Scientific Reproducibility in Biomedical Research Through **Transparent Reporting**

> The National Academies of SCIENCES · ENGINEERING · MEDICINE



Reproducibility and reliability of biomedical research: improving research practice

Symposium report, October 2015





Key ways to improve biomedical research



Key aspects of good experimental design

If experiments are not rigorously designed the results are unlikely to be reliable.

Aspects of good experimental design you can look out for:

- Randomisation
- Masking (blinding)
- Using both sexes
- Appropriate sample size
- Pre-planned statistical analysis method







Reporting of rigour criteria in animal research publications

NC 3R^s

Menke J et al (2020). iScience. https://doi.org/10.1016/j.isci.2020.101698

Randomisation



How the randomisation sequence is generated is important – haphazard is not random. Random sequences can be generated by computers, dice or flipping coins.

Pick a number between 1 and 10



http://scienceblogs.com/cognitivedaily/2007/02/05/is-17-the-most-random-number/

Randomisation



How the randomisation sequence is generated is important – haphazard is not random. Random sequences can be generated by computers, dice or flipping coins.

Random allocation to interventions

Randomisation is crucial for two reasons:

1. Minimise selection bias

e.g. haphazard selection may result in slowest mice allocated to the same group

 Key assumption of the statistical analysis
 Different groups should be drawn from the same background population using random sampling



Randomisation

- Animal models of multiple sclerosis
- Comparison of randomised and non-randomised studies



Studies not randomised overestimate treatment efficacy



Masking



Being unaware of group assignment (e.g. not knowing which intervention any animal has received, how they are grouped or which group data came from).

Masking can be implemented at different steps in an experiment

During allocation and intervention

Steps where the animals are assigned to experimental groups and steps when they receive the experimental intervention(s).

During conduct of the experiment

Housing and welfare management of the animals during the experiment.

When assessing the outcome

Steps where an outcome is measured, or a sample processed in preparation for a measurement.

When assessing results

Data processing and statistical analysis.

Reducing bias

157 veterinary medicine students

Trained to identify positive and negative social interactions in pigs

5-minute video clips

NC 3R^s

Control pigs, and 'High Social Breeding Value' pigs

Masking





Reducing bias

157 veterinary medicine students

Trained to identify positive and negative social interactions in pigs

5-minute video clips

Control pigs, and 'High Social Breeding Value' pigs

Masking





Tuyttens et al (2014). Ani Behav. doi: 10.1016/j.anbehav.2014.02.007

GUIDANCE

Using masking/blinding in in vivo experiments

On this page

- Why use masking in an animal experiment?
- How to make it work in practice?
- Examples of masking strategies for animal experiments
- References

Why use masking in an animal experiment?

Masking (also known as blinding) is a methodological process where an animal's allocation to a specific experimental group is concealed from the people running the study, caring for the animals or analysing the data.



Read the full article in PLOS Biology

A qualitative study of the barriers to using blinding in in vivo experiments and suggestions for improvement



https://nc3rs.org.uk/3rs-resources/using-maskingblinding-vivo-experiments Karp NA et al (2022). *PLOS Biol.* https://doi.org/10.1371/journal.pbio.3001873

Using both sexes



Most research is conducted on male subjects.

We cannot assume that findings in males translate to females.

Why it is important to use both sexes in preclinical research

- Issue recognised for clinical trials, which now include women in the research pipeline
- But limited change in preclinical research, despite role in informing clinical trials
- Example of Thalidomide teratogenic effects could have been predicted with in vitro testing of female human tissue
- Funders actively pushing the inclusion of sex as a biological variable (e.g. NIH)



Sex bias in animal experiments



b Choice and reporting of animal subjects in the journal *Pain*



Sex bias in animal experiments



b Choice and reporting of animal subjects in the journal *Pain*



Using both male and female animals https://eda.nc3rs.org.uk/experimental-design-animal-characteristics

In most cases researchers should include both sexes in their protocol

It should be clear from the protocol what researchers are trying to do

- Are they including both sexes to ensure results are generalisable?
- Are they directly comparing the sexes? Do they have adequate numbers to make this comparison? Will they need to include more animals?
- If the researcher is not directly comparing the sexes they may not need to increase animal numbers
- Suggest the researcher consults a statistician



Animal Characteristics



The characteristics of the animals used can have a large influence on both the results and the analysis methods used to analysis those results



Sample size



Using an appropriate sample size is essential to the reliability of results.

