

My experience with “Open Source” modelling

Howard Thom

Senior Lecturer, University of Bristol

Howard.thom@Bristol.ac.uk



Why the apostrophes??

I'm really talking about "**Public Domain**" models generally, rather than "**Open Source**" models specifically released under a license.

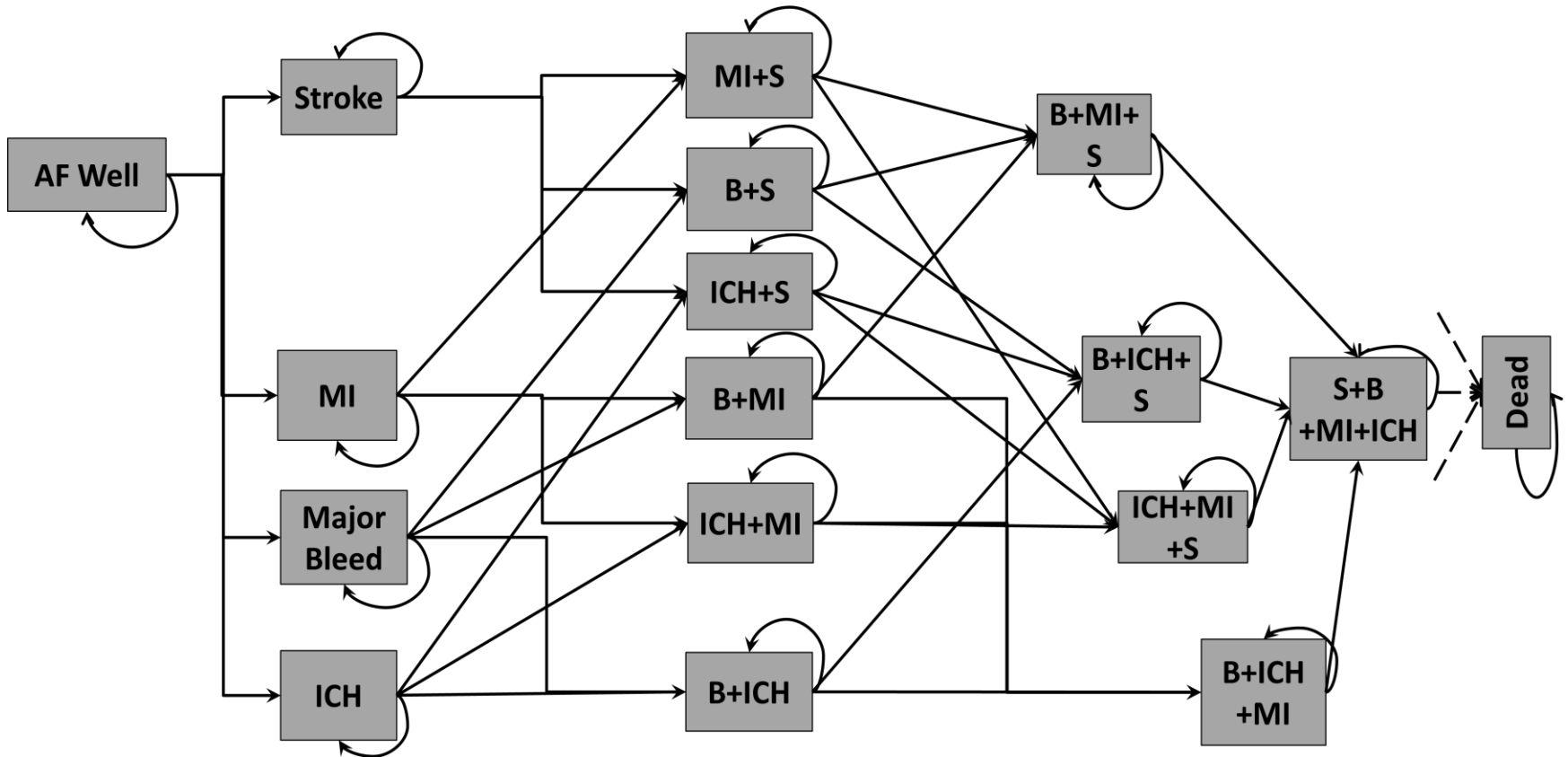
Now that that's cleared up...

Why am I talking to you?

- My research is primarily in network meta-analysis on limited evidence networks and value of information analysis
- With Gianluca Baio, I co-chair the R for Health Technology Assessment (HTA) scientific committee (<https://r-hta.org/>)
- We push for the adoption of “open source” R programming instead of Excel in HTA
- In 2015 I developed an R Markov model comparing directly acting oral anticoagulants (DOACs) for the prevention of stroke in Atrial Fibrillation (AF)



The AF Model – not the prettiest




S=Ischemic stroke, B=Major extracranial bleed, MI=Myocardial infarction, ICH=Intracranial haemorrhage
TIA and SE are transient events. Patients can switch treatment following events.

