THE UNIVERSITY OF MELBOURNE

Faculty of Medicine, Dentistry and Health Sciences

MACH Melbourne Academic Centre for Health

Statistics for your grant applications

Professor Julie Simpson

Methods and Implementation Support for Clinical and Health (MISC) research Hub

Melbourne School of Population and Global Health

Website:- https://clinicalresearch.mdhs.unimelb.edu.au/

Email:- misch-info@unimelb.edu.au

@MischHub

Housekeeping

- Please keep your microphone switched off during the presentation.
- You are welcome to leave your video on or off as you prefer.
- If you have any questions, please feel free to enter them in the chat box. We will review them throughout the presentation.
- Note that this presentation will be recorded and a link will be provided after the webinar.
- A copy of the slides will also be provided.

MISCH Hub

Collaborate with MISCH to maximise your research impact

Methods and **I**mplementation **S**upport for **C**linical and **H**ealth research Hub

Our aim is to provide support to researchers and affiliated researchers of the University of Melbourne in clinical and health research.

We provide support on core research methods of Biostatistics and Clinical Epidemiology, Health Economics, Clinical Trials, Implementation Effectiveness and Co-Design and Health Informatics (REDCap).

MISCH Staff

CLINICAL TRIALS LEAD
Adam Deane

MISCH MANAGER
Shaie O'Brien

'HEALTH ECONOMIST LEAD
An Duy Tran

BIOSTATISTICIAN
Aistair McLean

HEALTH INFORMATICIAN
David Ormiston-Smith

BIOSTATISTICIAN
Hannah Johns

DEPUTY HEAD, MISCH BIOSTATISTICS
Sabine Braat

BIOSTATISTICIAN
Dominic Italiano

MISCH OPERATIONS MANAGER
An Nguyen

BIOSTATISTICIAN
Anurika De Silva

HEALTH INFORMATICS LEAD
Douglas Boyle

BIOSTATISTICIAN
Patty Chondros

BIOSTATISTICIAN
Emily Karahalios

DIRECTOR OF MISCH
Julie Simpson

HEALTH ECONOMIST
Kim Dalziel

BIOSTATISTICS LEAD, AUSTIN HOSPITAL
Leonid Churilov

BIOSTATISTICIAN
Rob Mahar

BIOSTATISTICIAN
Karen Lamb

BIOSTATISTICIAN
Sophie Zaloumis

CO-DESIGN AND IMPLEMENTATION EFFECTIVENESS
Jennifer Bibb

EPIDEMIOLOGIST AND BIOSTATISTICIAN
Digsu Koye

EPIDEMIOLOGIST
Mark Jenkins

BIOSTATISTICIAN
Vanessa Pac Soo

CO-DESIGN AND IMPLEMENTATION EFFECTIVENESS LEAD
Victoria Palmer

HEALTH ECONOMIST
Paul Amores

CLINICAL TRIALS OFFICER
Katie Ozdowska

Why you need statistics for grant applications

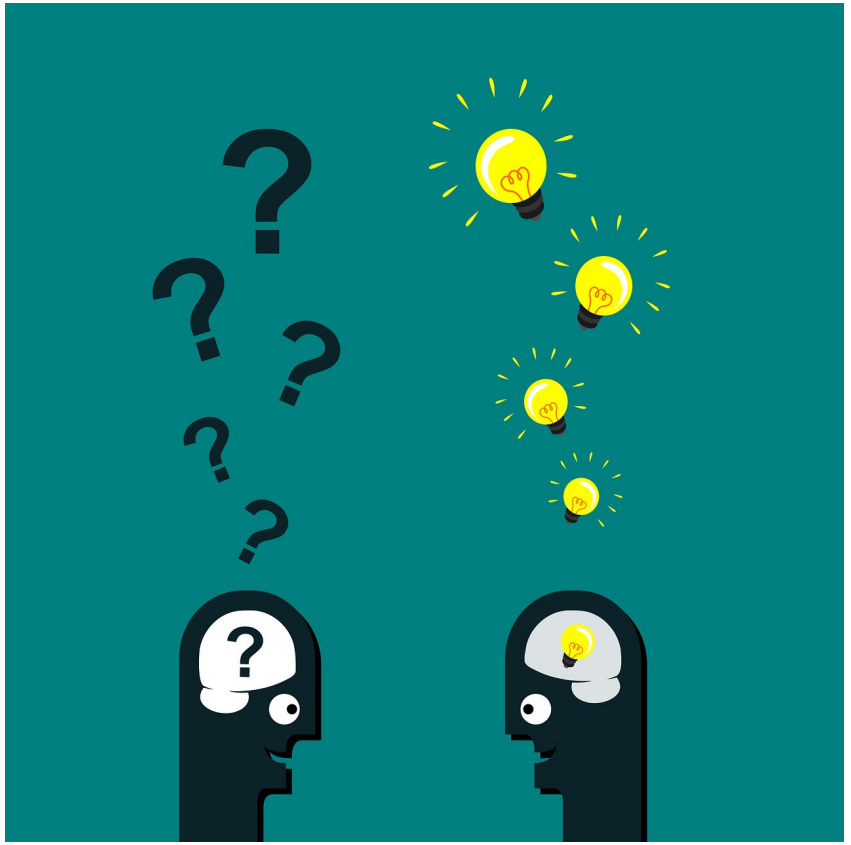
- Getting the statistical sections of your grant correct can be critical to the success of your application.
- Often only the **sample size calculation** and **statistical analysis sections** are considered when researchers think about the statistics for a grant application.
- Statistics is not solely about deciding what analysis method (e.g. regression, survival analysis) to use.
- A well crafted grant has statistical elements throughout.

Why you need statistics for grant applications

- Statistics (and statisticians) are important for the **design**, **conduct**, **analysis** and **interpretation** of study findings.

- Statisticians supporting grant applications check that the **grant flows** clearly from the research question and hypotheses to the study design to the methods.

