Piloting peer code review in a research consortium community of practice

Saras Windecker

The KIDS RESEARCH INSTITUTE AUSTRALIA

ACEFA
AUSTRALIA-AOTEAROA CONSORTIUM FOR EPIDEMIC FORECASTING & ANALYTICS

I wish to acknowledge that I am presenting today from the unceded land of the Wurundjeri Woi Wurrung and Bunurong peoples of the Kulin Nation. I live and work on the lands of the Turrbul and Jagera peoples. I pay my respect to all their Elders past and present.

Who are we?

The Australia-Aotearoa Consortium for Epidemic Forecasting and Analytics

Who are we?

The Australia-Aotearoa Consortium for Epidemic Forecasting and Analytics

Research groups from across Australia and New Zealand aiming to support **timely, effective response** to epidemic disease.

respiratory infection
transmission

problem

problem

what is the disease burden?

0110
0111

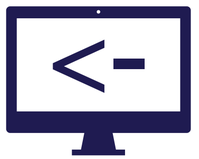
questions
& data

(e.g. cases, hospital admissions, deaths)

problem

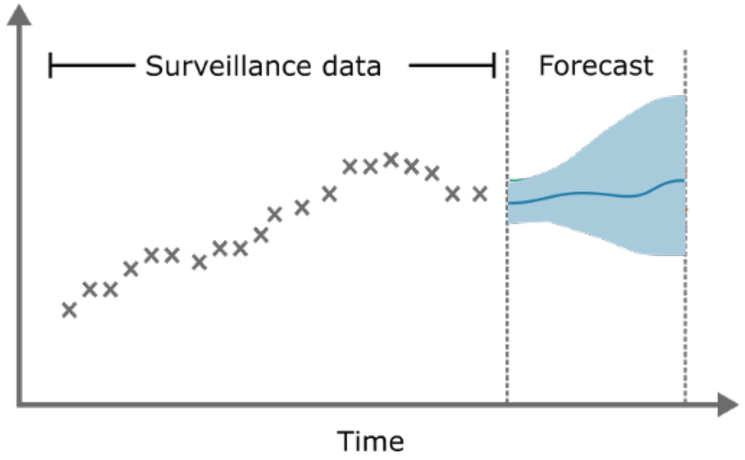
questions
& data

(e.g. auto-regressive model,
renewal model)