Widely published

HEALTH TECHNOLOGY ASSESSMENT


VOLUME 21 ISSUE 9 MARCH 2017
ISSN 1366-5278

 Check for updates



 Check for updates

Article

MDMP
Policy & Practice

MDM Policy & Practice
1–14
© The Author(s) 2019
Article reuse guidelines:
sagepub.com/journals-permissions
DOI: 10.1177/2381468319866828
journals.sagepub.com/home/mdm


Directly Acting Oral Anticoagulants for the Prevention of Stroke in Atrial Fibrillation in England and Wales: Cost-Effectiveness Model and Value of Information Analysis

Howard H. Z. Thom , Will Hollingworth, Reecha Sofat, Zhenru Wang, Wei Fang, Pritesh N. Botalia, Peter A. Bryden, Philippa A. Davies, Deborah M. Caldwell, Sofia Dias , Diane Eaton, Juli Aroon D. Hingorani, Jose A. Lopez-Lopez, George N. Alison Richards, Chris Salisbury, Jelena Savović, Anny Jonathan A. C. Sterne, and Nicky J. Welton

Abstract

Objectives. Determine the optimal, licensed, first-line anticoagulant non-valvular atrial fibrillation (AF) in England and Wales from and estimate value to decision making of further research. **Methods.** Compare warfarin (international normalized ratio target range 2–3) oral anticoagulants (DOACs) apixaban 5 mg, dabigatran 150 mg twice daily, edoxaban 60 mg once daily, and rivaroxaban 20 mg once daily. In addition to death, the 17-sta

 OPEN ACCESS

Oral anticoagulants for prevention of stroke in atrial fibrillation: systematic review, network meta-analysis, and cost effectiveness analysis

José A López-López,¹ Jonathan A C Sterne,^{1,2} Howard H Z Thom,¹ Julian P T Higgins,^{1,2} Aroon D Hingorani,³ George N Okoli,¹ Philippa A Davies,^{1,4} Pritesh N Botalia,^{5,6} Peter A Bryden,¹ Nicky J Welton,^{1,2} William Hollingworth,¹ Deborah M Caldwell,¹ Jelena Savović,^{1,4} Sofia Dias,¹ Chris Salisbury,¹ Diane Eaton,⁷ Annya Stephens-Boal,⁸ Reecha Sofat³

ABSTRACT OBJECTIVE

To compare the efficacy, safety, and cost effectiveness of direct acting oral anticoagulants (DOACs) for patients with atrial fibrillation.

DESIGN

Systematic review, network meta-analysis, and cost effectiveness analysis.

rivaroxaban 20 mg once daily (1.35, 1.03 to 1.78) than with dabigatran 150 mg twice daily. The risk of all-cause mortality was lower with all DOACs than with warfarin. Apixaban 5 mg twice daily (0.71, 0.61 to 0.81), dabigatran 110 mg twice daily (0.80, 0.69 to 0.93), edoxaban 30 mg once daily (0.46, 0.40 to 0.54), and edoxaban 60 mg once daily (0.78, 0.69 to 0.90) reduced the risk of major bleeding compared

Oral anticoagulants for primary prevention, treatment and secondary prevention of venous thromboembolic disease, and for prevention of stroke in atrial fibrillation: systematic review, network meta-analysis and cost-effectiveness analysis

Jonathan AC Sterne, Pritesh N Botalia, Peter A Bryden, Philippa A Davies, Jose A Lopez-Lopez, George N Okoli, Howard H Z Thom

RESEARCH

For numbered affiliations see end of article.

Correspondence to: J A C Sterne
jonathan.sterne@bristol.ac.uk

Additional material is published online only. To view please visit the journal online.

Cite this as: *BMJ* 2017;359:j5058
<http://dx.doi.org/10.1136/bmj.j5058>

Importantly – I uploaded it to GitHub

- This was my first upload of a model to GitHub

<https://github.com/Bogdasayen/DOACs-AF-Economic-model>

Why did I do it?

- I wanted to be transparent
- No commercial interest in the model
- I wanted to teach and learn from others...

Open Source \neq Transparency

- First strike: not a project.
 - Programmed in base R and not RStudio
- At least the model is modularized into functions, split across multiple scripts, and uses a consistent style throughout
- And I included lots of comments
- Opening lines explain what changes have been made and thus the ".X" after every script.
- Not exhaustive though and fuller use of GitHub commit comments would have been better.

Open Source \neq Transparency

- Not *tidyverse* style.
- For example, `n.cycles` instead of `n_cycles`.
- Tabs are used to indicate function/loop level but there are no aesthetic spaces.

- For example

```
total.qalys <- matrix(0, nrow = n.samples, ncol = n.treatments - 1))
```

- instead of

```
total.qalys <- matrix(0, nrow = n.samples, ncol - n.treatments - 1))
```


Open Source \neq Transparency

Tangle of file names

- `NOAC.AF.model.main.3.R`: Primary script which runs main and sensitivity analyses
- `NOAC.AF.net.benefit.3.R`: Core Markov model
- `generate.transition.matrix.7.R`: Actually, has functions to generate some input parameters (`age.independent.generate.probabilities`) as well as the transition matrices (`generate.probabilities`)
- Neither function is actually called `generate.transition.matrix()` or really describes what the function does, so needlessly confusing
- There are also various utility scripts/functions like `next.state.name()` and `generate.hr.death()`.

