

**Guidance Initiative:—
a route map, and the journeying of the missing data topic group (TG1)**

James Carpenter

London School of Hygiene & Tropical Medicine *and* MRC Clinical Trials Unit

Email: james.carpenter@lshtm.ac.uk

August 28, 2013

Acknowledgements

Acknowledgements

● **Acknowledgements**

● Outline

Possible route map

Planning for the journey

Embarkation

Journeying

Discussion

Willi Sauerbrei and the Steering Group

Topic group members to date:

James Carpenter (LSHTM, UK)

Els Goethgebeur (Gent, Belgium),

Ian White (MRC Biostatistics Unit, Cambridge)

Joint chairs: James + Els +

Outline

Acknowledgements

- Acknowledgements
- **Outline**

Possible route map

Planning for the journey

Embarkation

Journeying

Discussion

1. The route map
2. Planning for the journey —
 - Topics
 - Destination
3. Embarkation
4. Journeying
 - ‘We’ve been here before...’
 - Diversions
 - Avoiding collisions
 - Arrival
5. Discussion

Initial route map

Steering Group (SG) identifies topics, agrees the brief for each Topic Group (TG), allocates SG members to TGs and agrees TG Group Chairs (ie founder members)

Founder members scope topic, agree with SG

Founder members seek to recruit further members: target size 6-8

Initial discussions on the issues to be covered, relevant existing work, and the expertise required

Initial TG teleconference:
- identify areas to cover;
- identify relevant literature;
- allocate areas to group members;
- set up central storage resource
- agree time frame

SG comments on draft

TG teleconference to discuss SG comments

Revisions by email

Drafts circulated by email

TG to agree revision

Approval by SG

Face-to-face TG meeting: discussion of draft and agreement to submit to steering group

Publication and dissemination

[Acknowledgements](#)

[Possible route map](#)

● [Initial route map](#)

[Planning for the journey](#)

[Embarkation](#)

[Journeying](#)

[Discussion](#)

Acknowledgements

Possible route map

Planning for the journey

● **Topics**

● First port of call

● Subsequently:

Embarkation

Journeying

Discussion

Topics

The steering group has identified the following topics on the basis of their importance for reliable inference from medical data:

- TG1 Missing data
- TG2 Selection of variable and functional form
- TG3 Descriptive and initial data analysis
- TG4 Measurement error
- TG5 Design
- TG6 Evaluating diagnostic tests and prediction models
- TG7 Causal inference
- TG8 ...
- TG9 ...
- TG10 ...

The numbering of the topic groups does not represent either the relative importance of the topics or the order in which they should be considered!

First port of call

Acknowledgements

Possible route map

Planning for the journey

● Topics

● **First port of call**

● Subsequently:

Embarkation

Journeying

Discussion

- Level 2: Experienced statistician
(typically Masters level plus work experience
= ‘somewhat below state-of-the-art’)
 - outline issues
 - outline methodological approaches, and their pros and cons
 - guidance on software

Output:

1. SG publication about initiative aims *etc*
2. For each TG:
 - a longer report (around 10,000 words) on the Guidance Initiative’s website—open for external comments, and
 - publications deriving for this.

Subsequently:

Acknowledgements

Possible route map

Planning for the journey

- Topics
- First port of call
- **Subsequently:**

Embarkation

Journeying

Discussion

Guidance for:

1. Level 1: Less experienced analysts, eg
 - those without formal statistical training
 - recently qualified Masters students
2. Level 3: Experienced statisticians, eg
 - active researchers
 - research group leaders

Level 1: probably the same TG

Level 3: some additional experts maybe added

Assembling Topic Groups

Acknowledgements

Possible route map

Planning for the journey

Embarkation

● **Assembling Topic Groups**

● Missing data TG

● Recruitment

Journeying

Discussion

TG founding members

- liaise with Steering Group on initial scope
- invite researchers to participate ↔ finalize scope

Acknowledgements

Possible route map

Planning for the journey

Embarkation

- Assembling Topic Groups

- **Missing data TG**

- Recruitment

Journeying

Discussion

Missing data TG

Scope (principally clinical, but also other application areas):

- Likely effects of missing data:
 - marginal statistics;
 - regression;
 - implications for more complex settings....
- Methodology:
 - Complete Records
 - Multiple imputation
 - EM algorithm
 - Inverse probability weighting
 - Combinations of the above
- Reporting

Recruitment

Acknowledgements

Possible route map

Planning for the journey

Embarkation

- Assembling Topic Groups

- Missing data TG

- **Recruitment**

Journeying

Discussion

We seek to recruit a 6–8 applied researchers with an established interest in missing data, who between them are

- representative of the range of methodological approaches;
- able to collaborate effectively, and
- internationally representative.

Recruitment by

- initial e-mail (brief)
- subsequent discussion, and more formal letter giving full details.

Thus far:

Ian White (MRC-BSU Cambridge) - agreed

One 'positive refusal'

One pending...

'We've been here...'

Acknowledgements

Possible route map

Planning for the journey

Embarkation

Journeying

● 'We've been here before...'

● ...or have we?

● Diversions

● E.g. sensitivity analysis for missing data

● 'Are we nearly there yet?'

● Possible missing data guidance structure

Discussion

Do existing review papers in a particular area mean there is no need for further work—just dissemination?