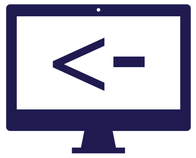


analysis &
code

problem



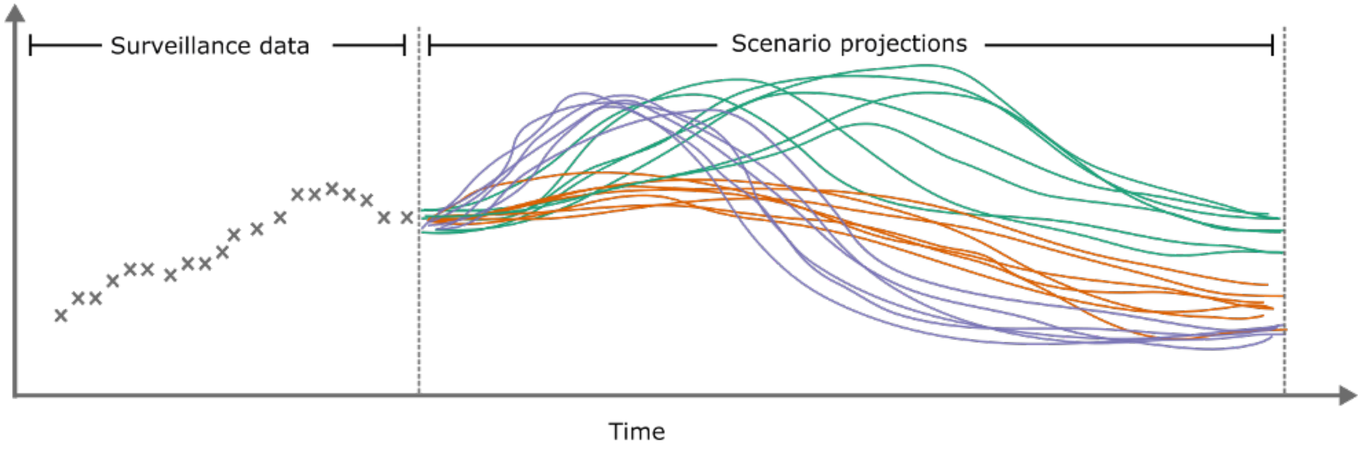
questions & data



analysis & code

outputs

problem



questions & data

analysis & code

outputs

problem

questions & data

analysis & code

outputs

recommendations

expected hospital admission
short-term case forecast

problem

questions
& data

analysis &
code

outputs

recommendations

Reliable recommendations

problem

questions & data

analysis & code

outputs

recommendations

Reliable recommendations
depend on **quality control**

problem

questions
& data

analysis &
code

outputs

recommendations

Reliable recommendations depend on quality control **throughout research cycle**

problem

questions & data

analysis & code

outputs

recommendations

Ensuring reliable research…

isn’t that where **peer review** comes in??

Ensuring reliable research…

isn’t that where **peer review** comes in??

SCIENTIFIC PUBLISHING

A Scientist’s Nightmare: Software Problem Leads to Five Retractions

The Credibility Crisis In Scientific Publishing

A Texas A&M public health researcher examines journal research quality assurance methods in the field of addiction.

By Rae Lynn Mitchell, Texas A&M University School of Public Health • JULY 29, 2019

problem

questions & data

analysis & code

outputs

recommendations

What does peer review check?

problem

questions & data

analysis & code

outputs

recommendations

What does peer review check?

