

WOMBAT Safer Birth at Term Outcomes and Sample Size

Introduction – John Simes



THE UNIVERSITY OF
SYDNEY

Celebrating 175 *years*



N H M R C
Clinical Trials Centre

Issues to consider in planning / designing platform trials for safer birth at term

- What is the aim of each trial question
 - To identify optimal treatment / care for future practice (phase III practice changing)
 - To assess some comparative measures of efficacy and safety (earlier phase study)
- What are the proposed interventions and how will they impact outcomes
- Is the primary outcome a measure efficacy, safety, effectiveness, net clinical benefit
- Potential outcome measures (single outcome, composite outcome, ranked / ordinal outcome, etc)

For the proposed outcome (ordinal, composite, etc)

- How sensitive is it? To improve study power.
- Does it capture overall benefit
- Does it capture net clinical benefit?
- Will it be sufficient to guide future clinical practice?
- Will it be helpful more in shaping future research?
- Are there problems with these outcomes with different effects on different outcomes?
- How important is the biological mechanism / treatment pathway in choice of outcomes?
- How important are the values to participants of the outcomes?

Novel endpoints

- Novel endpoints are becoming increasingly common in platform trials
- Binary endpoints (e.g. death/survival) may not be sensitive enough
- Traditional composite endpoints do not take account of event hierarchies
- Ordinal endpoints are increasingly in favour, as they take account of event hierarchies that distinguish between events of varying severity
- Ordinal endpoints are more sensitive than binary endpoints (greater power)
- Statistical analysis: greater complexity and more assumptions

WOMBAT Safer Birth at Term Outcomes and Sample Size

Presentation – Ian Marschner



THE UNIVERSITY OF
SYDNEY

Celebrating 175 *years*



N H M R C
Clinical Trials Centre

Ordinal endpoint features

- Categorical endpoint: K categories
- Categories are ordered for severity 1, 2, 3, ... , K , but are not treated as quantitative
- In traditional statistical analysis K is fairly small e.g. $K=5$ or 6
- Modern trials are choosing K to be very large e.g. 20 or more
- Large K introduces additional complexities and statistical assumptions are more likely to be violated

A traditional ordinal outcome: Modified Rankin Scale in stroke

The Modified Rankin Scale

Grade	mRS
0	No symptoms at all
1	No significant disability: despite symptoms, able to carry out all usual duties and activities
2	Slight disability: unable to perform all previous activities but able to look after own affairs without assistance
3	Moderate disability: requiring some help but able to walk without assistance
4	Moderately severe disability: unable to walk without assistance and attend to own bodily needs without assistance
5	Severe disability: bedridden, incontinence, and requiring constant nursing care and attention
6	Death

A novel ordinal outcome: PLATIPUS versus APTS

Infant outcome	APTS binary	PLATIPUS ordinal
Liveborn, no conditions below	No	1
Admitted to neonatal unit, <48h	No	2
Admitted to neonatal unit, 48h+	No	3
Non-invasive respiratory support, <5 days	No	4
Non-invasive respiratory support, 5 days or more	No	5
Mechanical ventilation, <7 days	No	6
Mechanical ventilation, 7 days+	No	7
Moderate respiratory morbidity	No	8
Necrotising enterocolitis and/or Sepsis	Yes	9
Severe respiratory morbidity	Yes	10
Major surgery	Yes	11
Brain injury	Yes	12
Two of: Brain injury, Surgery or Severe respiratory morbidity	Yes	13
Brain injury <u>and</u> Surgery <u>and</u> Severe respiratory morbidity	Yes	14
Death	Yes	15

Ordinal and binary endpoints

- How do we compare ordinal endpoints?
- **Key principle:**
 - An ordinal endpoint with k levels is a collection of $k-1$ binary endpoints
 - Each binary endpoint can be compared using a standard treatment effect measure such as an odds ratio
- **Key assumption:**
 - The treatment effect on each binary endpoint is identical