E.g. missing data in clinical trials, where in 2010 the National Academies Press published a report 'The prevention and Treatment of Missing Data in Clinical Trials' prepared by a panel of experts chaired by Rod Little [2].

This

- convened an expert group;
- took evidence;
- wrote a report with recommendations for action, and
- published commentaries published in various journals

Further, the European Medicines Association published a 'Guideline on Missing Data in Confirmatory Clinical Trials' (2010)[1].

What do we add to these?

Acknowledgements

Possible route map

Planning for the journey

Embarkation

Journeying

● 'We've been here before...'

● ...or have we?

● Diversions

● E.g. sensitivity analysis for missing data

● 'Are we nearly there yet?'

● Possible missing data guidance structure

Discussion

...or have we?

- the US National Academies of Science report did not aim to be a guideline; rather it was commissioned by the FDA to inform their guidelines.
- The EMEA guidelines are closer to our remit, but—partly because of process—are less statistically credible.
- Neither
 - has focus on observational data;
 - sets out to provide a systematic guide for analysts at an appropriate level;
 - is part of a unified initiative;
 - has a web-based home providing an up-to-date links to the literature, and documenting comments

The aim is an accessible document raising awareness of the issues to be considered and how various methods may be applied, not a cookbook.

Diversions

Acknowledgements

Possible route map

Planning for the journey

Embarkation

Journeying

- 'We've been here before...'

- ...or have we?

- **Diversions**

- E.g. sensitivity analysis for missing data

- 'Are we nearly there yet?'

- Possible missing data guidance structure

Discussion

Each topic has many threads, relating to specific adaptations of the methodology for particular application areas. Clearly it is not possible to relate to all these.

There is, though a limited number of key concepts that inform these developments. Awareness of these, how they relate, their pros and cons, and their accessibility with current software, are all key areas for guidance.

It may be helpful to consider:

- presentation: a limited number of specific quite detailed examples (esp levels 1 and 2).
- evidence production: common simulation designs for comparisons (level 3)

E.g. sensitivity analysis for missing data

Acknowledgements

Possible route map

Planning for the journey

Embarkation

Journeying

- 'We've been here before...'

- ...or have we?

- Diversions

- E.g. sensitivity analysis for missing data

- 'Are we nearly there yet?'

- Possible missing data guidance structure

Discussion

Need to accessibly convey why this is important, and this almost certainly needs an example.

Then present the three methodological streams:

- selection factorisation / modelling
- pattern mixture factorisation / modelling
- shared parameter modelling

Show what each means for the example.

Discuss pros and cons of each in context of example and audience.

Give software suggestions for each.

Give suggestions for further reading.

'Are we nearly there yet?'

Acknowledgements

Possible route map

Planning for the journey

Embarkation

Journeying

- 'We've been here before...'

- ...or have we?

- Diversions

- E.g. sensitivity analysis for missing data

- 'Are we nearly there yet?'

- Possible missing data guidance structure

Discussion

TG members inevitably have different opinions, interests, and personal investments in different approaches.

Most also feel some pressure to publish, and may become impatient with an open ended initiative.

So it's going to be important to have

- clearly defined phases;
- a plausible timetable for each phase, indicating the commitment involved, and
- good communication within the TG at each stage.

It's better for potential TG members to be aware of these and not sign up, rather than sign up and then be dissatisfied at a later date.

Possible missing data guidance structure

Acknowledgements

Possible route map

Planning for the journey

Embarkation

Journeying

- 'We've been here before...'

- ...or have we?

- Diversions

- E.g. sensitivity analysis for missing data

- 'Are we nearly there yet?'

- Possible missing data guidance structure

Discussion

1. Issues raised by missing data
2. Analysis aims
3. Methodologies, including critique
4. Software
5. Reporting
6. Examples

Probably a 10,000 word document.

Suggested timeframe: 18 months.

Acknowledgements

Possible route map

Planning for the journey

Embarkation

Journeying

Discussion

● **SG aims:**

● Summary

● References

SG aims:

At the end of today, the SG hopes to have

1. convinced you the initiative is worthwhile
2. got some good suggestions for taking it forward, which will input into a more detailed form for the overall structure, and
3. interested you in signing up to a Topic Group!

Summary

Acknowledgements

Possible route map

Planning for the journey

Embarkation

Journeying

Discussion

- SG aims:
- **Summary**
- References

Progress so far

- SG in place; key TGs identified and chairs in place
- TGs recruiting, in parallel with initial scoping of work
- ...

Challenges

- SG agreeing—and appropriately formalising—an overall structure for the TGs to follow; together with an appropriate timetable;
- SG providing appropriate guidance to ensure cohesiveness of resulting guidance documents;
- TG recruitment
- Effectively monitoring progress and troubleshooting...
- Obtaining funding for ‘central services’

Acknowledgements

Possible route map

Planning for the journey

Embarkation

Journeying

Discussion

- SG aims:
- Summary
- **References**

References

- [1] Committee for Medicinal Products for Human Use (CHMP). *Guideline on Missing Data in Confirmatory Clinical Trials*. London: European Medicines Agency; download from www.emea.europa.eu (accessed), 2010.
- [2] National Research Council. *The prevention and Treatment of Missing Data in Clinical Trials*. Panel on Handling Missing Data in Clinical Trials. Committee on National Statistics, Division of Behavioral and Social Sciences and Education. Washington, DC: The National Academies Press., 2010.