problem

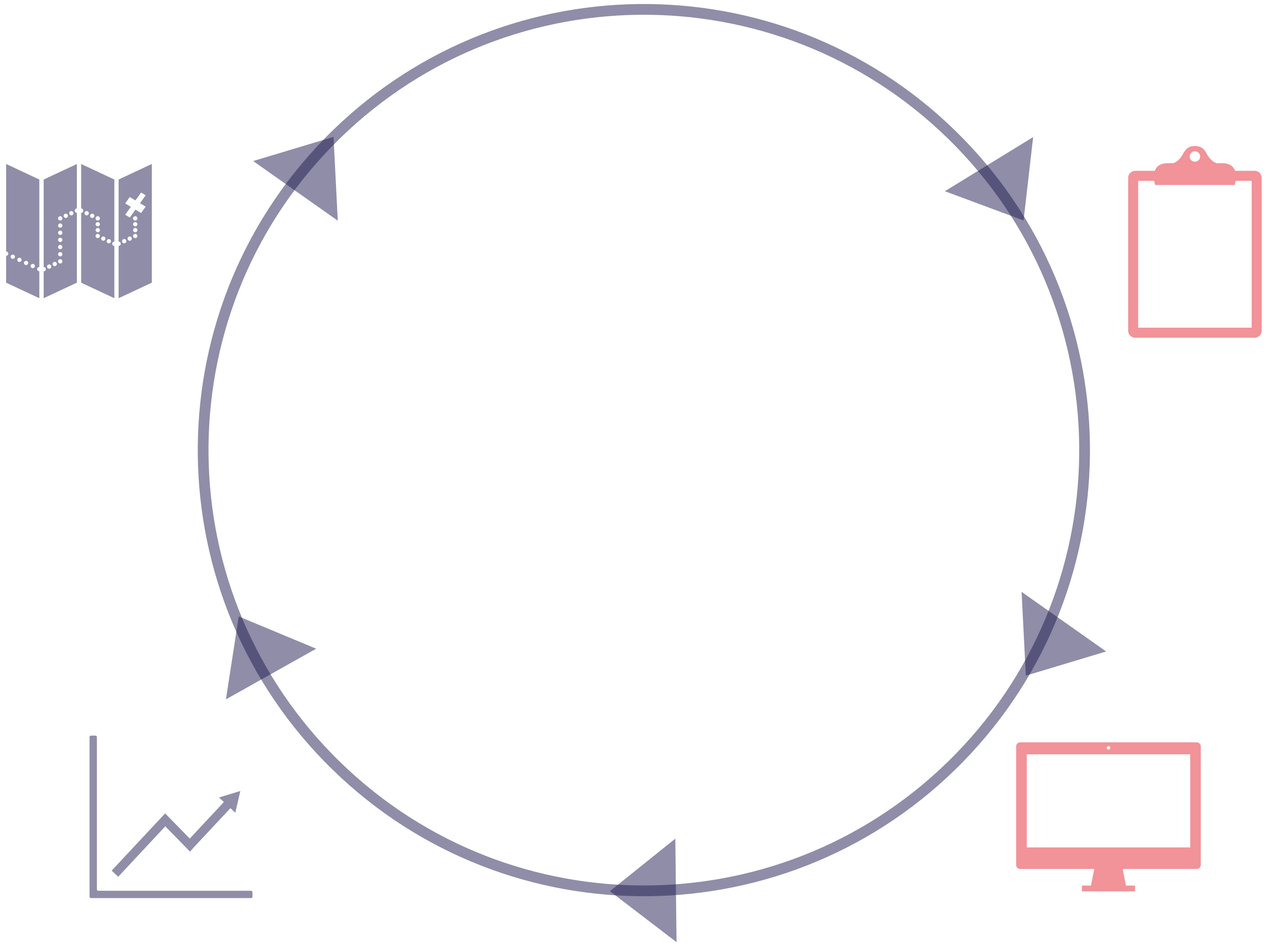
questions & **data**

analysis & **code**

outputs

recommendations

What does peer review check?