The number of animals in an experiment must be justified

If data will be compared with a statistical test (e.g. tests that give a pvalue or F statistic) the sample size should be determined with a power calculation.





Sample size



Using an appropriate sample size is essential to the reliability of results.

The number of animals in an experiment must be justified

If data will not be compared with statistical tests it is still important to justify the number of animals needed for the study.

For example, if the study aims to establish if a new surgical technique works in mice the number of animals needed depends on how many animals the researcher needs to ascertain if the technique would work and be practical for the future applications (e.g. a further study).



Analysis plan



It is important that a statistical analysis plan is made before the study starts

Analysis plans are an integral part of experimental design.

- Promote careful consideration of variables and outcome measures
- Identify appropriate analysis path (inferential vs descriptive statistics)
- Helps prevent p-hacking
- Identify statistical support needs



The EDA: benefits for researchers and IACUCs



Experimental Design Assistant (EDA)

Online tool for researchers to design *in vivo* experiments

Free to use

Secure

Focuses on internal validity

Developed by *in vivo* researchers and statisticians

eda.nc3rs.org.uk



The Experimental Design Assistant

A free resource from the NC3Rs used by over 12,000 researchers worldwide to help design robust experiments more likely to yield reliable and reproducible results.

The EDA helps you build a diagram representing your experimental plan, which can be critiqued by the system to provide **bespoke feedback**. The EDA also:

- · Recommends statistical analysis methods
- Provides support for randomisation and blinding
- · Performs sample size calculations

For an overview of how the EDA works, watch our one minute video.



The EDA website also provides information about the different concepts of experimental design, and how to apply these in your experiments.





Features of the EDA

- Advice to improve the experimental plan
- Recommendations for the statistical analysis
- Power calculation
- Randomisation and masking
- Improve transparency
- Teach experimental design





FDΔ





The EDA diagram



 \times

https://eda.nc3rs.org.uk/RT0011

The diagram does not include any nuisance variables

In the EDA, nuisance variable nodes are used to indicate other sources of variability or con influence the outcome, such as blocking factors or covariates.

Having no nuisance variable implies that only the independent variable(s) of interest (e.g. th influence the outcome measure. Thus it is rarely appropriate as the result of an experiment by many variables. Identifying these nuisance variables and accounting for them increases experiment to detect changes induced by the variable(s) of interest.

The type of things to consider may include cages or rooms, if the animals are not all housed of the intervention or measurement, or the person doing it if animals are not all processed the experimenters with different levels of skills. The list could be endless but the important thing relevant to a particular experiment, based on common sense and past experimental results, to identify new sources of variability.

These should be indicated on the EDA diagram using nuisance variable nodes; then the us best to account for each of the nuisance variables identified. Depending on the type of nuisa objective of the experiment, there are different options to account for the variability; the variable variable variables are used to be account for the variability of the variability.

- · standardised for example record all measurements on the same piece of equipment
- randomised across for example the effect of the location of the cage in the room co randomly allocating each cage to the housing rack
- blocked for example the day of the experiment can be used as a blocking factor in analysis
- nested within another variable for example individual neurons can be nested within 'mouse' when multiple neurons are recorded for each mouse
- · used as a covariate for example baseline locomotor activity could be included as a
- · if none of these things are done, then the variable is deemed uncontrolled.



×

👫 Indicate the blinding status during assessment of the outcome | NC3Rs EDA - Google Chrome

https://eda.nc3rs.org.uk/RT0093

Indicate the blinding status during assessment of the outcome

Information crucial to the internal validity of the experiment is missing. In the properties of this node, in the field 'blind measurement' please indicate whether the experimenter will be aware of the group allocation when assessing the results. Note that this only concerns the measurement stage; blinding before, during and after the intervention, and during the data analysis should be indicated in the properties of the allocation and analysis nodes, respectively. Choose from the dropdown menu to indicate how the investigator will be blinded to the group allocation or whether they will be aware of the group allocation during the measurement.