- Our MISCH team provides support to researchers to ensure that the design is appropriate to answer the proposed research question.



PICOTs

- Consider **PICOT** to help frame your **research question**

[See recording of seminar by Karen Lamb: *Kicking off your research question: how to craft a well-defined research question* https://machaustralia.org/]

- This helps identify important statistical information to include in your grant

Population	Who should be in the study?
Intervention	Intervention/Exposure
Comparator	Control
Outcome	What is the outcome of interest?
Time	Over what time period?
(study design)	Study design features (e.g. parallel-group RCT)

Study design (PICOTs)

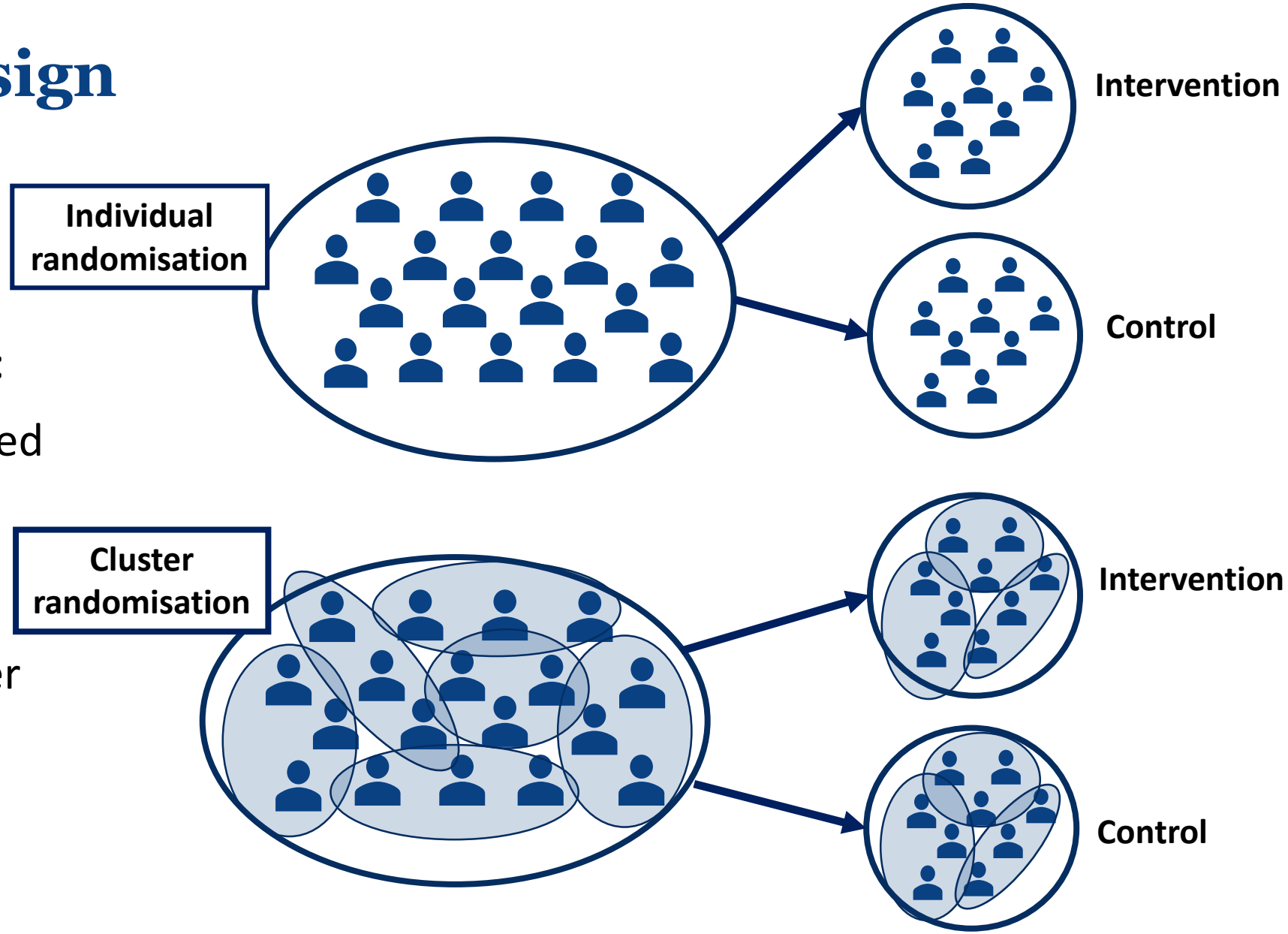
- It is essential to provide details of the methods proposed to tackle the research question.
- Reviewers need to see not only that the question under investigation is important but that it is **methodologically sound** and **feasible**.
- The topics that should be covered depend on the nature of the study.
- It is necessary to clearly state the overall **study design** early in the grant proposal (e.g., cohort study, cluster randomised controlled trial).

St George's University of London 'Statistics Guide for Research Grant Applications' https://www-users.york.ac.uk/~mb55/guide/describ.htm#intro

Study design

Some RCT study designs:

- Individually randomised
- Cross-over
- Cluster randomised
- Stepped wedge cluster
- …



Sample size section

Sample size section

- Sample size calculations are required for **most quantitative studies**.
- Failing to include a sample size calculation for a quantitative study is a common reason for rejecting a grant application so be sure to prioritise this aspect of your grant.
- If you have multiple primary outcomes, the sample size has to be sufficient to detect the effect of interest for **all** outcomes (i.e., you have to choose the larger of the sample sizes from the calculations to ensure sufficient power for all outcomes).
- Sample size considerations for Pilot and Feasibility studies - **[See recording of seminar by Sabine Braat: *What's it going to take to get your study started? Pilot and Feasibility studies.* https://machaustralia.org/]**

St George's University of London 'Statistics Guide for Research Grant Applications' https://www-users.york.ac.uk/~mb55/guide/size.htm

Sample size section – Hard Quiz Question 1

G'day Steve.