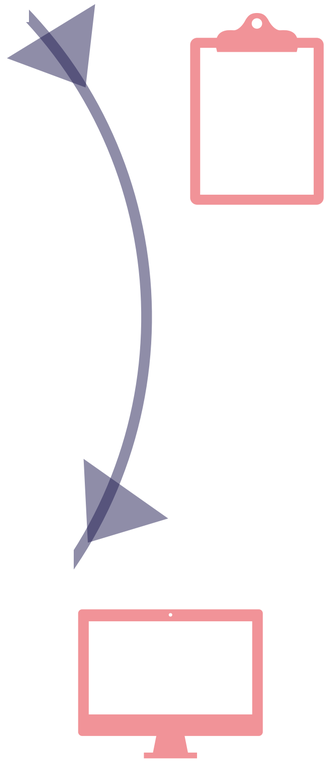


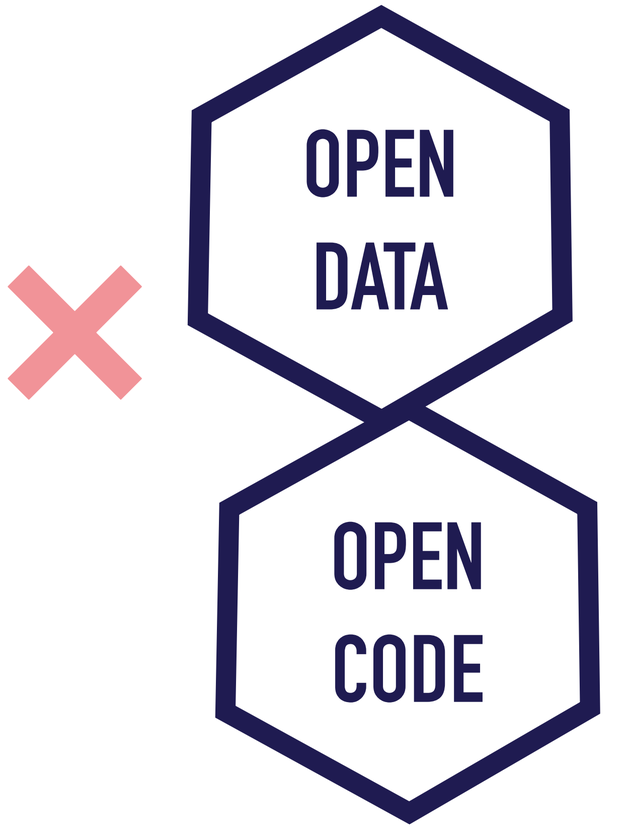
OPEN
DATA

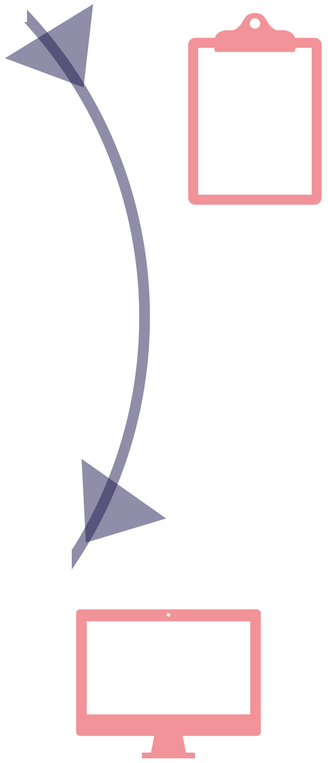
OPEN
CODE

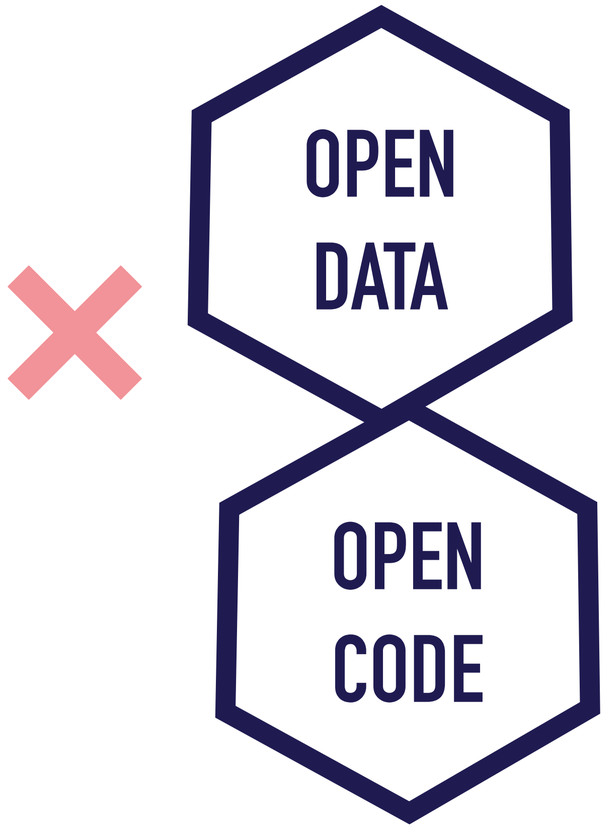
DRYAD

zenodo









Biological science practices

From policy to practice: progress towards data- and code-sharing in ecology and evolution

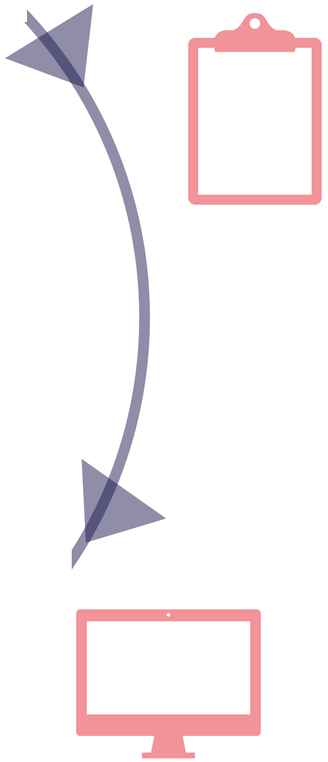
Edward R. Ivimey-Cook[†], Alfredo Sánchez-Tójar[†], Ilias Berberi, Antica Culina, Dominique G. Roche, Rafaela A. Almeida, Bawan Amin, Kevin R. Bairos-Novak, Heikel Balti, Michael G. Bertram, Louis Bliard, Ilha Byrne, ... See all authors

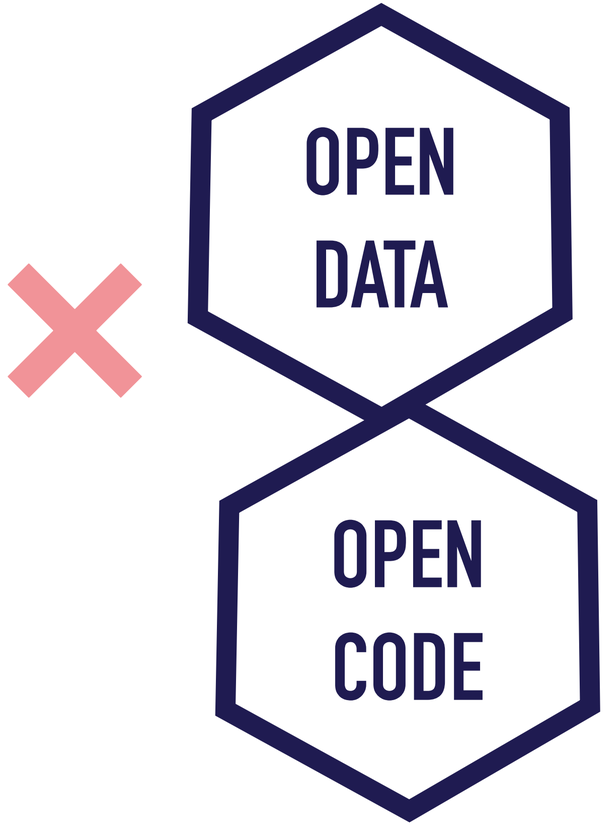
Published: 17 September 2025 https://doi.org/10.1098/rspb.2025.1394

