



HAL
open science

Towards reproducible neuroimaging across different analysis pipelines

Camille Maumet

► **To cite this version:**

Camille Maumet. Towards reproducible neuroimaging across different analysis pipelines. CogBases 2023, Oct 2023, Paris, France. 2023. hal-04236791

HAL Id: hal-04236791

<https://inria.hal.science/hal-04236791>

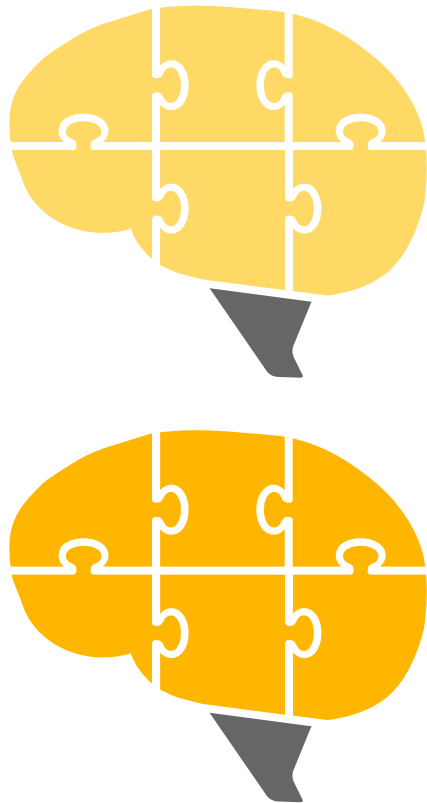
Submitted on 11 Oct 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution 4.0 International License



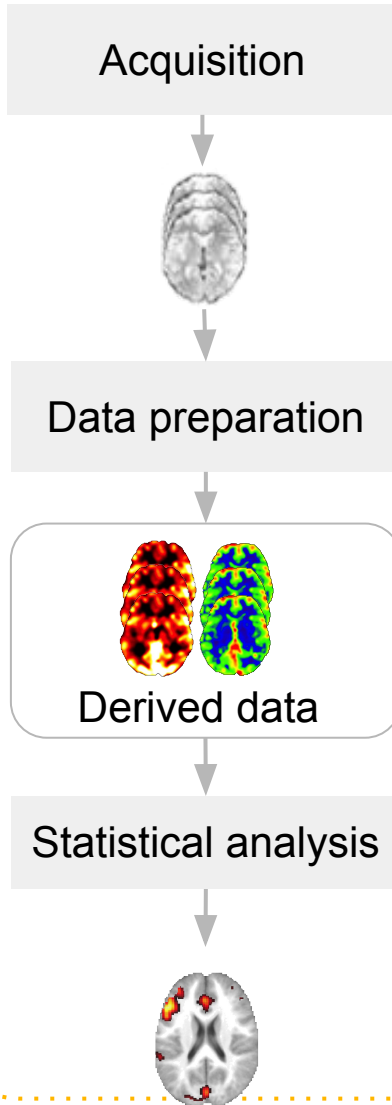
Towards reproducible neuroimaging across different analysis pipelines