Sample size section – Hard Quiz Question 1

Which of the following are excellent examples of the sample size text in a grant application?

a) A previous trial in this same area recruited 150 patients and found significant results (p=0.014), and therefore a sample size of 150 patients has been selected for the proposed study.

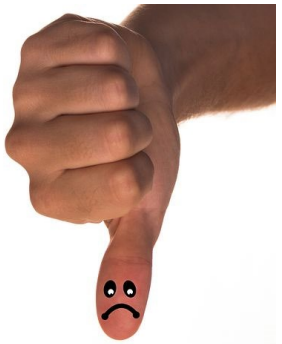
b) Sample sizes are not provided for this Covid treatment trial because there is no information on which to base them.

c) The clinic attends to around 50 patients per year, of whom 10% may refuse to take part in the study. Therefore over the 2 years recruitment phase of the study, the sample size will be 90 patients.

d) All of the above

e) None of the above

Sample size section – what NOT to do



"A previous study in this area recruited 150 subjects and found significant results (p=0.014), and therefore a similar sample size should be sufficient here."

- This does NOT mean the prior study was sufficiently powered!
- This could be a chance finding.
- Sample size calculations must be calculated for your specific study and the effect you wish to detect.

St George's University of London 'Statistics Guide for Research Grant Applications' https://www-users.york.ac.uk/~mb55/guide/size.htm

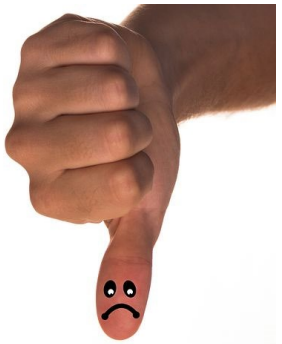
Sample size section – what NOT to do

- This is something statisticians commonly hear when working with researchers.
- It is important to make a concerted effort to find prior relevant published information.
- Alternatively, a small study could be conducted to obtain the required information.
- General sample sizes can still be undertaken if some information (e.g., standard deviation of the outcome) is not available.

"Sample sizes are not provided because there is no information on which to base them."

St George's University of London 'Statistics Guide for Research Grant Applications' https://www-users.york.ac.uk/~mb55/guide/size.htm

Sample size section – what NOT to do



"The throughput of the clinic is around 50 patients a year, of whom 10% may refuse to take part in the study. Therefore over the 2 years of the study, the sample size will be 90 patients."

- This is valuable information to include in your grant and helps highlight the **feasibility** of the study being undertaken.
- It does **NOT** mean your study will be **powered**. Underpowered studies are a waste of resources and are unlikely to get funded.
- A sample size calculation is required to determine the power to detect differences of interest.
- If the sample size is too small, you may want to extend the study length or collaborate with other centres.

St George's University of London 'Statistics Guide for Research Grant Applications' https://www-users.york.ac.uk/~mb55/guide/size.htm

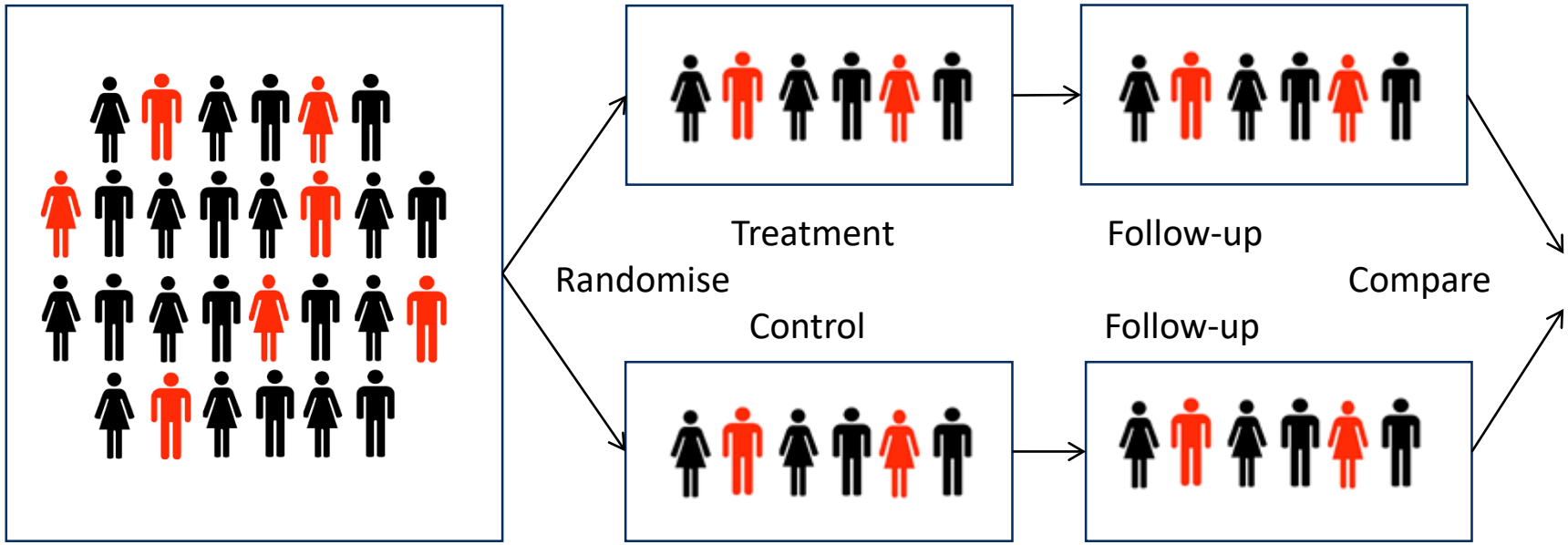
How many participants do I need in my study?

- In order to think about sample size for a study, there must be a clearly articulated **research question.**

- Sample size estimates are based on the **primary outcome** that the study is investigating, so
  - ***Outcome*** measure must be clearly articulated
  - How will the ***outcome*** be measured?
  - Is the ***outcome*** categorical or numerical?
  - How will the o***utcome*** be analysed?