DRYAD

zenodo

goal == transparency





Biological science practices

From policy to practice: progress towards data- and code-sharing in ecology and evolution

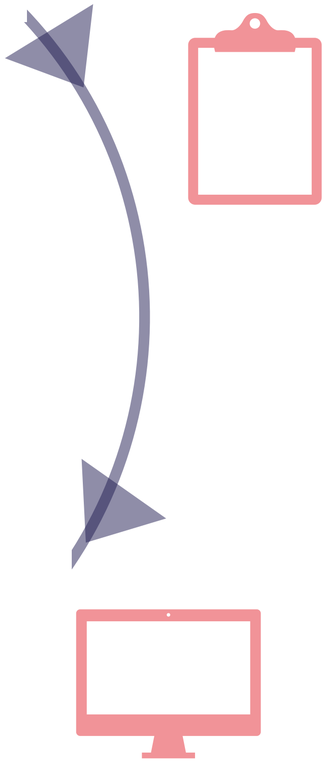
Edward R. Ivimey-Cook†, Alfredo Sánchez-Tójar†, Ilias Berberi, Antica Culina, Dominique G. Roche, Rafaela A. Almeida, Bawan Amin, Kevin R. Bairos-Novak, Heikel Balti, Michael G. Bertram, Louis Bliard, Ilha Byrne, ... See all authors

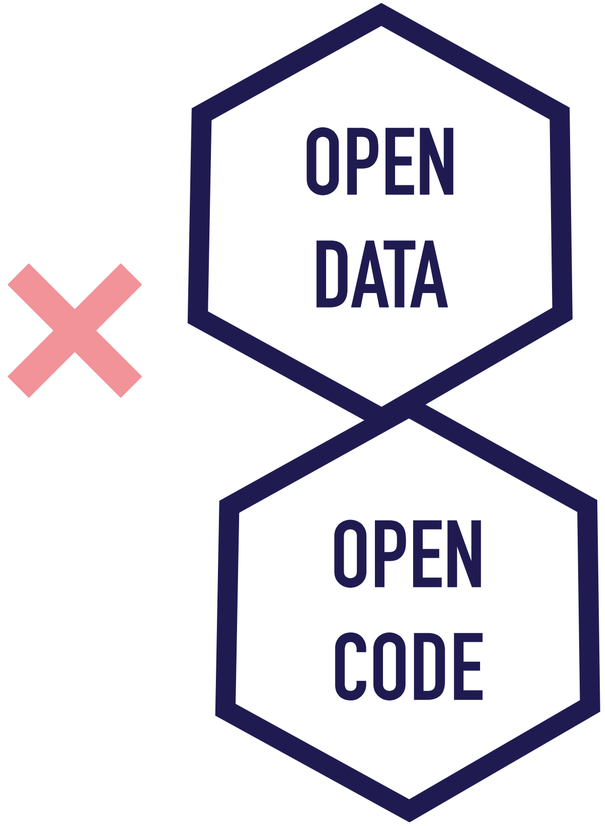
Published: 17 September 2025 https://doi.org/10.1098/rspb.2025.1394

DRYAD

zenodo

transparency is not enough





Biological science practices

From policy to practice: progress towards data- and code-sharing in ecology and evolution

Edward R. Ivimey-Cook[†], Alfredo Sánchez-Tójar[†], Ilias Berberi, Antica Culina, Dominique G. Roche, Rafaela A. Almeida, Bawan Amin, Kevin R. Bairos-Novak, Heikel Balti, Michael G. Bertram, Louis Bliard, Ilha Byrne, ... See all authors

Published: 17 September 2025 https://doi.org/10.1098/rspb.2025.1394

Code review is part of the answer

Code review is part of the answer

Code review is **basic best practice** for software development.