CogBases workshop 2023

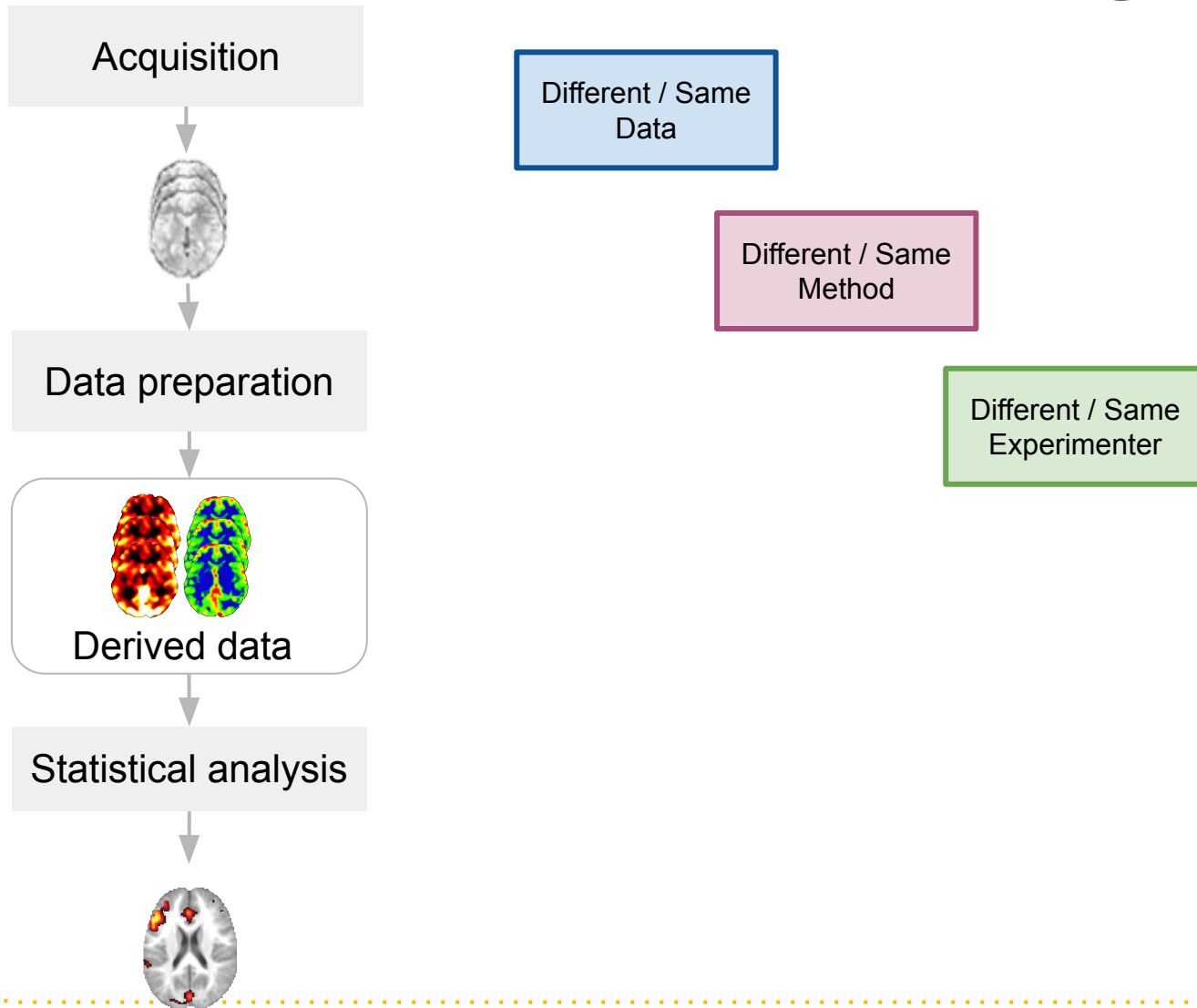
Camille Maumet

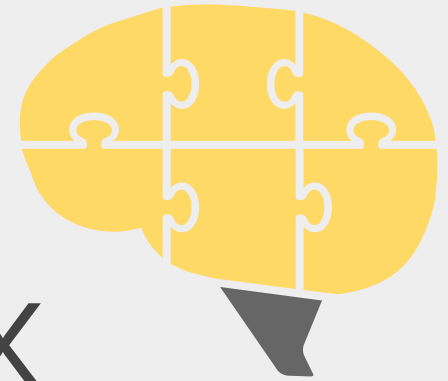
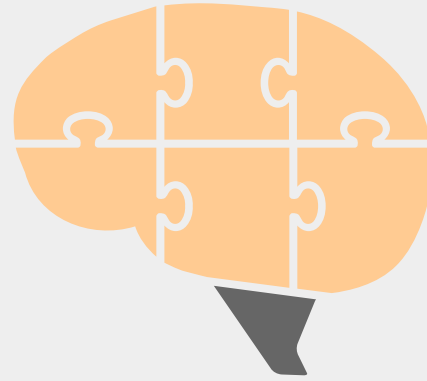
Inria, Univ Rennes, Inserm, CNRS, IRISA

Reproducible brain imaging



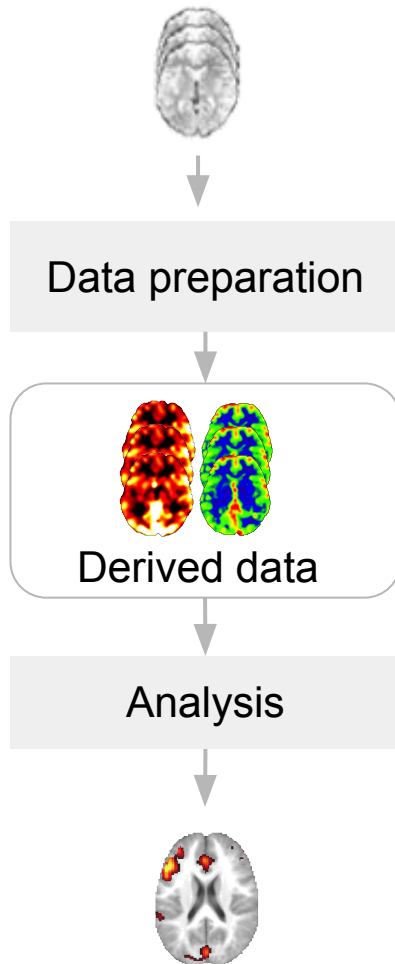
Reproducible brain imaging





On our way to fix
irreproducibility...

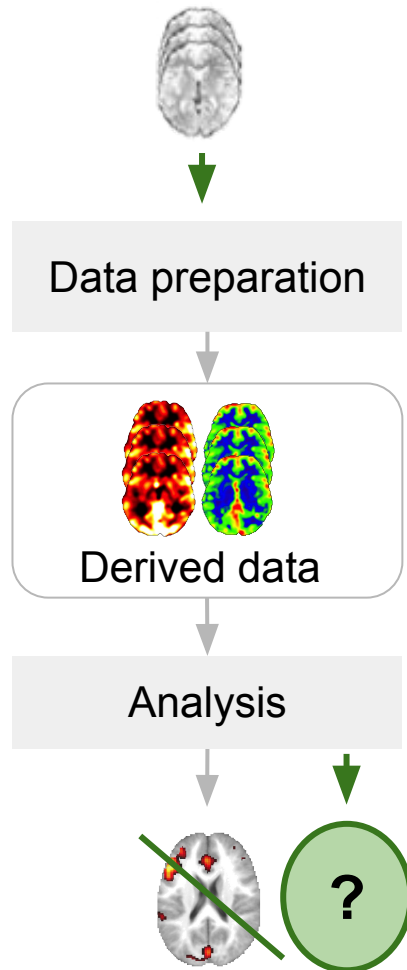
Towards **reproducible** neuroimaging



Towards **reproducible** neuroimaging

Irreproducible with...

Same Data, Same Methods

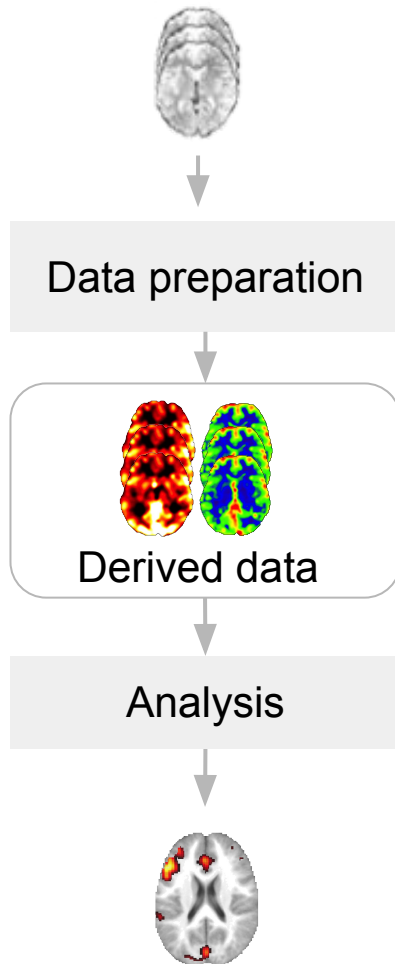


- Sharing code
- Sharing environments:
Docker, guix

Towards **reproducible** neuroimaging

Irreproducible with...