Open Source \neq Transparency

As an example of my very transparent coding...

```
# Change the probability.matrix into transition.matrix
# Need rules to assign event probabilities to state probabilities
# These are conditional on i.treatment and i.health.state
transition.matrix[[age]]<-array(0,dim=c(n.samples,n.states,n.states))
for(i.treatment in 1:n.treatments)
{
  for(i.health.state in 1:n.health.states)
  {
    # The old state is this
    old.state.name<-paste(treatment.names[i.treatment],nondeath.health.states[i.health.state])
    i.state<-which(state.names==old.state.name)
    # Need the name and index of the state patients in this state go
    # to if they have each event.state.codes

    # Find the name and index (in the state.names vector) of the new state (a bit messy)
    new.state.indices<-new.state.name<-rep(NA,n.events)
```

I'm not sure I could debug this anymore...

```
# State following Clinically relevant bleeding (B)
new.state.name[event.state.codes==" B "]<-next.state.name(old.state.name," B ")
new.state.indices[event.state.codes==" B "]<-which(next.state.name(old.state.name," B ")==state.names)
```

```
# If no discontinuation/switching
```

```
transition.matrix[[age]][,i.state,new.state.indices[event.state.codes==" B "]]<-
transition.matrix[[age]][,i.state,new.state.indices[event.state.codes==" B "]]+
probability.matrix[[age]][,i.state,event.state.codes==" B "]*(1-
event.switch.probs[,event.state.codes==" B "])
```

```
# If discontinuation/switching
```

```
transition.matrix[[age]][,i.state,new.state.indices[event.state.codes==" B "]]+
treatment.switch.indices[i.treatment]<-
transition.matrix[[age]][,i.state,new.state.indices[event.state.codes==" B "]]+
treatment.switch.indices[i.treatment]+probability.matrix[[age]][,i.state,event.state.codes==" B "]*
event.switch.probs[,event.state.codes==" B "]
```

What happened – upsides

- Constructive comments from others on my coding style, which have helped me to improve my skills
- Many invitations to collaborate
- Most importantly, the model was used in the 2021 NICE Clinical Guidelines on DOACs
 - Updated the model to better account for stroke risk and use registry evidence for baseline probabilities

The screenshot shows a web browser window with the URL <https://www.nice.org.uk/guidance/ng196>. The page header includes the NICE logo (National Institute for Health and Care Excellence), a search bar, and a 'Sign in' button. The main navigation menu is dark green with white text, listing: NICE Pathways, NICE guidance (selected), Life sciences, Standards and indicators, Evidence search, BNF, BNFC, CKS, and Journals and databases. A yellow banner below the navigation menu reads: 'Read about [our approach to COVID-19](#)'. The breadcrumb trail is: Home > NICE Guidance > Conditions and diseases > Cardiovascular conditions > Heart rhythm conditions. The main heading is 'Atrial fibrillation: diagnosis and management'. Below the heading, it says 'NICE guideline [NG196] Published: 27 April 2021 Last updated: 30 June 2021'. At the bottom, there is a secondary navigation menu with tabs: Guidance (selected), Tools and resources, Information for the public, Evidence, and History.

What happened – one big downside

- DOACs are highly commercially sensitive
- Pharma and consultants dissecting the code for errors
- And **yes** there were errors, just none that affected conclusions
 - Probability of switching after stroke should be 30% but coded as 10%
 - Multiple definitions of redundant objects
- These are helpful to know but weren't identified to be helpful.
- Instead used to feed a narrative that the model was somehow deficient.

- Naïve of me not to anticipate the asymmetric interest in conclusions:
 - Little scientific difference if treatment X or treatment Y is most cost-effective for preventing stroke, so long as it really is most cost-effective
 - But this is a vital difference for commercial stakeholders.

What I've learned

- Use projects, TidyVerse style, and useful naming of functions and scripts
- All code has errors – that's why you test and validate
- The benefits of Open Source modelling outweighed the costs, even in this commercially sensitive case study
- A better understanding of the asymmetric interests of academic and commercial modellers

Acknowledgments

- *The HTMR Collaboration and innovation in Difficult and Complex randomised controlled Trials In Invasive procedures (ConDuCT-II) hub provided support for this research.*
- *This study was supported by the National Institute for Health Research (NIHR) Biomedical Research Centre (BRC) at the University Hospitals Bristol National Health Service (NHS) Foundation Trust and the University of Bristol, and the NIHR Manchester BRC.*
- *Funding for the Atrial Fibrillation model was provided by NIHR HTA grant 11/92/17 and NIHR Senior Investigator award (to Jonathan Sterne) NF-SI-0611-10168.*



Thank you!