- Sample size depends on the researcher's knowledge & assumptions such as those arising from systematic reviews (as much as technical statistical calculations).
  - It is important to carry out sample size calculations for several different scenarios, not just one.

Sample size: Demonstration for a RCT



Aim: – To compare some outcome measure between treatment and control group

- randomised at individual level

- superiority trial

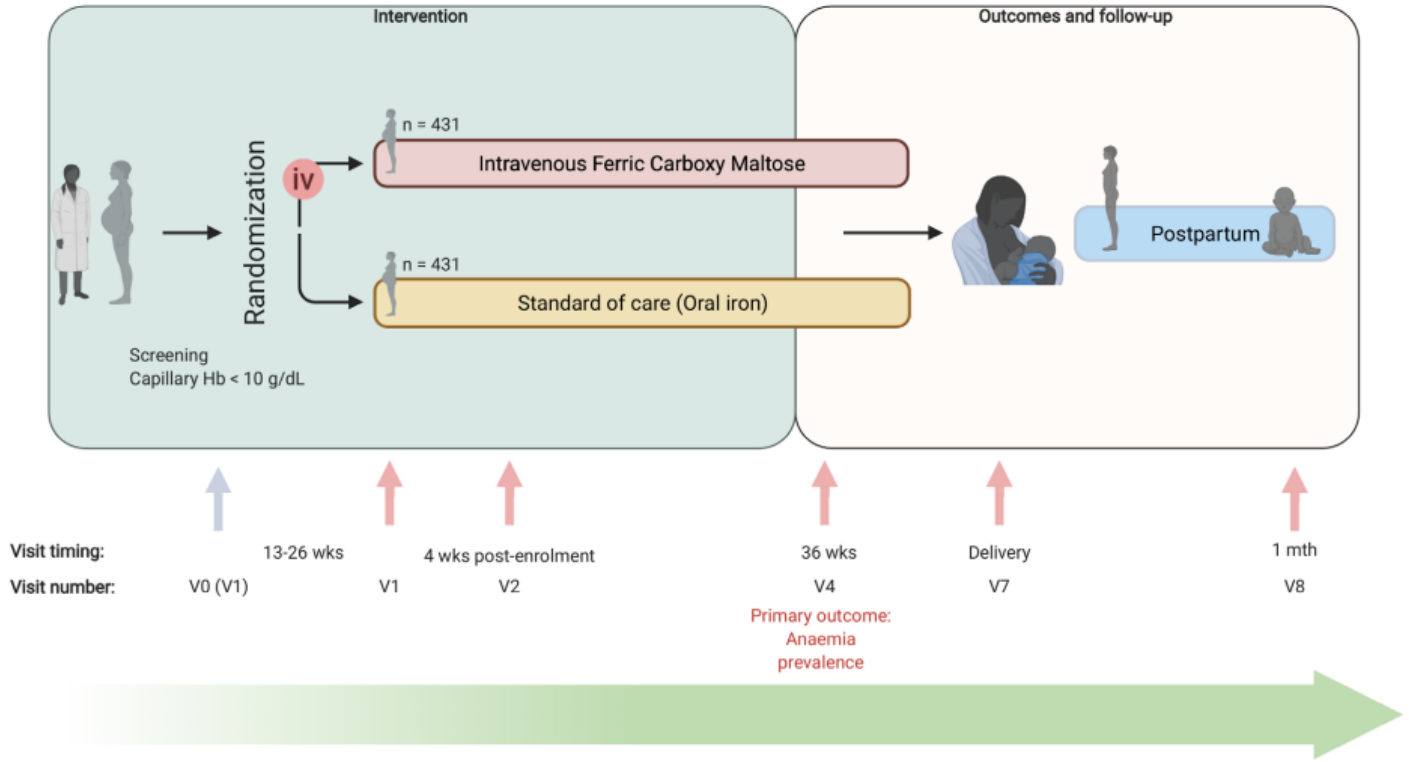
Sample size: Demonstration for a RCT
- randomised at individual level
- superiority trial

REVAMP trial – *Mwangi M et al. BMJ Open, 2021*

Population	Malawian pregnant women (recruited at 13-26 weeks' gestation)
Intervention	Intra-venous iron (once over 15 minutes after randomisation)
Comparator	Standard of care – oral iron treatment course (two times per day for 90 days)
Outcome	Prevalence of maternal anaemia (venous haemoglobin concentration < 110 g/L) at 36 weeks gestation
Time	Primary outcome assessed at 36 weeks gestation (note:- follow-up until 1 month post-partum for secondary outcomes)
(study design)	Randomised Controlled Trial (*multi-centre, open-label, superiority, two-arm, parallel group, individually randomised*)

Sample size: Demonstration for a RCT
- randomised at individual level
- superiority trial

REVAMP trial – *Mwangi M et al. BMJ Open, 2021*



Sample size: Demonstration for a RCT

- randomised at individual level
- superiority trial

Information (**ingredients**) required for sample size calculation:-

1) Baseline information
   The **proportion** with the feature in the control group
   (**categorical outcome**)
   OR
   Measure of variability in the control group
   (**numerical outcome**)

2) Minimum clinically important difference
   The smallest difference in outcome between the treatment and control groups that would be deemed to be of 'clinical/public health' relevance.

3) Relative sizes of the two groups
   Ratio of number of treatment / control

4) Significance level
5) Power

Sample size – *information needed for REVAMP trial*

1) Baseline information

We need…..