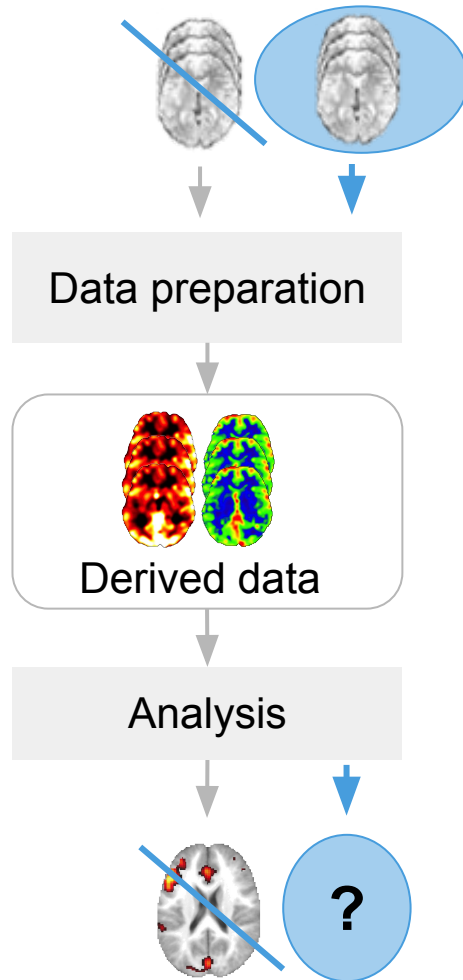
Different Data



Towards **reproducible** neuroimaging

Irreproducible with...

Different Data



Towards **reproducible** neuroimaging

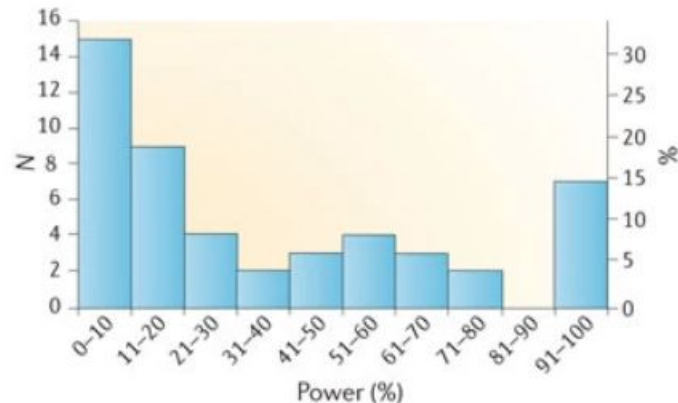
Irreproducible with...

Different Data

Lack of power

Power of neuroscience studies

Power = Prob. to correctly find a significant effect when a the alternative hypothesis is true.



[Button et. al, Nat Rev Neurosci 2013]

We need bigger datasets

Towards **reproducible** neuroimaging

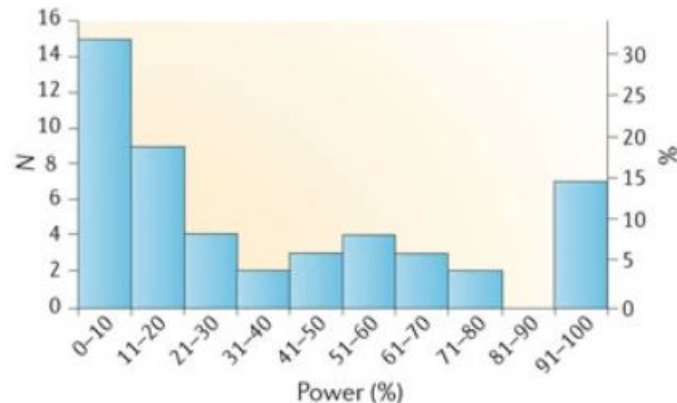
Irreproducible with...

Different Data

Lack of power

Power of neuroscience studies

Power = Prob. to correctly find a significant effect when a the alternative hypothesis is true.



[Button et. al, Nat Rev Neurosci 2013]

Generalizability and fairness

X-ray: Lung opacity detection

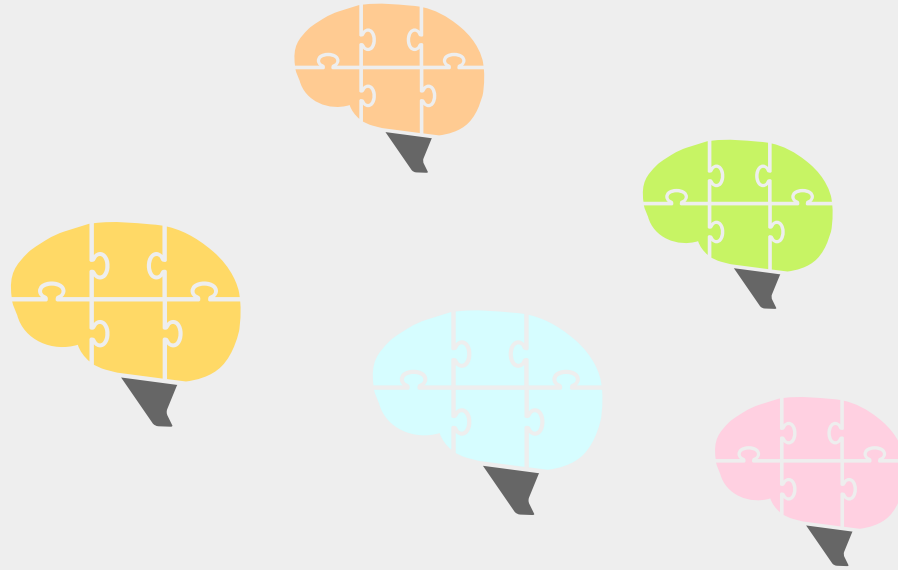
Model trained on male images, tested on female images



[Larrazabal et. al, PNAS 2020]

We need bigger datasets

We need representative and diverse datasets

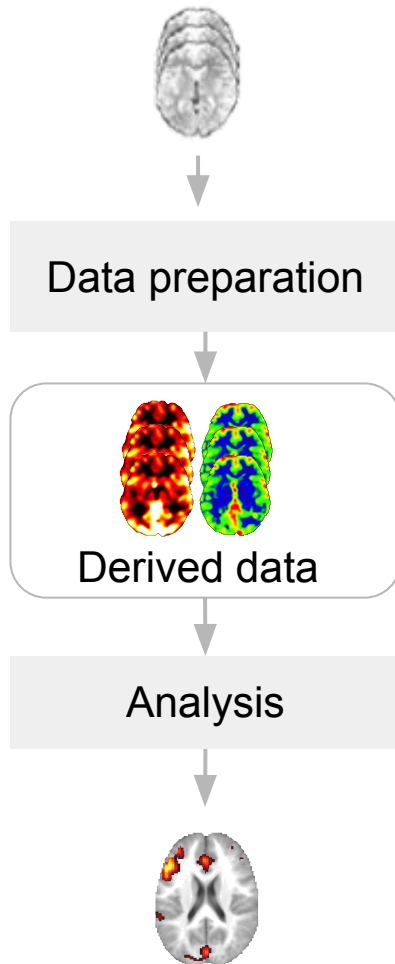


Different pipelines...
different results?