But that is **not yet the case** for code related to research.

Code review is part of the answer

Code review is basic best practice for software development.

But that is not yet the case for code related to research.

Regular academic practice == layers of review before submission

Code review also needs to happen during development

Many limitations to code review

Limited by:

Many limitations to code review

Limited by:

- Identifying reviewers with the right **context knowledge**,

Many limitations to code review

Limited by:

- Identifying reviewers with the right context knowledge,
- **Technical expertise**,

Many limitations to code review

Limited by:

- Identifying reviewers with the right context knowledge,
- Technical expertise,
- and **time**

Many limitations to code review

Limited by:

- Identifying reviewers with the right context knowledge,
- Technical expertise,
- and time
- And **code authors** themselves are still **less prepared to open up**

Many limitations to code review

Limited by:

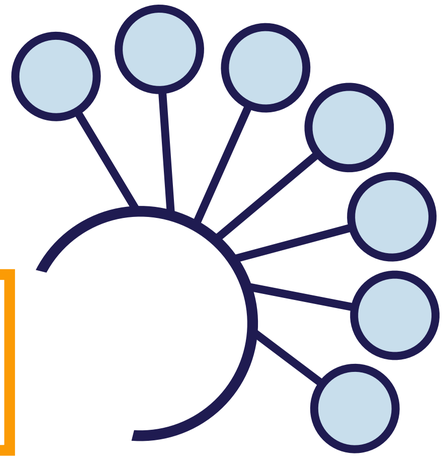
- Identifying reviewers with the right context knowledge,
- Technical expertise,
- and time
- And **code authors** themselves are still **less prepared to open up**

How can we address these limitations?

Many limitations to code review

Limited by:

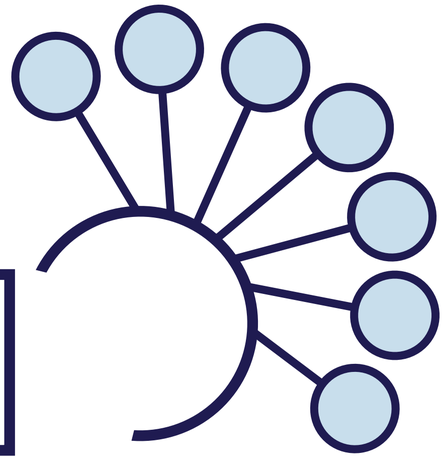
- Identifying reviewers with the right context knowledge,
- Technical expertise,
- and time
- And **code authors** themselves are still **less prepared to open up**



Foster culture of code review within communities

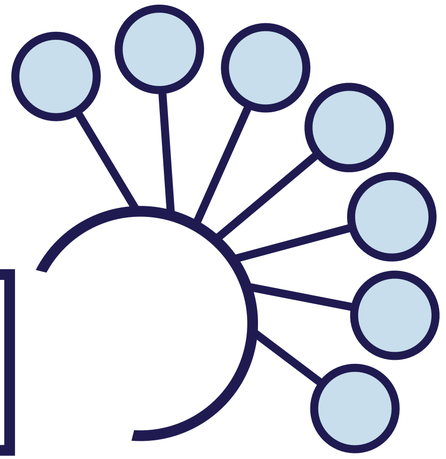
Context knowledge

Foster culture of code review within communities



Context knowledge

✓ **Commitment to one another**



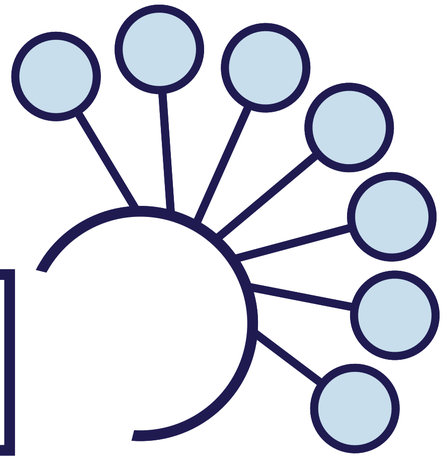
Foster culture of code review within communities

Context knowledge

Commitment to one another

Trust

Foster culture of code review within communities



problem

questions
& data

analysis &
code

outputs

recommendations

Our community of practice

We want to **foster a community** per the aims of the CoP:

"coming together to learn, share knowledge, and network to

improve their practice and expertise"

Our community of practice

We want to foster a community per the aims of the CoP:

> “coming together to learn, share knowledge, and network to
>
> improve their practice and expertise”

We have a unique opportunity. In many instances, implementing code review is challenging because of being limited by having people with the right expertise, context knowledge, and time to dedicate to reviewing.