Outcome - Prevalence of maternal anaemia (venous haemoglobin concentration < 110 g/L) at 36 weeks gestation ***for the***

Control group - Standard of care – oral iron treatment course (two times per day for 90 days)

Set at 60% - based on study in Gambia

Source picture: https://www.mdedge.com/obgyn/article/153051/obstetrics/recognize-and-treat-iron-deficiency-anemia-pregnant-women

Sample size – *information needed for REVAMP trial*

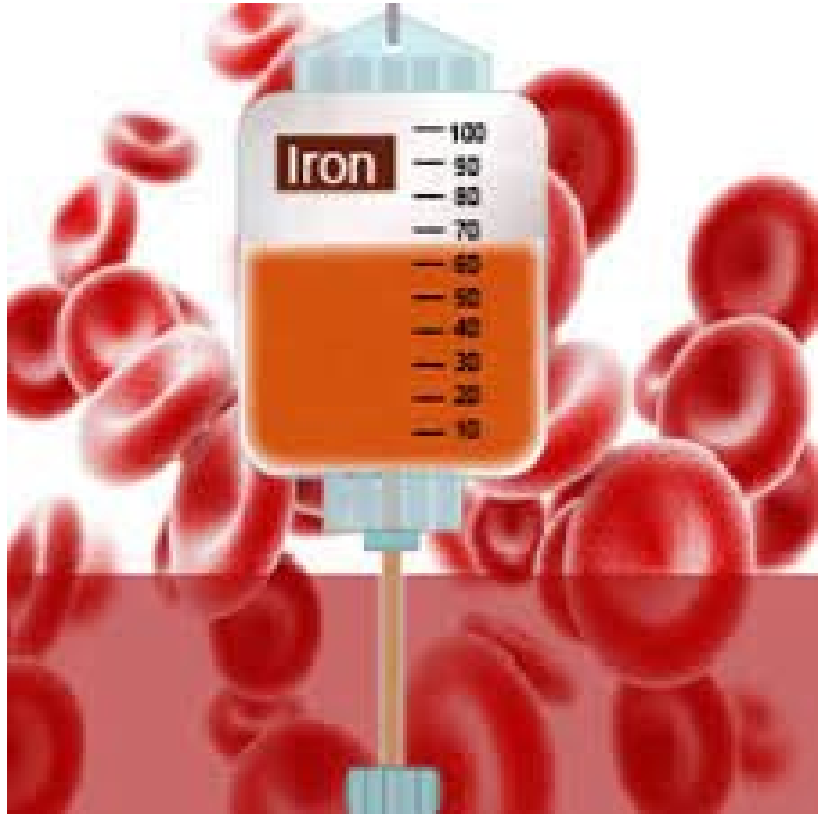
2) Minimum clinically important difference

Hypothesise that intra-venous iron will result in an absolute decrease of 10% units in prevalence of anaemia at 36 weeks gestation.

[10% - justified based on a similar trial in a high income setting that observed 14%]

IV group – prevalence of 50%

Standard of care oral group – prevalence of 60%

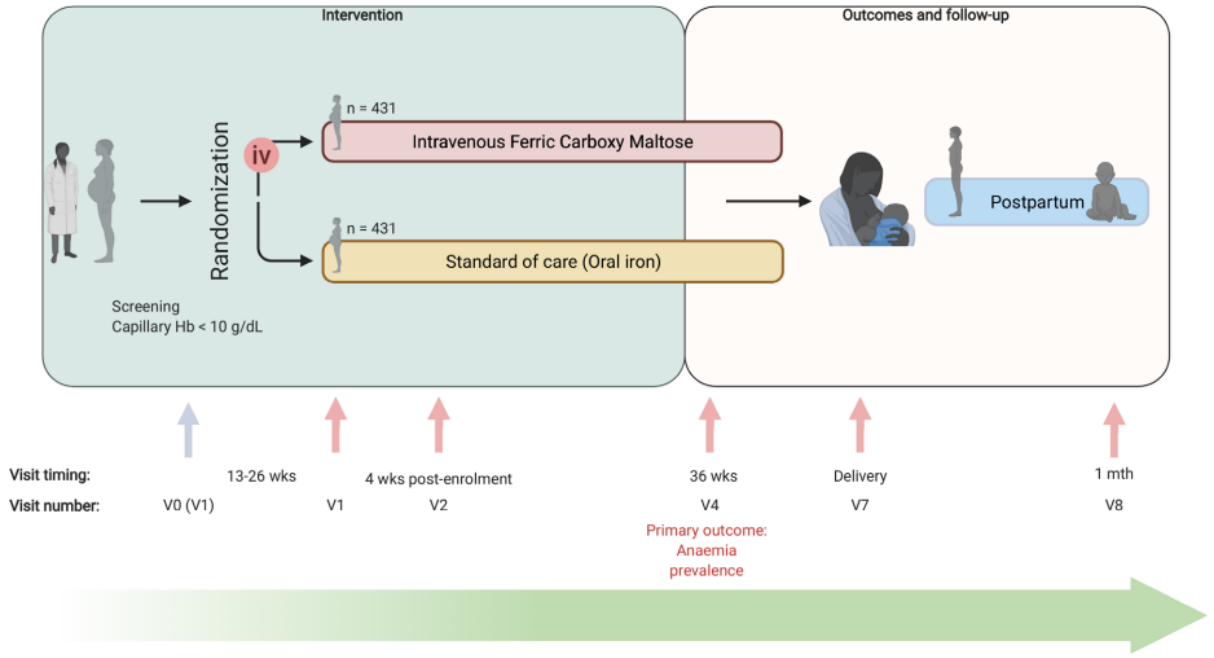


Source picture: https://podtail.com/podcast/dr-chapa-s-obgyn-pearls/iv-vs-oral-iron-therapy-in-pregnancy/

Sample size – *information needed for REVAMP trial*

3) Relative sizes of the two groups

Ratio of number of pregnant women randomised to IV versus oral iron – **1:1**



Sample size – *information needed for REVAMP trial*

4) Significance level & 5) Power

Null hypothesis in superiority question

A null hypothesis is one that proposes there is no difference in population parameter between groups

REVAMP null hypothesis:

- The prevalence of anaemia at 36 weeks gestation is the same for those in the population who receive IV iron or oral iron.