Towards **reproducible** neuroimaging

Irreproducible with...

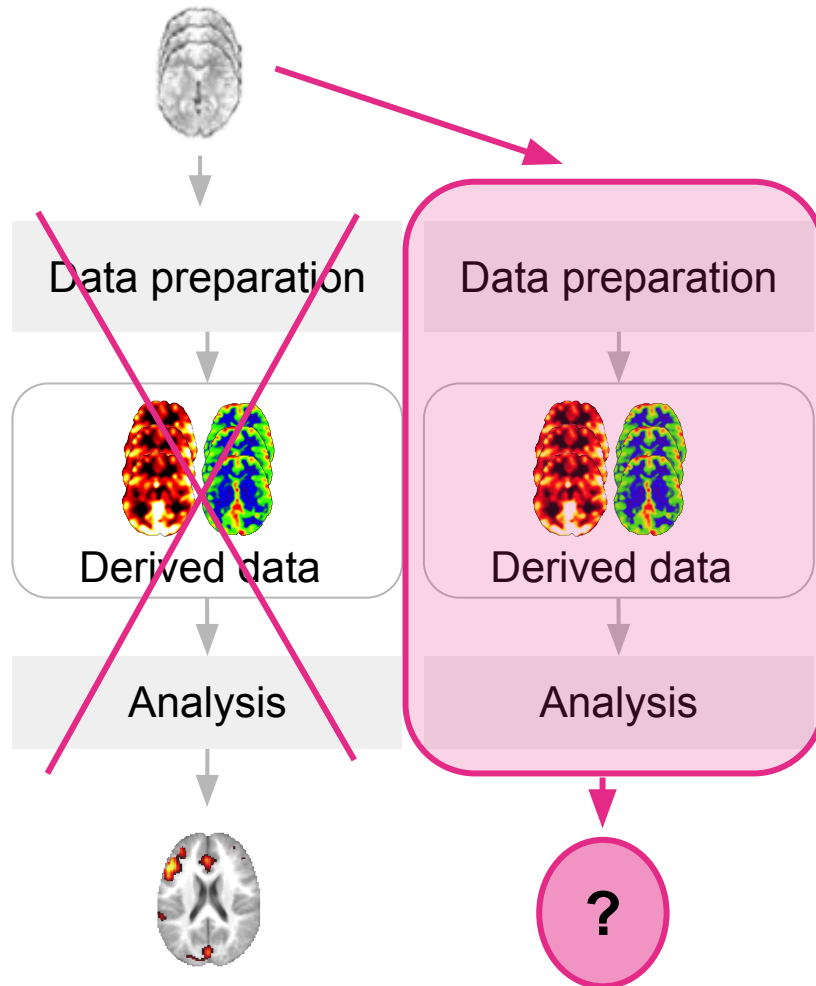
Different Analysis



Towards **reproducible** neuroimaging

Irreproducible with...

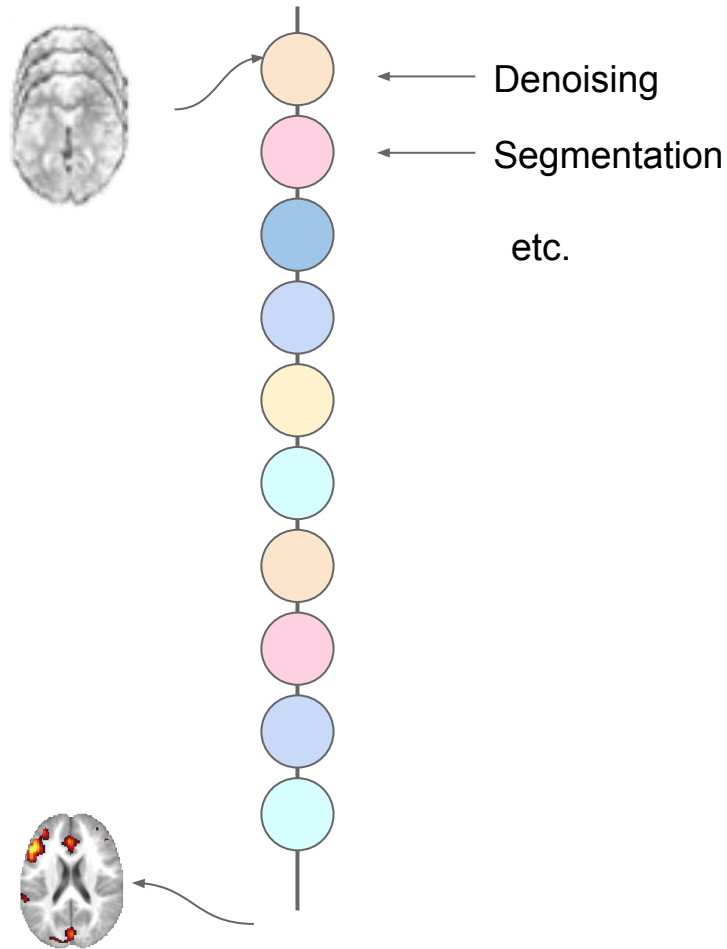
Different Analysis



Towards **reproducible** neuroimaging

Irreproducible with...