Blinding is especially important when it comes to assessment, particularly if there is a subjective element in assessing the outcome of the treatments, for example when assessing behavioural changes or reading histological slides. The person taking care of the animals and the person assessing the outcome should not know which intervention each of the animals received (i.e. the group allocation) and which animals are grouped together. Randomising the order of examination can help with this. For further information about blinding, click here.

Options in the dropdown menu

Investigator aware of the group allocation (not blinded) – the investigator taking the measurement knows what treatment each animal has received, or what animals are grouped together. Sometimes the person assessing the outcome cannot be blinded to the group allocation, for example if there are obvious phenotypic differences between groups of genetically modified animals; this could be mitigated by, for example, taping the behaviours and sending them to a third party, who has no vested interest in the experiment and does not know whether the transgenic should improve or worsen the outcome. Such an approach would at least counter the directional expectation.

Key features for rigorous research





Key experimental information:

- Objectives and hypotheses
- Animal numbers and justification for sample size

EDA Report

The Experimental Design Assistant (https://eda.nc3rs.org.uk) is an online tool which guides researchers through the design and analysis of in vivo experiments. Information is provided by the researcher to build an EDA diagram – see Annex. Depending on the information inputted specific prompts are triggered by the EDA which provide tailored advice and feedback on the experimental plan.

This report summarises the information provided by the researcher and the feedback from the EDA.

Section 1: Summary

Title of EDA diagram	Effect of drug A on plasma glucose level	
Date report generated	06/12/2019	

Section 2: Information provided by the researcher

1: Objectives

Null hypothesis	Drug A has no effect on plasma glucose levels in diabetic mice
Alternative hypothesis	Drug A modulates plasma glucose levels in diabetic mice
Effect of interest	Change in plasma glucose level
Effect size	100
Justification for effect size	A difference smaller than 100 mg/dL would not be biologically relevant

2: Groups and sample size

Total number of animals in the experiment	12
Groups included in the primary analysis	2 groups:
• Group 1	role=control/comparator, n=6
Group 2	role=test; n=6
Justification for sample size	Power calculation to detect a difference of 100 with 85% power (sig level: 0.06, SD: 50)





EDA

Key experimental information:

- Steps taken to minimise the effect of bias
- Primary and secondary outcome measures
- Planned statistical analysis

3: Randomisation and blinding				
Experimental unit	animal			
There is one step in this experiment where experimental units are allocated to groups:				
Allocation: complete randomisation				
Randomisation strategy	complete randomisation			
Randomisation procedure	EDA spreadsheet			
Allocation concealment	treatments coded for individual animals			
There is one step in this experiment where measu	urements are taken:			
Measurement: Plasma glucose				
Blinding during result assessment	animals individually coded			
There is one analysis in this experiment:				
Analysis: unpaired t-test				
Blinding during analysis of the data	groups coded			
4: Analysis				
Details of the primary analysis (Analysis: unpaired	d t-test)			
Statistical analysis method	unpaired t-test			
Factor of interest	Independent variable of interest: Drug A, categorical, with 2 levels (Drug A, Vehicle)			
Blocking factor	NONE			
Covariate	NONE			
Outcome measures				
Outcome measures in the primary analysis	Outcome measure: glucose concentration, treated as continuous			
Other outcome measures	NONE			
	Page 2 of 4			



EDA

Key experimental information:

• Characteristics of the animals in the study

The Experim in vivo experi	3: Randomisa Experimental u	5: Characteristics of animals in this ex	cperiment	EDA	
specific prom		Species			
This report su	Allocation: com	Strain	mouse DC \(angle (
Section	Randomisation	Sev	Bo.v-Lepab/J		
	Randomisation	Ago			
Title of EU	Allocation con	Age	mean=20, range=19-21, unit=week		
Date repo	There is one step	weight	mean=25, range=20-30, unit=g		
	 Measurement: 	Animal characteristics: diabetic mice, male			
Section	Blinding during	Species	mouse		
1: Object	There is one ana	Strain	B6.V-Lepab/J		
Null hypot		Sex	male		
Alternativ	- Analysis, unpai	Age	mean=20, range=19-21, unit=week		
Effect of i	Blinding during	Weight	mean=25, range=20-30, unit=g		
Justificati 2: Groups	ize 4: Analysis ati Details of the print Details of the print Critique (Table 6) and advice (Table 7) from the EDA is dependent on the quality, including accuracy and completeness, of the information inputted by the researcher. Where the researcher has not addressed issues detected by the EDA, it is important to consider whether this undermines the design of the study.				
Groups in		6: Critique			
• Groups III	Blocking factor	Total number of issues	1		
• Group	Covariate Outcome measur	Issues related to the diagram structure, which might compromise the accuracy of this report	0		
oustinedti	Outcome meas	Issues related to internal consistency	0		
37.	Other outcome	Issues related to missing information	0		
	ouler outcome	Issues suggesting improvements to the	1		
	Other sources of variability are not accounted for in the design of this experiment				
	7: Advice for the primary analysis				
		Suggestion for a method of analysis appropriate for the design	Unpaired t-test or Mann-Whitney test		
NC 3R ^s					
			Page 3 of 4	37	



Information about feedback:

- Summary of the EDA feedback that has not been addressed
- Statistical analysis method recommended by the EDA