Example of key assumption:
HERO-2 trial of anti-coagulation in heart attack

Key safety outcome:

5-level ordinal bleeding endpoint:

1. No bleeding
2. Minor bleeding
3. Moderate bleeding
4. Major bleeding
5. Intracranial
haemorrhage

Bleed severity	Bivalirudin	Heparin	Total
1	7218	7633	14 851
2	1076	763	1839
3	117	89	206
4	58	40	98
5	47	32	79
Total	8516	8557	17 073

HERO-2: 5-level ordinal endpoint = 4 binary endpoints

Bleed severity	Bivalirudin	Heparin	Total
1	7218	7633	14 851
2	1076	763	1839
3	117	89	206
4	58	40	98
5	47	32	79
Total	8516	8557	17 073

Binary endpoint 1

1 versus 2-5

$7218/8516 = 84.6\%$

$7633/8557 = 89.2\%$

Odds ratio = 1.49

More bleeds on bivalirudin

HERO-2: 5-level ordinal endpoint = 4 binary endpoints

Bleed severity	Bivalirudin	Heparin	Total
1	7218	7633	14 851
2	1076	763	1839
3	117	89	206
4	58	40	98
5	47	32	79
Total	8516	8557	17 073

Binary endpoint 2

1-2 versus 3-5

$8294/8516 = 97.4\%$

$8396/8557 = 98.1\%$

Odds ratio = 1.40

More bleeds on bivalirudin

HERO-2: 5-level ordinal endpoint = 4 binary endpoints

Bleed severity	Bivalirudin	Heparin	Total
1	7218	7633	14 851
2	1076	763	1839
3	117	89	206
4	58	40	98
5	47	32	79
Total	8516	8557	17 073

Binary endpoint 3

1-3 versus 4-5

$8411/8516 = 98.7\%$

$8485/8557 = 99.2\%$

Odds ratio = 1.47

More bleeds on bivalirudin

HERO-2: 5-level ordinal endpoint = 4 binary endpoints

Bleed severity	Bivalirudin	Heparin	Total
1	7218	7633	14 851
2	1076	763	1839
3	117	89	206
4	58	40	98
5	47	32	79
Total	8516	8557	17 073

Binary endpoint 4

1-4 versus 5

$8469/8516 = 99.4\%$

$8525/8557 = 99.6\%$

Odds ratio = 1.48

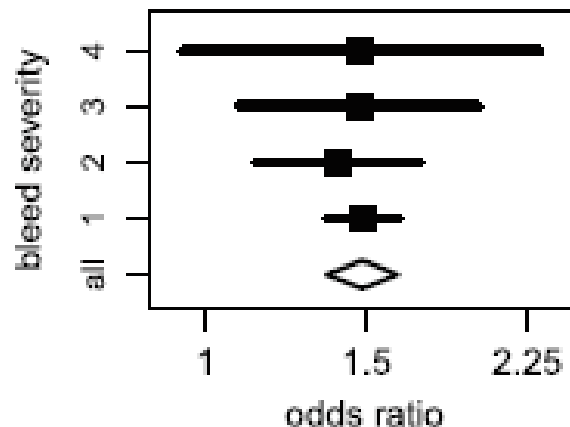
More bleeds on bivalirudin

HERO-2: Overall analysis and key assumption

Bleed severity	Bivalirudin	Heparin	Total	Cumulative difference	
				Odds ratio	95% CI
1	7218	7633	14 851	1.49	1.36–1.63
2	1076	763	1839	1.40	1.14–1.71
3	117	89	206	1.47	1.09–1.99
4	58	40	98	1.48	0.94–2.32
5	47	32	79	–	–
Total	8516	8557	17 073	–	–

Key assumption:

- All 4 treatment effects identical
- Proportional odds assumption



Key assumption in PLATIPUS

- 15-level ordinal endpoint comprised of different measures of patient experience
- Equivalent to a collection of 14 binary endpoints
- **Key assumption**
 - The treatment effect on all 14 binary endpoints is identical
 - The treatment acts identically on all components of the ordinal endpoint
 - *Mortality, neonatal admission, ventilation etc.*
- **What if the key assumption is not met?**
 - Treatment affects different components differently
 - Standard pre-specified analyses will produce an average effect that over-estimates the effect on some components and under-estimates it on others
 - E.g. effect on mortality may be over-estimated in pre-specified primary analysis

Hypothetical ordinal endpoint for WOMBAT

Maternal ordinal acuity outcome

Combination of maternal length of stay and acuity

(precise definitions of severities still to be determined)

LEVEL	Post-partum stay	Acuity
1	Short	Low
2	Long	Low
3	Any	Moderate
4	Any	Severe or death

Hypothetical ordinal endpoint for WOMBAT

Infant ordinal acuity outcome

Combination of NICU/SC length of stay and acuity
(*precise definitions of severities still to be determined*)

LEVEL	NICU/SCN stay	Acuity
1	None	Not severe
2	Short	Not severe
3	Medium	Not severe
4	Long	Not severe
5	Any	Severe or death

WOMBAT: Hypothetical binary outcomes

Both WOMBAT hypothetical ordinal endpoints can be reduced to binary endpoints

For example:

- **Maternal:**

- Levels 1-2 versus Levels 3-4
- Favourable outcome: **Low acuity**

- **Infant:**

- Level 1 versus Levels 2-5
- Favourable outcome: **Survival free of NICU/SCN admission**
- Analogous to **event-free survival** endpoints used in many diseases
 - Cancer: **Progression-free survival**
 - Cardiovascular disease: **Death or non-fatal MI**

Indicative sample size calculations for WOMBAT

Infant outcomes

Requires assumptions on the distribution of infants across levels

LEVEL	NICU/SCN stay	Acuity	Hypothetical numbers per 1000 control infants
1	None	Not severe	930 (93%)
2	Short	Not severe	35 (3.5%)
3	Medium	Not severe	20 (2.0%)
4	Long	Not severe	9 (0.9%)
5	Any	Severe or death	6 (0.6%)

Indicative sample size calculations for WOMBAT

*Infant **binary** outcome*

- Investigator preference for high power and low alpha
- 2 arm comparison with 90% power and 1% alpha (2-sided p-value < 0.01)
- Approximately equivalent to 97% power with 5% alpha
- 93% event-free survival (EFS) in control arm
- Risk difference = 2% (95% EFS in treatment arm)
Sample size = 4200 per arm, **n=8,400 total**
- Risk difference = 1.5% (94.5% EFS in treatment arm)
Sample size = 7750 per arm, **n=15,500 total**
- Risk difference = 1% (94% EFS in treatment arm)
Sample size = 18000 per arm, **n=36,000 total**
- With 5% alpha (and keeping power at 90%): **n=6000, 11000, 25500**

Indicative sample size calculations for WOMBAT

Infant ordinal outcome

- Sample size for an ordinal endpoint is smaller than for a binary endpoint
- Assuming proportional odds we can redo the illustrative sample size calculations
- Risk difference = 2% 95% vs 93% Odds ratio = 1.43
Ordinal sample size = 3600 per arm, **n=7,200 total (~15% reduction)**
- Risk difference = 1.5% 94.5% vs 93% Odds ratio = 1.29
Ordinal sample size = 6900 per arm, **n=13,800 total (~10% reduction)**
- Risk difference = 1% 94% vs 93% Odds ratio = 1.18
Ordinal sample size = 16800 per arm, **n=33,600 total (~7.5% reduction)**
- **Moderate sample size reductions in the range 7.5% to 15%**

RECOVERY

trial

philosophy:

streamlined

pragmatic

platform trials

Table 1. Key design features of the RECOVERY trial, exemplifying the application of quality-by-design principles to clinical trials and the production of rapid, reliable results