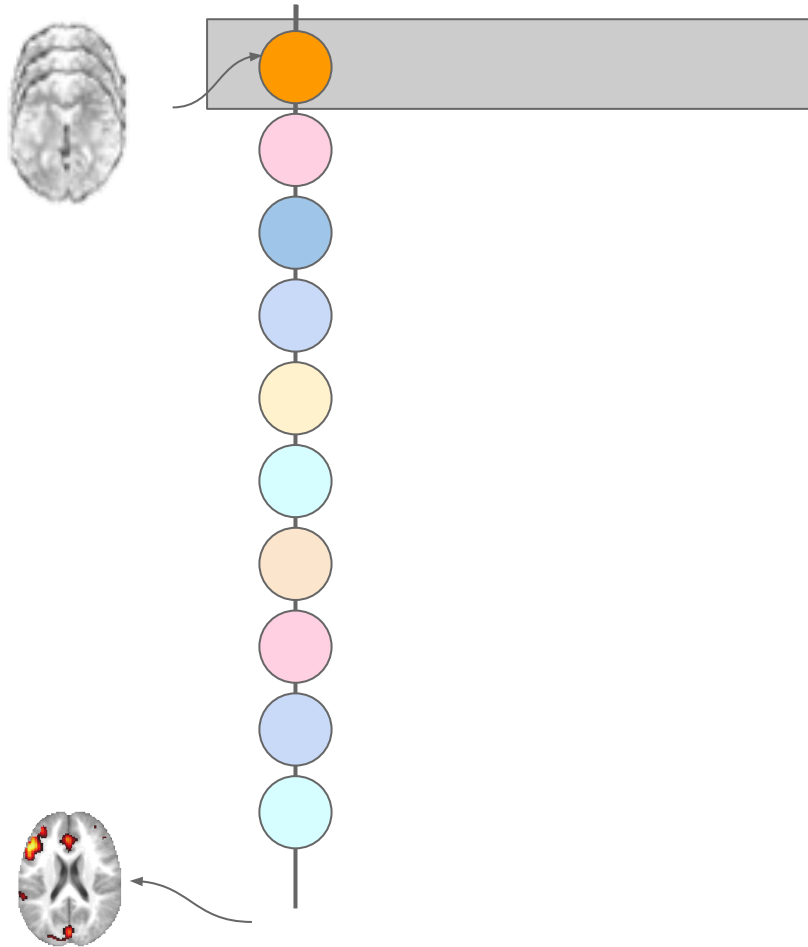
Different Analysis



Towards **reproducible** neuroimaging

Irreproducible with...

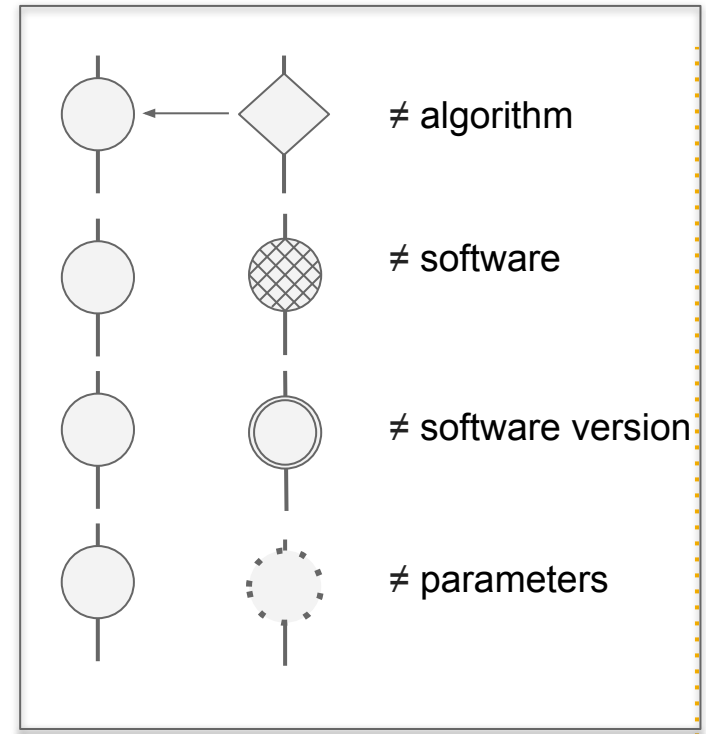
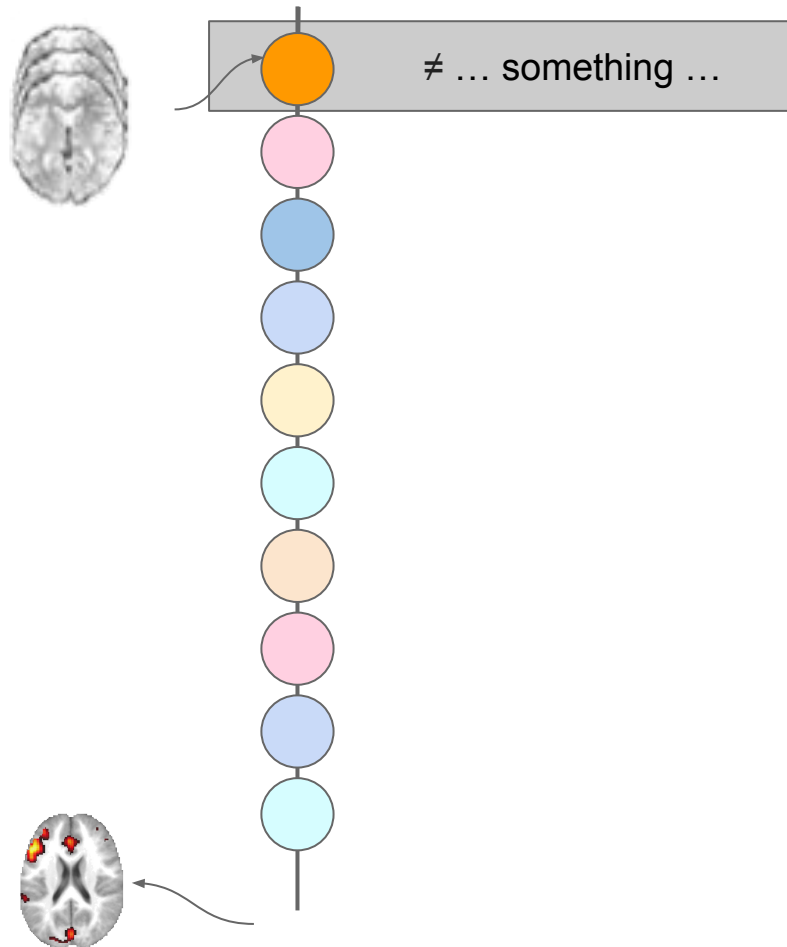
Different Analysis



Towards **reproducible** neuroimaging

Irreproducible with...