Species mouse strain B6.V-Lepab/J Sex female Age mean=20, range=19-21, unit=week Weight mean=25, range=20-30, unit=g Animal characteristics: diabetic mice, male Species Shrain B6.V-Lepab/J Age mean=20, range=19-21, unit=week Weight mean=25, range=20-30, unit=g Animal characteristics: diabetic mice, male Species Strain B6.V-Lepab/J Sex male Age mean=20, range=19-21, unit=week Weight mean=20, range=19-21, unit=week Weight mean=20, range=19-21, unit=week Weight mean=20, range=20-30, unit=g Analysis Section 3: Summary of the feedback provided by the EDA Critique (Table 6) and advice (Table 7) from the EDA is dependent on the quality, including accuracy and completeness, of the inputted by the researcher. Where the researcher has not addressed issues detected by the EDA, it is important to consider v undermines the design of the study. Actor of inter S: Critique totor of inter S: Critique totor of inter S: Critique totor of inter S: Critique i	information whether this
Strain B6.V-Lepab/J Sex female Age mean=20, range=19.21, unit=week Weight mean=25, range=20.30, unit=g Animal characteristics: diabetic mice, male Sec Age mouse Strain B6.V-Lepab/J Sex male Animal characteristics: diabetic mice, male Strain Section 3: Summary of the feedback provided by the EDA Velight mean=20, range=19-21, unit=week Weight mean=20, range=20-30, unit=g Age mean=20, range=20-30, unit=g Age mean=20, range=20-30, unit=g Age mean=20, range=20-30, unit=g Age mean=20, range=20-30, unit=g Analysis Section 3: Summary of the feedback provided by the EDA Critique (Table 6) and advice (Table 7) from the EDA is dependent on the quality, including accuracy and completeness, of the inputed by the researcher. Where the researcher has not addressed issues detected by the EDA, it is important to consider v undermines the design of the study. ctor of interce 6: Critique bring to compromise the accuracy of this report 0 variate Issues related to the diagram structure, whick might compromise the accuracy of this report 0 </th <th>information</th>	information
Sex female Age mean=20, range=19-21, unit=week Weight mean=25, range=20-30, unit=g Animal characteristics: diabetic mice, male Species Strain B6.V-Lepab/J Sex male Age mean=20, range=19-21, unit=week Weight mean=20, range=20-30, unit=g nalysis Section 3: Summary of the feedback provided by the EDA Critique (Table 6) and advice (Table 7) from the EDA is dependent on the quality, including accuracy and completeness, of the inputted by the researcher. Where the researcher has not addressed issues detected by the EDA, it is important to consider v undermines the design of the study. tor of intered cking factor rariate 5: Critique toral number of issues 1 Issues related to the diagram structure, which might compromise the accuracy of this report 0 Issues related to internal consistency 0 Issues related to missing information 0 Issues related to mis	information whether this
Age mean=20, range=19-21, unit=week weight mean=25, range=20-30, unit=g Animal characteristics: diabetic mice, male Animal characteristics: diabetic mice, male Species mouse Strain 86.V-Lepab/J Sex male Age mean=20, range=19-21, unit=week Weight mean=20, range=19-21, unit=week Weight mean=20, range=19-21, unit=week Weight mean=20, range=20-30, unit=g Sex male Age mean=25, range=20-30, unit=g Critique (Table 6) and advice (Table 7) from the EDA is dependent on the quality, including accuracy and completeness, of the inputted by the researcher. Where the researcher has not addressed issues detected by the EDA, it is important to consider v undermines the design of the study. tor of intered 5: Critique Citing factor Total number of issues 1 Issues related to the diagram structure, which might compromise the accuracy of this report 0 Issues related to internal consistency 0 Issues suggesting improvements to the design of this experiment 1	information whether this
weight mean=25, range=20-30, unit=g Animal characteristics: diabetic mice, male Animal characteristics: diabetic mice, male Species mouse Strain B6.V-Lepab/J Sex male Age mean=20, range=19-21, unit=week Weight mean=25, range=20-30, unit=g Sexting factor Section 3: Summary of the feedback provided by the EDA Critique (Table 6) and advice (Table 7) from the EDA is dependent on the quality, including accuracy and completeness, of the inputted by the researcher. Where the researcher has not addressed issues detected by the EDA, it is important to consider violation of interese the design of the study. Cortique Total number of issues 1 Issues related to the diagram structure, which might compromise the accuracy of this report 0 Issues related to internal consistency 0 Issues suggesting improvements to the design of this experiment 1 •Other sources of variability are not accounted for in the design of this experiment 1	information whether this
Animal characteristics: diabetic mice, male Section 3: Age Meight mean=20, range=19-21, unit=week Weight mean=20, range=19-21, unit=week Weight mean=25, range=20-30, unit=g Section 3: Summary of the feedback provided by the EDA Critique (Table 6) and advice (Table 7) from the EDA is dependent on the quality, including accuracy and completeness, of the inputted by the researcher. Where the researcher has not addressed issues detected by the EDA, it is important to consider v undermines the design of the study. For of intere effective exing factor ariate me measure come measure er outcome er outcome er outcome er outcome er outcome	information vhether this
ding during Species mouse strain B6.V-Lepab/J sex male Age mean=20, range=19-21, unit=week Weight mean=25, range=20-30, unit=g Section 3: Summary of the feedback provided by the EDA Critique (Table 6) and advice (Table 7) from the EDA is dependent on the quality, including accuracy and completeness, of the inputted by the researcher. Where the researcher has not addressed issues detected by the EDA, it is important to consider v undermines the design of the study. or of intered rariate Sc Critique issues related to the diagram structure, which might compromise the accuracy of this report 1 issues related to internal consistency 0 issues related to internal consistency 0 issues suggesting improvements to the design of this experiment 1 • Other sources of variability are not accounted for in the design of this experiment 1	information vhether this
Strain B6.V-Lepab/J Sex male Age mean=20, range=19-21, unit=week Weight mean=25, range=20-30, unit=g Section 3: Summary of the feedback provided by the EDA Critique (Table 6) and advice (Table 7) from the EDA is dependent on the quality, including accuracy and completeness, of the inputted by the researcher. Where the researcher has not addressed issues detected by the EDA, it is important to consider v undermines the design of the study. Corr of intered arriate 5: Critique Exking factor Total number of issues 1 Issues related to the diagram structure, which might compromise the accuracy of this report 0 Issues related to internal consistency 0 Issues suggesting improvements to the design of this experiment 1 • Other sources of variability are not accounted for in the design of this experiment 1	information
Sex male alysis: unpai Age mean=20, range=19-21, unit=week Weight mean=25, range=20-30, unit=g softhe print Section 3: Summary of the feedback provided by the EDA Critique (Table 6) and advice (Table 7) from the EDA is dependent on the quality, including accuracy and completeness, of the inputted by the researcher. Where the researcher has not addressed issues detected by the EDA, it is important to consider v undermines the design of the study. or of intered 6: Critique sking factor Total number of issues 1 ariate Issues related to the diagram structure, which might compromise the accuracy of this report 0 successer of the study 0 Issues related to internal consistency 0 Issues suggesting improvements to the 1 . . • Other sources of variability are not accounted for in the design of this experiment 1	information
Age mean=20, range=19-21, unit=week Weight mean=25, range=20-30, unit=g Section 3: Summary of the feedback provided by the EDA Critique (Table 6) and advice (Table 7) from the EDA is dependent on the quality, including accuracy and completeness, of the inputted by the researcher. Where the researcher has not addressed issues detected by the EDA, it is important to consider v undermines the design of the study. or of intere 6: Critique King factor 1 Issues related to the diagram structure, which might compromise the accuracy of this report 0 Issues related to internal consistency 0 Issues related to missing information 0 Issues suggesting improvements to the design of this experiment 1 • Other sources of variability are not accounted for in the design of this experiment 1	informatio vhether thi
Weight mean=25, range=20-30, unit=g Palysis Section 3: Summary of the feedback provided by the EDA s of the print Critique (Table 6) and advice (Table 7) from the EDA is dependent on the quality, including accuracy and completeness, of the inputted by the researcher. Where the researcher has not addressed issues detected by the EDA, it is important to consider wundermines the design of the study. or of interementariate Total number of issues 1 Issues related to the diagram structure, which might compromise the accuracy of this report 0 Issues related to internal consistency 0 Issues suggesting improvements to the design of this experiment 1 Other sources of variability are not accounted for in the design of this experiment 1	informatio vhether thi
halysis Section 3: Summary of the feedback provided by the EDA s of the print Critique (Table 6) and advice (Table 7) from the EDA is dependent on the quality, including accuracy and completeness, of the inputted by the researcher. Where the researcher has not addressed issues detected by the EDA, it is important to consider with undermines the design of the study. or of interdation 6: Critique king factor 1 ariate 1 imputted by the researcher the accuracy of this report 0 Issues related to the diagram structure, which might compromise the accuracy of this report 0 Issues related to internal consistency 0 Issues suggesting improvements to the design of this experiment 1 • Other sources of variability are not accounted for in the design of this experiment 0	informatio vhether thi
Init report Init report Issues related to internal consistency 0 Issues related to missing information 0 Issues suggesting improvements to the design 1 • Other sources of variability are not accounted for in the design of this experiment	
Issues related to internal consistency 0 Issues related to missing information 0 Issues suggesting improvements to the design 1 • Other sources of variability are not accounted for in the design of this experiment 0	
Issues related to missing information 0 Issues suggesting improvements to the design 1 • Other sources of variability are not accounted for in the design of this experiment	
design • Other sources of variability are not accounted for in the design of this experiment	
· Other sources of variability are not accounted for in the design of this experiment	
7: Advice for the primary analysis	
Suggestion for a method of analysis Unpaired t-test or Mann-Whitney test appropriate for the design Image: Comparison of the design	