We have an **ideal context to trial code review protocols**.

Our plan

Review checklist

Ability to understand the code, run the code, reproduce results, follow best practices

Collaborative code refinement

But we won’t get it right the first time

Review checklist

Ability to understand the code, run the code, reproduce results, follow best practices

Collaborative code refinement

But we won’t get it right the first time

1. **Pre-review survey**
   An initial survey to reviewers and code authors to self assess baseline exposure and interest.
2. **Review checklist**
   Ability to understand the code, run the code, reproduce results, follow best practices
3. **Collaborative code refinement**
4. **Post-review survey**
5. **Revision for the future**

We’d love your feedback

Saras Windecker & Rob Moss

Please get in touch with any suggestions

saras.windecker@thekids.org.au

rgmoss@unimelb.edu.au

The KIDS RESEARCH INSTITUTE AUSTRALIA

ACEFA

AUSTRALIA-AOTEAROA CONSORTIUM FOR EPIDEMIC FORECASTING & ANALYTICS

problem

questions
& data

individual model

analysis &
code

outputs

recommendations

problem

questions
& data

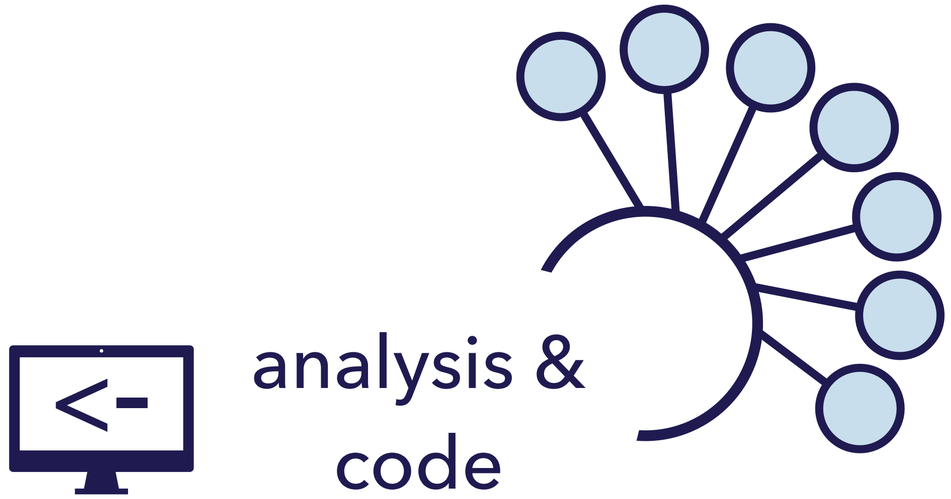
multiple models

analysis &
code

outputs

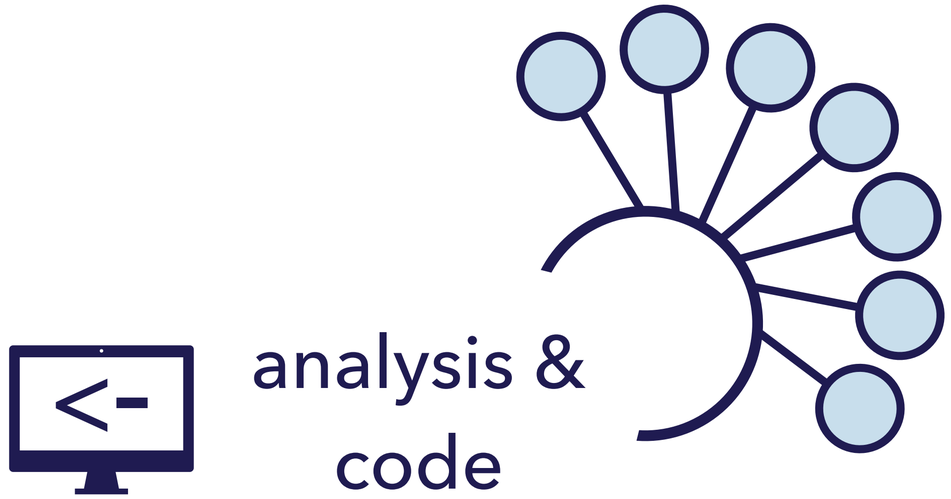
recommendations

Why use multiple models?