Hypothesis testing

Type I error

Investigator concludes from sample:
"IV iron reduces the prevalence of anaemia at 36 weeks gestation compared to standard oral iron"
(i.e. reject the null hypothesis)

WHEN

There is NO difference in prevalence of anaemia at 36 weeks gestation between IV and oral iron in the population
(i.e. null hypothesis is true)

Type II error

Investigator concludes from sample:
"There is no difference between the prevalence of anaemia at 36 weeks gestation for IV and oral iron groups"
(i.e. do not reject the null hypothesis)

WHEN

There is a REAL difference in prevalence of anaemia at 36 weeks gestation following IV iron compared to oral iron in the population
(i.e. null hypothesis is not true)

Hypothesis testing

Type I error (α)

Convention to fix at 5%

Two-sided significance level = α = 5%†

We will incorrectly interpret a difference as a real difference on less than 5% of occasions

(false positive)

Type II error (β)

Convention to fix at 10 or 20%

Power = 1-β
= 90% or 80%

We will be able to detect an important difference on 80/90% of occasions and will miss it on 20/10% of occasions

(false negative)

† - leads to $p<0.05$ convention for "statistical significance", note that $p=0.049$ & $p=0.051$ are not in reality different, but for planning purposes it is necessary to have a cut-off.

Sample size – *information needed for REVAMP trial*

1) Baseline information

   The proportion of anaemia at 36 weeks gestation for standard oral iron group - 60%

2) Minimum clinically important difference

   The smallest difference in outcome between IV and oral iron groups that would be deemed to be clinically relevant

   – 10% (i.e. prevalence of anaemia in IV iron group = 50%)

3) Relative size of IV to oral iron group: 1 to 1

4) Significance level (two-tailed) 5%

5) Power 80%

Sample size – *REVAMP trial*

REVAMP trial:- Comparing prevalence of anaemia at 36 weeks gestation between IV and oral iron groups

	5% significance level		1% significance level	
	Minimum clinically important difference		Minimum clinically important difference	
Power	50% vs 60% **10%** (absolute reduction)	55% vs 60% **5%**	50% vs 60% **10%**	55% vs 60% **5%**
80%	388 (no. of patients per group)			
90%				

Sample size section – Hard Quiz Question 2

If we reduce the minimum clinically important difference from 10% to 5% for the absolute reduction in prevalence of anaemia at 36 weeks gestation, will the sample size required?

a) Increase

b) Decrease

c) Don’t know

Sample size – *REVAMP trial*

REVAMP trial:- Comparing prevalence of anaemia at 36 weeks gestation between IV and oral iron groups

Power	5% significance level		1% significance level	
	Minimum clinically important difference		Minimum clinically important difference	
	50% vs 60% **10%** (absolute reduction)	55% vs 60% **5%**	50% vs 60% **10%**	55% vs 60% **5%**
80%	388 (no. of patients per group)	1534		
90%				

Sample size section – Hard Quiz Question 3

If we increase the power from 80% to 90%, will the sample size required?

a) Increase

b) Decrease

c) Don’t know

Sample size – *REVAMP trial*

REVAMP trial:- Comparing prevalence of anaemia at 36 weeks gestation between IV and oral iron groups

	5% significance level		1% significance level	
	Minimum clinically important difference		Minimum clinically important difference	
Power	50% vs 60% **10%** (absolute reduction)	55% vs 60% **5%**	50% vs 60% **10%**	55% vs 60% **5%**
80%	388 (no. of patients per group)			
90%	519			

Sample size section – Hard Quiz Question 4

If we decrease the significance level from 5% to 1%, will the sample size required?

a) Increase

b) Decrease

c) Don’t know

Sample size – *REVAMP trial*

REVAMP trial:- Comparing prevalence of anaemia at 36 weeks gestation between IV and oral iron groups

	5% significance level		1% significance level	
	Minimum clinically important difference		Minimum clinically important difference	
Power	50% vs 60% **10%** (absolute reduction)	55% vs 60% **5%**	50% vs 60% **10%**	55% vs 60% **5%**
80%	388 (no. of patients per group)		577	
90%				

Sample size – *REVAMP trial*

REVAMP trial:- Comparing prevalence of anaemia at 36 weeks gestation between IV and oral iron groups

Power	5% significance level		1% significance level	
	Minimum clinically important difference		Minimum clinically important difference	
	50% vs 60% **10%** (absolute reduction)	55% vs 60% **5%**	50% vs 60% **10%**	55% vs 60% **5%**
80%	388 (no. of patients per group)	1534	577	2282
90%	519	2053	735	2907

REVAMP sample size – grant application

Grant application statement:-

> *"This study, with 388 pregnant women per group, has 80% power, i.e. an 80% chance of producing a statistically significant finding at a two-sided 5% significance level, to detect an absolute difference in prevalence of anaemia at 36 weeks gestation of 10% between IV and oral iron groups, assuming a prevalence of 60% in the oral iron group (i.e. 50% versus 60% respectively)."*

Note: *You need to provide further text in the grant application to justify the values of 60% for the control (standard of care) group, and the absolute reduction of 10%.*