EDA diagram provides a visual overview to easily see:

- How many groups are being compared
- What variables have been included
- If any variability is being taken into account as a blocking factors or covariates





Read-only EDA diagram

eda.nc3rs.org.uk/eda/modelPublicExport/index/D465A36951173BAE0BE404B82F5A7E41

22/03/2023

100

12

2 groups:

role=test: n=6

level: 0.05, SD: 50)

Example 1: Effect of drug A on plasma glucose level

Drug A has no effect on plasma glucose levels in diabetic mice

A difference smaller than 100 mg/dL would not be biologically relevant

Power calculation to detect a difference of 100 with 85% power (sig

Drug A modulates plasma glucose levels in diabetic mice

Change in plasma glucose level

role=control/comparator; n=6



3: Randomisation and blinding

Justification for sample size

 \rightarrow С \cap

Experimental

Key Experimental Design Features

Section 1: Information provided by the researcher

Desian Assistant

Title of Experiment

1: Objectives

Null hypothesis

Effect of interest

Effect size

Group 1

Group 2

Alternative hypothesis

Justification for effect size

2: Groups and sample size

Total number of animals in the experiment

Groups included in the primary analysis

Date report generated

This figure was generated with the Experimental Design Assistant (EDA, eda.nc3rs.org.uk, RRID: SCR_017019, doi.org/10.1371/journal.pbio.2003779), software to help researchers design in vivo experiments. EDA diagrams are machine-readable 0 the EDA analyses them to provide researchers with tailored advice and feedback to improve the design of their experiments.

ment: Plasma

I recorded as

ndependent

variable of interest: Drug A

cose concen

is output for

The EDA: Benefits for researchers

- Bespoke advice on experimental plans, leading to better experimental design
- Analysis recommendations for the appropriate statistical analysis method

fect of drug

- Help determining the appropriate sample size
- Support for randomisation and masking



FDA



The EDA: Benefits for IACUCs



You could request an EDA report, or read-only diagram, as part of an IACUC application.