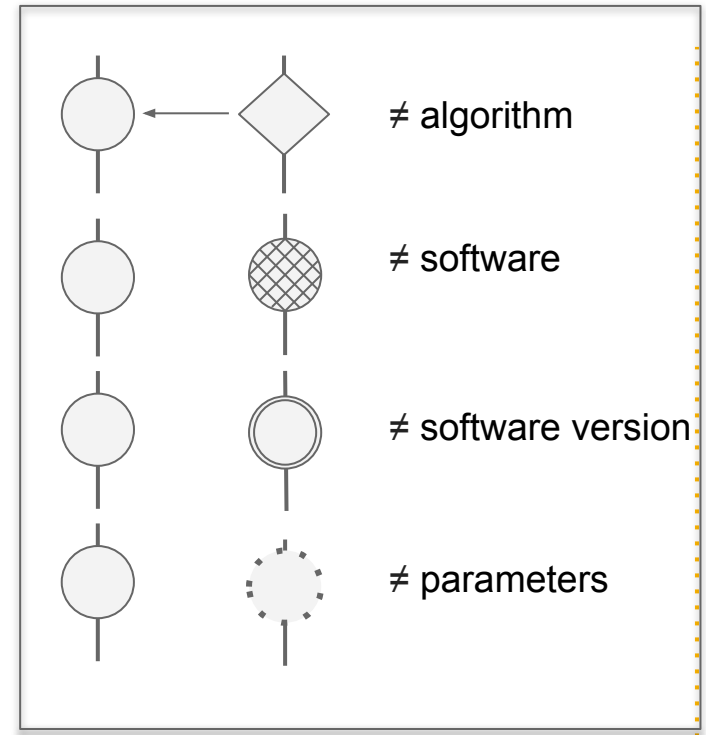
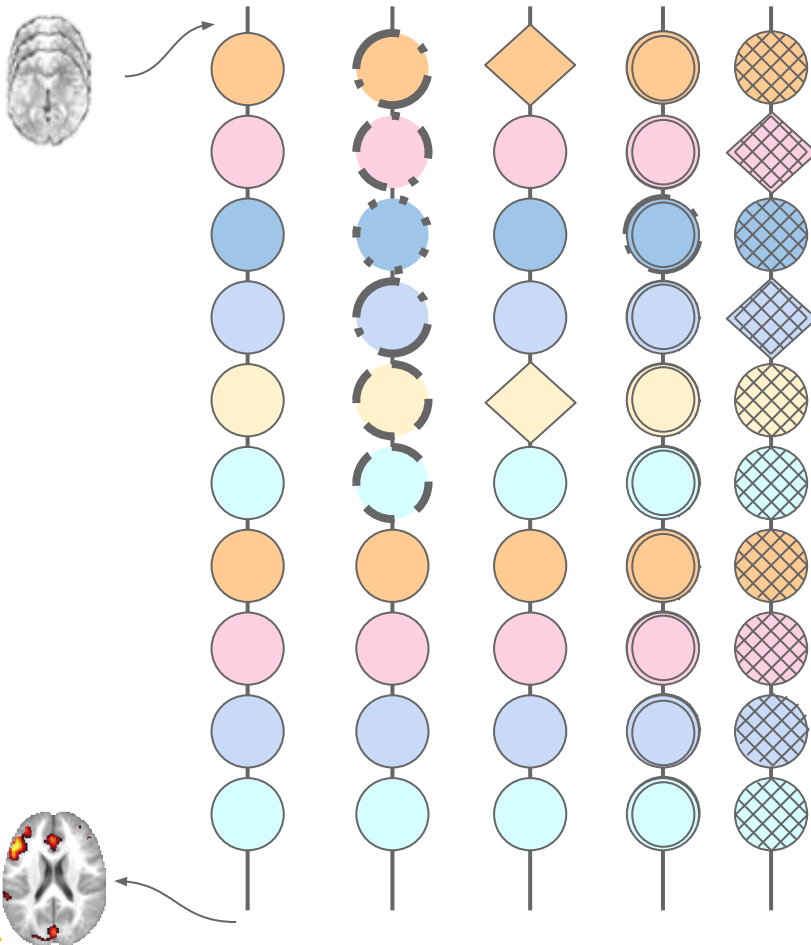
Different Analysis



Towards **reproducible** neuroimaging

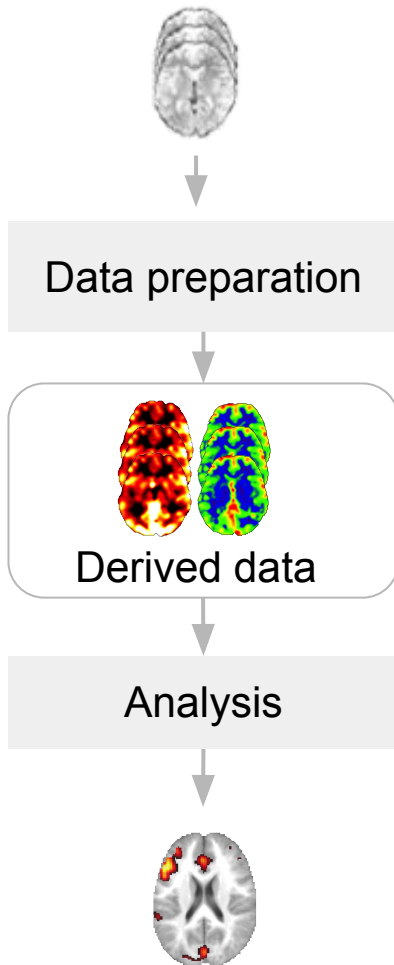
Irreproducible with...