Design feature

Randomised

Large

Simple

Inclusive

Objective clinical outcomes

Linkage to national healthcare datasets

Collaborative

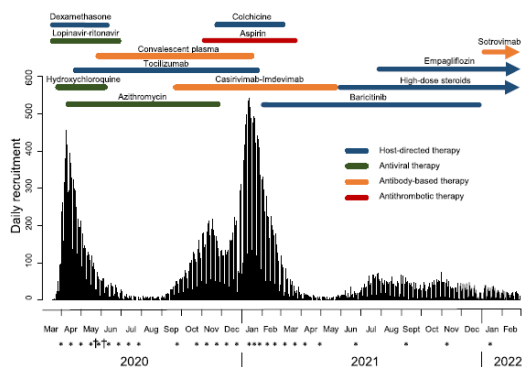
Robust web-based systems

Data quality monitoring

Trial oversight

Rationale

- > Avoidance of systematic error (bias).
- > Avoidance of play of chance.
- > Provide adequate statistical power to detect moderate, but important, treatment effects across a broad range of circumstances.
- > Focus on critical components of quality that preserve safety and reliability.
- > Do not include extraneous procedures.
- > Streamlined study procedures at set-up, training, recruitment, randomisation, treatment delivery and follow-up facilitate timely results for patients and protect busy healthcare staff.
- > Facilitates prompt recruitment.
- > Provides translatable results for all patients (both in and outside the UK).
- > Clinically relevant results.
- > Resistant to bias.
- > Allow data linkage for complete follow-up.
- > Reduced data collection by local research teams.
- > Complete follow-up improving study reliability.
- > Rapid collection of data to inform data monitoring committee analyses.
- > Low-cost long-term follow-up.
- > Buy-in from stakeholders at every level into study quality.
- > Transparency facilitates trust in study procedures, safety and results.
- > Resources are not wasted.
- > Accessible to site staff.
- > Reliable randomisation.
- > Built-in options on online forms encourage consistency and completeness of data.
- > Study procedures are subject to ongoing and continuous quality evaluation during the trial.
- > Investigators at local hospital best placed to oversee high-quality study conduct locally.
- > Experienced data monitoring committee meeting regularly provide reassurances on efficacy, accountability and safety.



Which type of outcome for WOMBAT?

- Novel ordinal endpoints can combine different types of information into the same endpoint and make efficient use of information on event hierarchies
- Statistical methods make important assumptions which may be questionable when different types of endpoints are combined in a single endpoint, which complicates pre-specification and interpretation
- When power (sample size) is limited, ordinal endpoints can be more sensitive and provide greater power (smaller sample size), under assumptions
- When a large sample size is feasible, a well powered trial with a clinically relevant binary endpoint may be preferable

Grant Outcomes E,F,G,H

- **E.** Develop the core protocol and submit it for ethics approval.
- **F.** After consultation, decide the primary composite outcome.
- **G.** Calculate sample sizes in each domain to show a 20% relative risk reduction with 90% power at $p=0.01$, using RECOVERY-style calculations.
- **H.** Harmonise data with similar trials.

Grant outcome I:

Set up governance and an Independent Data and Safety Monitoring Committee

- Professor Neena Modi has agreed to be chair of the IDSMC
- Suggestions for other members (and how many) welcome.
- What principles should an IDSMC Charter follow?
- What are its limitations? What the IDSMC can and cannot do

Grant outcome J

Develop a Knowledge Translation plan to disseminate the outcomes of the Incubator Grant Phase and National Adaptive Platform Trial

Co-design outcomes and implementation priorities with:

- Aboriginal and Torres Strait Islander representatives
- Clinicians, midwives, consumers, parent partners, policymakers

Engage national and international stakeholder groups early

Identify barriers and facilitators to implementation

Build consensus around feasible, scalable and clinically meaningful outcomes

Develop communication strategies for clinicians, consumers and health with guideline bodies and professional colleges from the outset