- The report and read-only diagram highlight missing information.
- The EDA identifies issues with experimental design.
- You can refer researchers to the EDA application and supporting website for experimental design guidance.





Live: inside the EDA





EDA demonstration

www.nc3rs.org.uk/EDAdemos

Regular live demonstrations

- Building an experiment diagram
- Getting feedback
- Generating a randomisation sequence
- Generating a pdf report
- Generating the read-only diagram
- Finding the built-in help



https://eda.nc3rs.org.uk/overview-demonstration



ARRIVE guidelines

Guidelines to improve the reporting of animal research –

maximising information published and minimising unnecessary studies.



Original guidelines published in 2010. Revised in 2020: ARRIVE 2.0. Supporting resources on the website:

www.ARRIVEguidelines.org



The AR	RIVE	(1) A first of the first of the second se	a) Andreaders (San Tang) Tana Salah Matakang Yang Ang	
guidelir Anirus Researche Departir	1es 2.0 g of the Water Departments	McMark & Typen shape are up of a set to a the set of the set of the set of the set of the set of the provide the the set of the set of the "A set of the set of the set of the set of the "A set of the set of	(1) Synchronia Marca, M. M. (2007), "An Arg. Status, Administration is Series Conservations of the Status, and W. (Research & Arg. & Arg. and Arg. (2007)) (2007) (2007) (2007) (2007), "An Arg. (2007) (2007) (2007) (2007), "Arg. (2007) (2007), "Arg. (2007)) (2007) (2007) (2007), "Arg. (2007), "Arg. (2007), "Arg. (2007)) (2007) (2007) (2007), "Arg. (2007), "Arg. (2007), "Arg. (2007)) (2007) (2007) (2007), "Arg. (2007),	
The dataStructure to a state of the state of	 a dividual in a constraint of each of the constraint of the	(c) In placeton societing or place of the placeton societing or placeton and the relation of the relation o	 Determine patients of 24 of 100 mm sectors called a sector and a sector called a	Plant to botto spacebase Plant to botto spa



Checklist containing key information necessary to describe a study comprehensively and transparently. Can be used as a framework for planning research studies.





In 2015, more than 150 journals signed up to the ARRIVE checklist for animal studies — the highest number of signatories in a single year since it was released.



Endorsed internationally by over 1,000 journals, major research funders, universities, learned societies and scientific organisations.



ARRIVE guidelines



Essential 10 1. Study design 2. Sample size 3. Inclusion and exclusion criteria 4. Randomisation

5. Blinding

6. Outcome measures 7. Statistical methods 8. Experimental animals

9. Experimental procedures

10. Results

Recommended Set

Glossary

2. Sample size

ESSENTIAL 10

2a

Specify the exact number of experimental units allocated to each group, and the total number in each experiment. Also indicate the total number of animals used

Explanation Examples

For any type of experiment, it is crucial to explain how the sample size was determined. For hypothesis-testing experiments, where inferential statistics are used to estimate the size of the effect and to determine the weight of evidence against the null hypothesis, the sample size needs to be justified to ensure experiments are of an optimal size to test the research question [1,2] (see item 13 - Objectives). Sample sizes that are too small (i.e. underpowered studies) produce inconclusive results, whereas sample sizes that are too large (i.e. overpowered studies) raise ethical issues over unnecessary use of animals and may produce trivial findings that are statistically significant but not biologically relevant [3]. Low power has three effects: first, within the experiment, real effects are more likely to be missed; second, where an effect is an over-estimation of the true effect size [4]; and finally, when low power is combined with publication WWW.arriveguidelines.org the false positive rate in the published literature [5]. Consequently, low powered studies contribute to the

2b

research and risk wasting animals used in inconclusive research [6].