Different Analysis

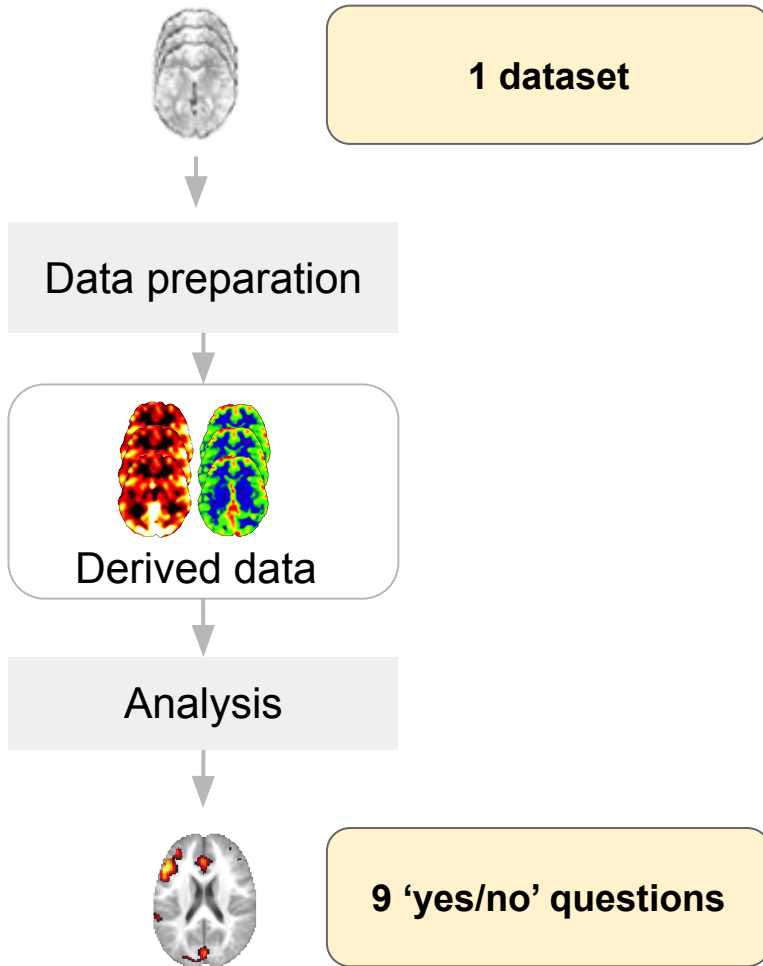


A family of acceptable pipelines
over 10^{30} combinations...