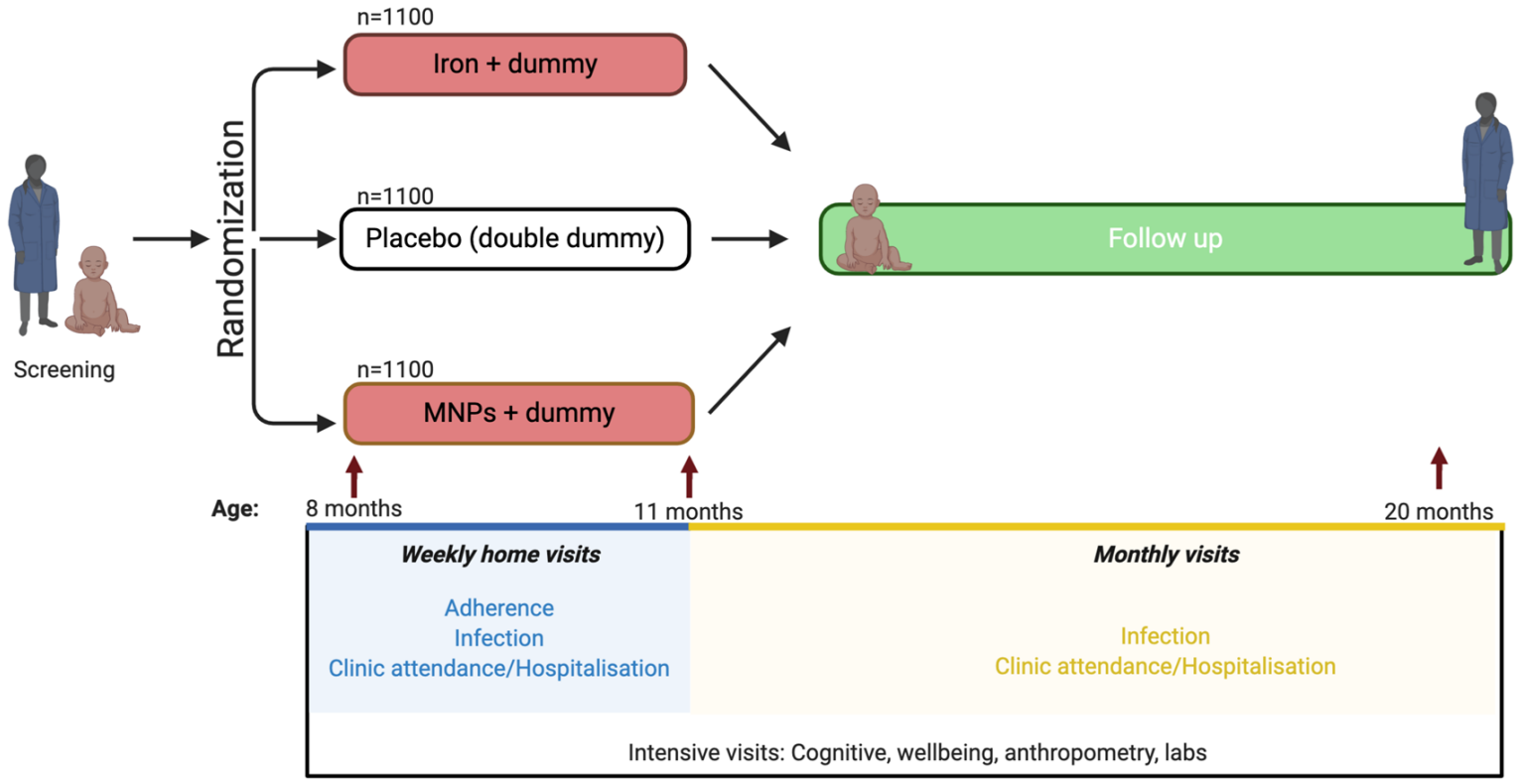
Sample size: Demonstration for a RCT
- randomised at individual level
- superiority trial

BRISC trial – *Pasricha SR et al. NEJM 2021*

Population	Bangladeshi 8-month old infants
Intervention	Intervention 1: Iron drops (3-months) Intervention 2: Multiple micronutrient powders (3-months)
Comparator	Placebo (3-months)
Outcome	Cognitive composite score on the Bayley Scales of Infant and Toddler Development
Time	Primary outcome assessed after 3-months of intervention (note:- follow-up until 12 months post-randomisation)
(study design)	Randomised Controlled Trial (*multi-centre, double-dummy, superiority, three-arm, parallel group, individually randomised*)

Sample size: Demonstration for a RCT
- randomised at individual level
- superiority trial

BRISC trial – *Pasricha SR et al. NEJM 2021*



BRISC sample size

Pasricha SR et al. NEJM 2021

Primary outcome – Cognitive composite score on the Bayley Scales of Infant and Toddler Development, assessed at completion of 3 month regimen.

1) Baseline information

   Measure of variability in the placebo group

   (numerical outcome) – standard deviation = 15 points

2) Minimum clinically important difference

   The smallest difference in mean outcome between the treatment and placebo group that would be deemed to be clinically relevant

   – 2-point difference

Factors important in sample size calculations

Basic ingredients for power-based sample size calculation in RCT:

- The proportion with the feature in the control group (categorical outcome) OR measure of variability in the control group (numerical outcome)
  - published data, pilot data, guess-timate of range
- Minimum clinically important difference
  - must be based on clinical (substantive) considerations
- Relative sizes of the two groups (usually 1:1 in most trials)
- Significance level, two-sided (near-universal convention to set this at 0.05)
- Power
  - Conventionally never less than 0.8 (80%); for more important studies many authorities recommend 90%

Factors important in sample size calculations

In addition to be incorporated:

- Missing data: - (illustrated no loss-to-follow-up)
  - Predicted response &/or loss to follow-up rates, e.g. REVAMP - loss to follow-up expected to 10%, increased sample size by 100/90.
  - Of note, beware of non-random dropout leading to *bias* (larger samples do not correct for bias)

- Study design:- (illustrated parallel group, individual level design)
  e.g. cross-over trial, cluster randomised trial

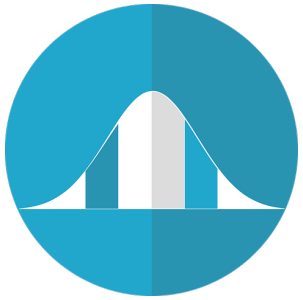
- Study question:- (illustrated difference question - superiority)
  to show difference, equivalence or non-inferiority

Finally:

- It is important that you calculate the sample size using the same/similar techniques to the primary analysis presented in your grant application.

Statistical methods

Statistical methods



- The statistical methods should include **sufficient details** about how each of the research questions posed will be addressed.

- The reviewer should be made aware that the researchers have spent time thinking about how the data collected will be used.