Study design can influence the statistical power of an experiment and the power calculation used ne design implemented. Statistical programs to help perform a priori sample size calculations exist for a designs and statistical analyses, both freeware (web based applets and functions in R) and commercial the appropriate calculator or algorithm to use depends on the type of outcome measures and indepe number of groups. Consultation with a statistician is recommended, especially when the experimenta unusual

Where the experiment tests the effect of an intervention on the mean of a continuous outcome meas calculated a priori, based on a mathematical relationship between the predefined, biologically relevant effect size.

PLOS BIOLOGY

Explain how the sample size was decided.

Provide details of any a priori sample size

calculation, if done.

COMMUNITY PAGE



Nathalie Percie du Sert^{1*}, Amrita Ahluwalia^{2,3}, Sabina Alam⁴, Marc T. Avey⁵, Monya Baker⁶, William J. Browne⁷, Alejandra Clark⁸, Innes C. Cuthill⁹, Ulrich Dirnagl¹⁰, Michael Emerson¹¹, Paul Garner¹², Stephen T. Holgate¹³, David W. Howells¹⁴, Viki Hurst¹, Natasha A. Karp¹⁵, Stanley E. Lazic¹⁶, Katie Lidster¹, Catriona J. MacCallum¹⁷, Malcolm Macleod¹⁸, Esther J. Pearl¹, Ole H. Petersen¹⁹, Frances Rawle²⁰, Penny Reynolds²¹, Kieron Rooney²², Emily S. Sena¹⁸, Shai D. Silberberg²³, Thomas Steckler²⁴, Hanno Würbel²⁵

ity actimated from prior data, chocon significance loval, nower and comple size (See "Information used in a r



Six recommendations to increase the methodological rigour and reliability of *in vitro* studies

Developed by an international working group including:

Funders

NC 3R^s

- Journal editors / publishers
- Methodologists & statisticians
- In vitro researchers in industry, academia & government

The RIVER recommendations Reporting *In Vitro* Experiments Responsibly









Experimental design resource links

Resource	link
Experimental Design Assistant	https://eda.nc3rs.org.uk/
Experimental Design Assistant demonstrations	Register for a live demonstration - <u>https://nc3rs.org.uk/EDAdemos</u> Recorded demonstration - <u>https://eda.nc3rs.org.uk/index.php/overview-demonstration</u>
ARRIVE guidelines	https://arriveguidelines.org/
RIVER recommendations	https://doi.org/10.31222/osf.io/x6aut
Webinar: Best practices in experimental design	https://vimeo.com/442640803/09faa99012
MRC/NC3Rs Webinar: Using both sexes in animal research	https://nc3rs.org.uk/3rs-resources/mrc-nc3rs-webinar-using- both-sexes-animal-experiments
NC3Rs experimental design and reporting resources	https://nc3rs.org.uk/3rs-resources/search?topic[]=497
NC3Rs resources for masking	https://www.nc3rs.org.uk/3rs-resources/using- maskingblinding-vivo-experiments
British Pharmacological Society animation – blinding	https://www.youtube.com/watch?v=hbU5kHC9yH0
British Pharmacological Society animation – experimental unit	https://www.youtube.com/watch?v=WQSWJLDcy5M
British Pharmacological Society eLearning resources	https://www.bps.ac.uk/education-engagement/research- animals/blinding-elearning-resource



National Centre for the Replacement Refinement & Reduction of Animals in Research

Thank you!

For more information

Sesther.pearl@nc3rs.org.uk

www.nc3rs.org.uk

🕑 @NC3Rs

🕑 @DrEJPearl

Keep in touch

Our monthly newsletter provides the latest updates from the NC3Rs, including funding calls and events www.nc3rs.org.uk/register



In case you missed it!



OLAW Online Seminar: March 9, 2023 Foundations for Evaluating Study Design and Statistical Approaches for the IACUC

Recording and materials at: <u>https://olaw.nih.gov/education/education</u> <u>al-resources/webinar-2023-03-09.htm</u>



OLAW Online Seminar: June 15, 2023 The ARRIVE 2.0 Essential 10: Guidance for NIH-sponsored Research

Recording and materials at: <u>https://olaw.nih.gov/education/educational-</u> <u>resources/webinar-2023-06-15.htm</u> 51

Next Webinar: Winter 2023 Topic TBD