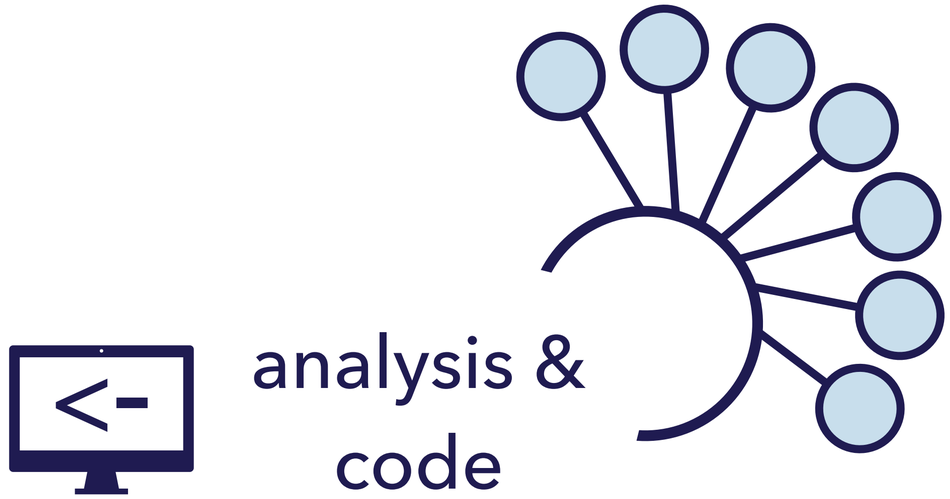
Why use multiple models?

- Highest standard of model-based policy evidence



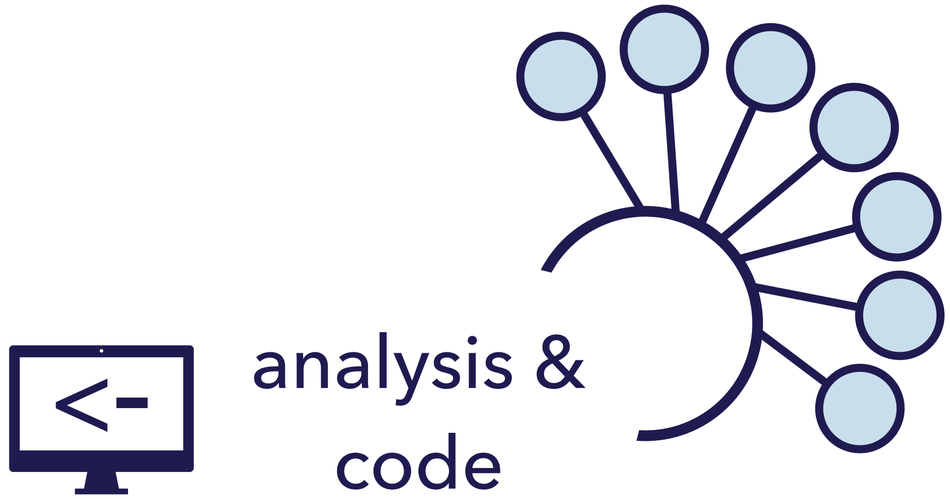
Why use multiple models?

- Highest standard of model-based policy evidence
- Ensemble forecasts consistently outperform any individual component model



Why use multiple models?

- Highest standard of model-based policy evidence
- Ensemble forecasts consistently outperform any individual component model
- Multi-model scenario projections capture variation in outcomes and how they depend on interventions and uncertainties.



Why use multiple models?

- Highest standard of model-based policy evidence
- Ensemble forecasts consistently outperform any individual component model
- Multi-model scenario projections capture variation in outcomes and how they depend on interventions and uncertainties.
- Extensively used by public health agencies to inform policy decisions (ECDC, US CDC, UK Gov, WHO)