- A grant is unlikely to be funded with only statements like:

A statistician will be employed to conduct the statistical analysis.

OR

Statistical analysis will be conducted using Stata.

St George's University of London 'Statistics Guide for Research Grant Applications' https://www-users.york.ac.uk/~mb55/guide/methods.htm

Statistical methods

- The statistical methods section should include **unambiguous statements** about how analyses will be conducted.

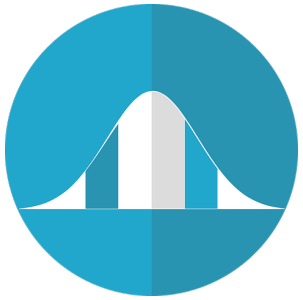
- For example:

Linear regression will be used to compare the mean difference in body mass index (primary outcome) between the two treatment arms at 6 month follow-up, controlling for baseline body mass index.

- The method described should be **appropriate for the study design** (e.g., clustered designs need to take into account clustering in the analysis) and the **data type of outcome** (e.g., continuous, categorical).

St George's University of London 'Statistics Guide for Research Grant Applications' https://www-users.york.ac.uk/~mb55/guide/methods.htm

Statistical methods



- Importantly, the statistical methods used for the primary outcome(s) should directly align with the sample size calculation included in the grant.

- If the primary outcome is continuous (e.g., body mass index) but the statistical analysis section includes only methods for dealing with a binary outcome (e.g., overweight or not overweight), this raises a red flag to reviewers.

- The grant should be **consistent** throughout.

Statistical methods

REVAMP Trial – Statistical method for primary outcome

"Anaemia will be analysed using a log-binomial regression model. The model will include the standard-of-care (oral iron) group as the reference group. The primary maternal hypothesis will be evaluated by obtaining the estimate of the prevalence ratio of intravenous iron versus standard-of-care (oral iron), 95% CI at 36 weeks' gestation, and p value."

Statistical methods – Hard Quiz Question 5

Is the following paragraph sufficient for the statistical methods section in a grant application?

"Treatment groups will be compared using a t-test (continuous outcomes) or chi-squared test (binary/categorical outcomes). We will declare our intervention successful if the p-value is less than 0.05."

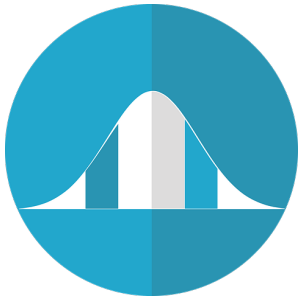
a) Yes

b) No

c) Maybe

d) Don't know

Statistical methods



- Topics typically covered in the statistical methods sections are:
  - Analysis sample
  - Model for primary and key secondary outcomes
  - Handling missing data
  - Handling multiple testing
  - Subgroup analyses
- The grant should allow for sufficient space to cover the above topics in the statistical methods section...

Investigator team

- Are your CI/AI team the winning team to deliver this project, with all the required expertise?
- Important to include a biostatistician.

Check out:-

Grant Success: Choosing a Winning Team for your Grant Application

Wednesday June 1, 2pm-5pm

https://clinicalresearch.mdhs.unimelb.edu.au/

https://www.eventbrite.com.au/e/grant-success-choosing-a-winning-team-for-your-grant-application-tickets-305186931327

2:00pm	Introductions from **Dr Bow Tauro** (MRFF Senior Project Officer) and **Prof Julie Simpson** (MISCH Director)
2:20pm–2:35pm	**Prof Kim Bennell:** *The key role of stakeholders – Examples from a successful partnership with Medibank*
2:35pm–2:50pm	**Dr An Duy Tran:** *Integration of health economic research into a grant application: the case of Targeted Translation Research Accelerator Program*
2:50pm–3:05pm	**Prof Victoria Palmer:** *At the crossroads, translation research needs transdisciplinary teams – Lessons from interdisciplinary engagements*
3:05pm–3:20pm	**Dr Adam Deane:** *Can a soufflé rise twice? Repurposing your application for different tastes*
3:20pm–3:35pm	**Prof Jane Hocking:** *Maximising your score – experiences from the NHMRC CT/CS scheme*
3:35pm–3:50pm	**Prof Leonid Churilov:** *Never leave home without a statistician (especially when applying for research funding…)*
3:50pm	Questions
4:00pm	Break
4:15pm–5:00pm	Break out rooms with facilitated discussion. Moderators are: Prof Kim Bennell, Prof Leonid Churilov, Dr Adam Deane, Dr An Duy Tran, Prof Victoria Palmer, Prof Jane Hocking

Funding

- It is important to consider the funding required to support the data and statistical aspects of your study.
- Do you need support with **data management** (e.g., database set up and design)?
- Will you need **ongoing statistical support** for the design, conduct, analysis and reporting of your study?
- Statisticians and data managers require funding to ensure they can continue to support research.
- It is important that these costs are built into the budget.

Ensure sufficient time for statistics!

- It is important to think about the statistics as early as possible in the planning stage.
- Statisticians can assist with framing the research question(s) and identifying the best design.
- This can take time!

THE UNIVERSITY OF MELBOURNE

Faculty of Medicine, Dentistry and Health Sciences

MACH Melbourne Academic Centre for Health

Next MACH webinar

12:30-1:30pm, 22nd June

Economic evaluation alongside clinical trials: principles of study design and decision analysis

Dr An Duy Tran, Head of MISCH Health Economics Node

Thank you

THE UNIVERSITY OF MELBOURNE

Faculty of Medicine, Dentistry and Health Sciences

MACH Melbourne Academic Centre for Health

- Recording:- https://machaustralia.org/
- MISCH Newsletter:- https://clinicalresearch.mdhs.unimelb.edu.au/collaborate/contact-us/misch-newsletter-sign-up
- Website:- https://clinicalresearch.mdhs.unimelb.edu.au/
- Email:- misch-info@unimelb.edu.au
- @MischHub