Many-analyst project : NARPS

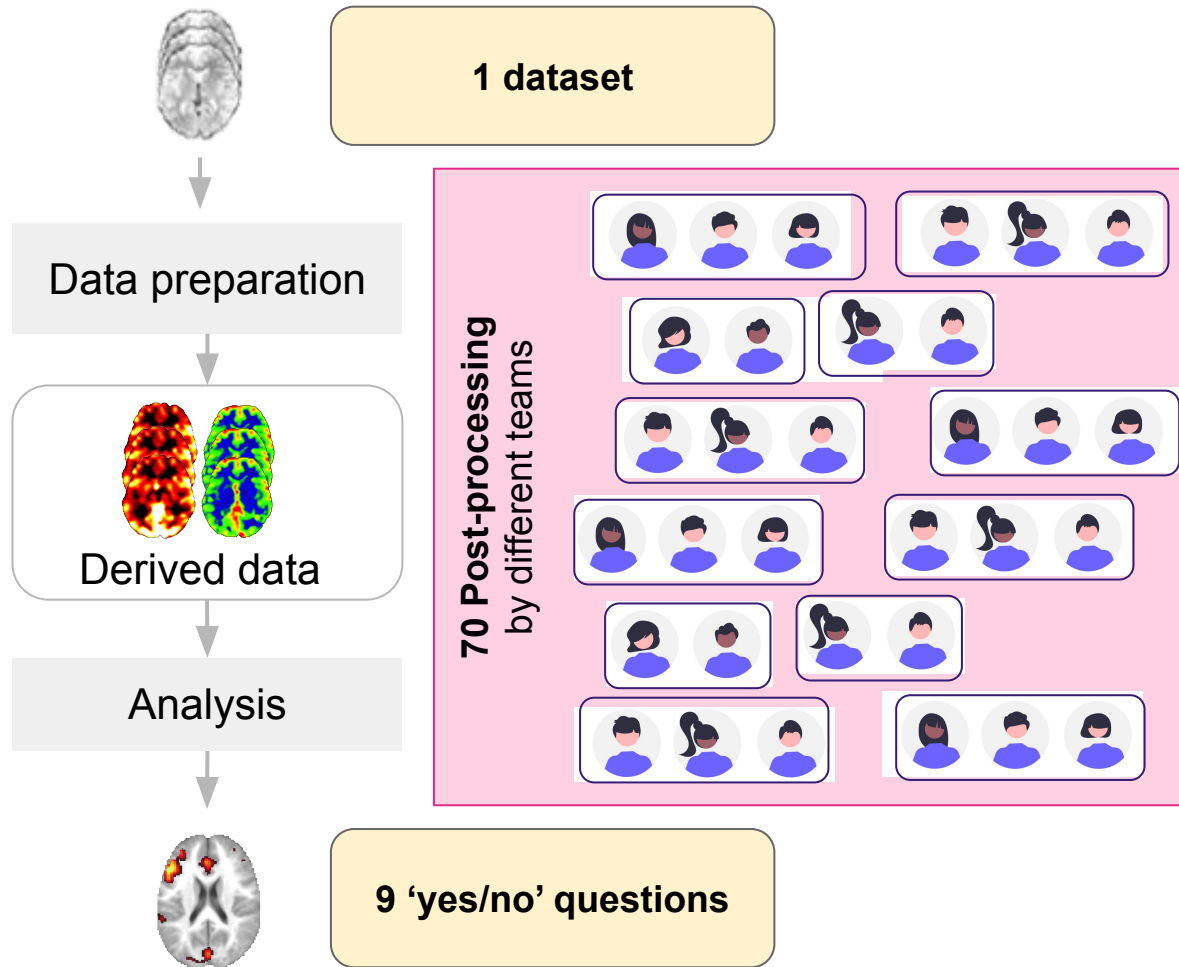


Many-analyst project : NARPS



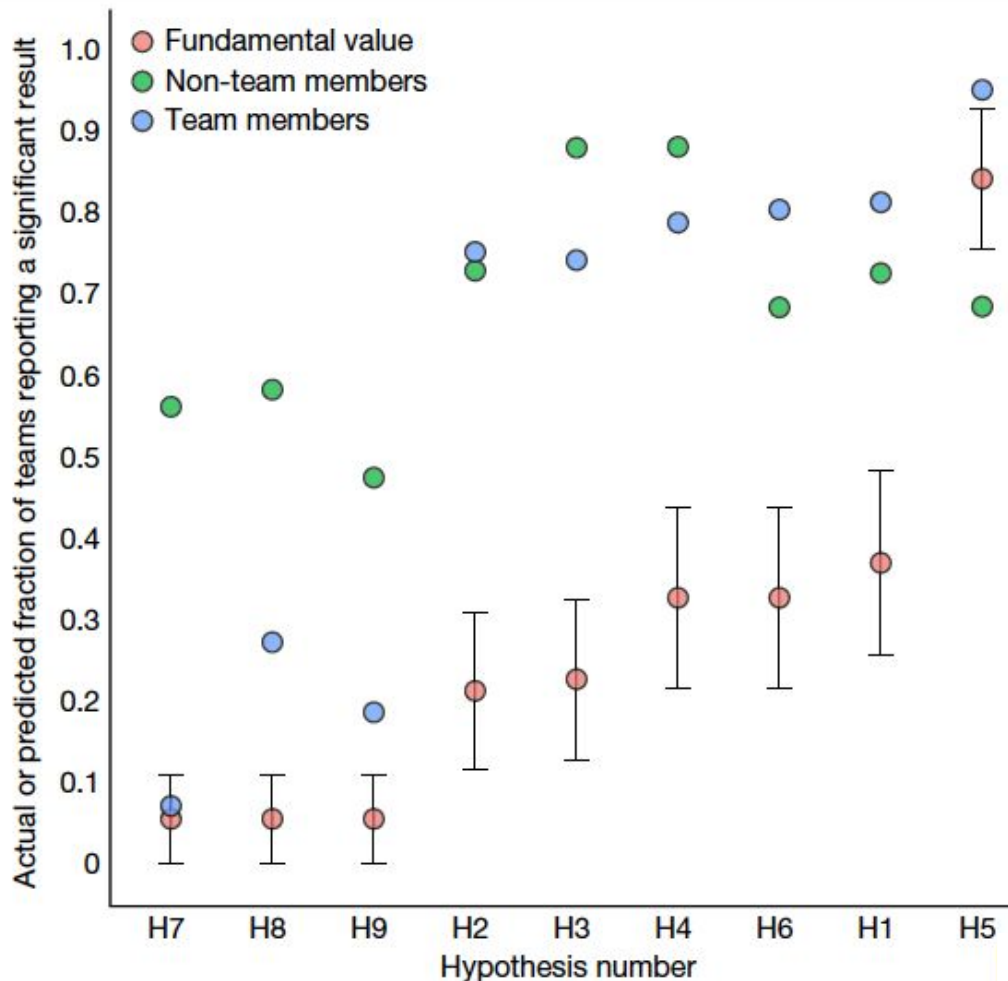
*e.g. Is there a parametric effect of gain:
Positive effect in ventromedial PFC -
for the equal indifference group?*

Many-analyst project : NARPS



Many-analyst project : NARPS

Results over the **70 post-processing pipelines**:

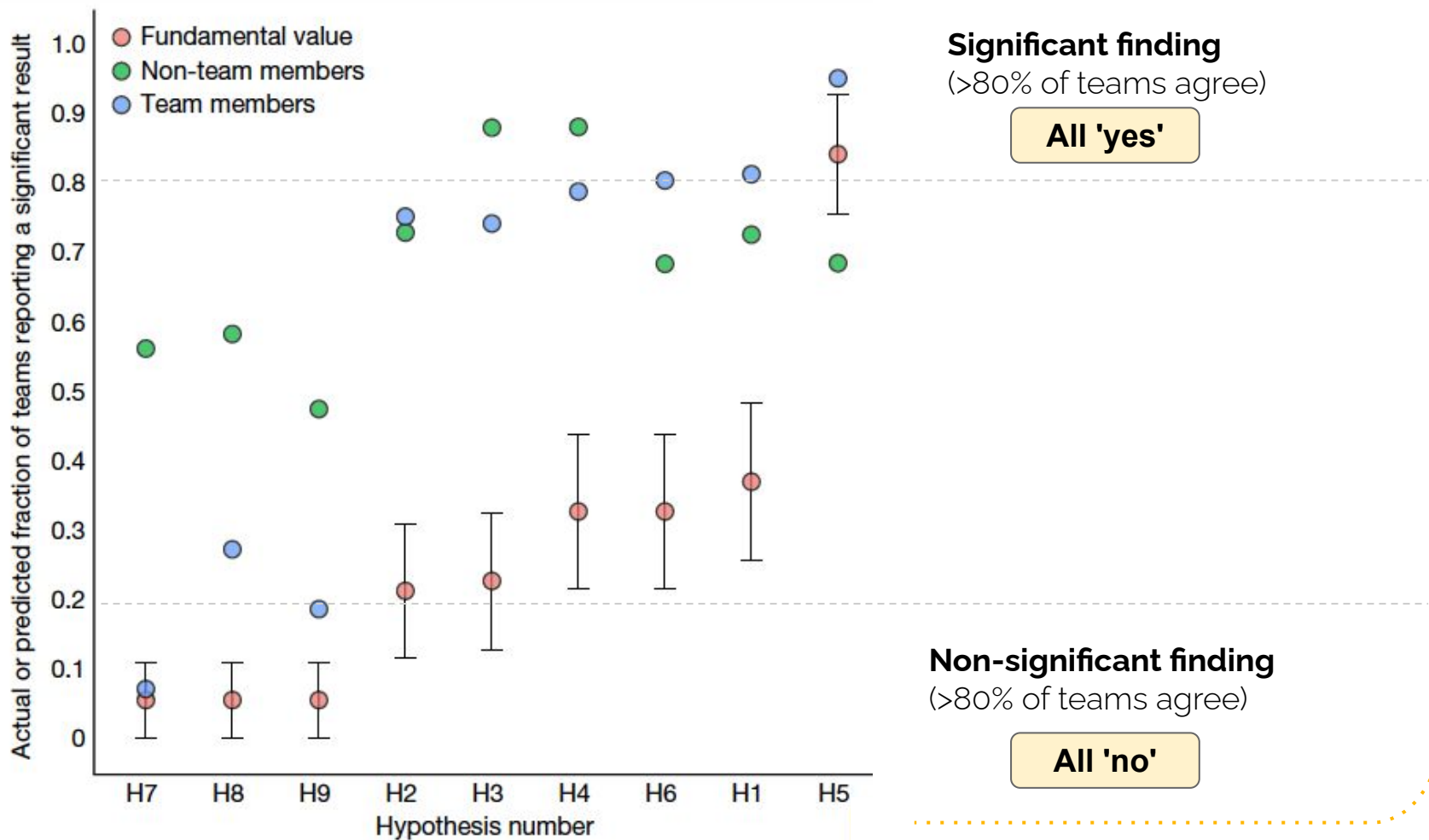


9 'yes/no' questions

[Botvinik-Nezer et. al, Nature 2020]

Many-analyst project : NARPS

Results over the **70 post-processing pipelines**:

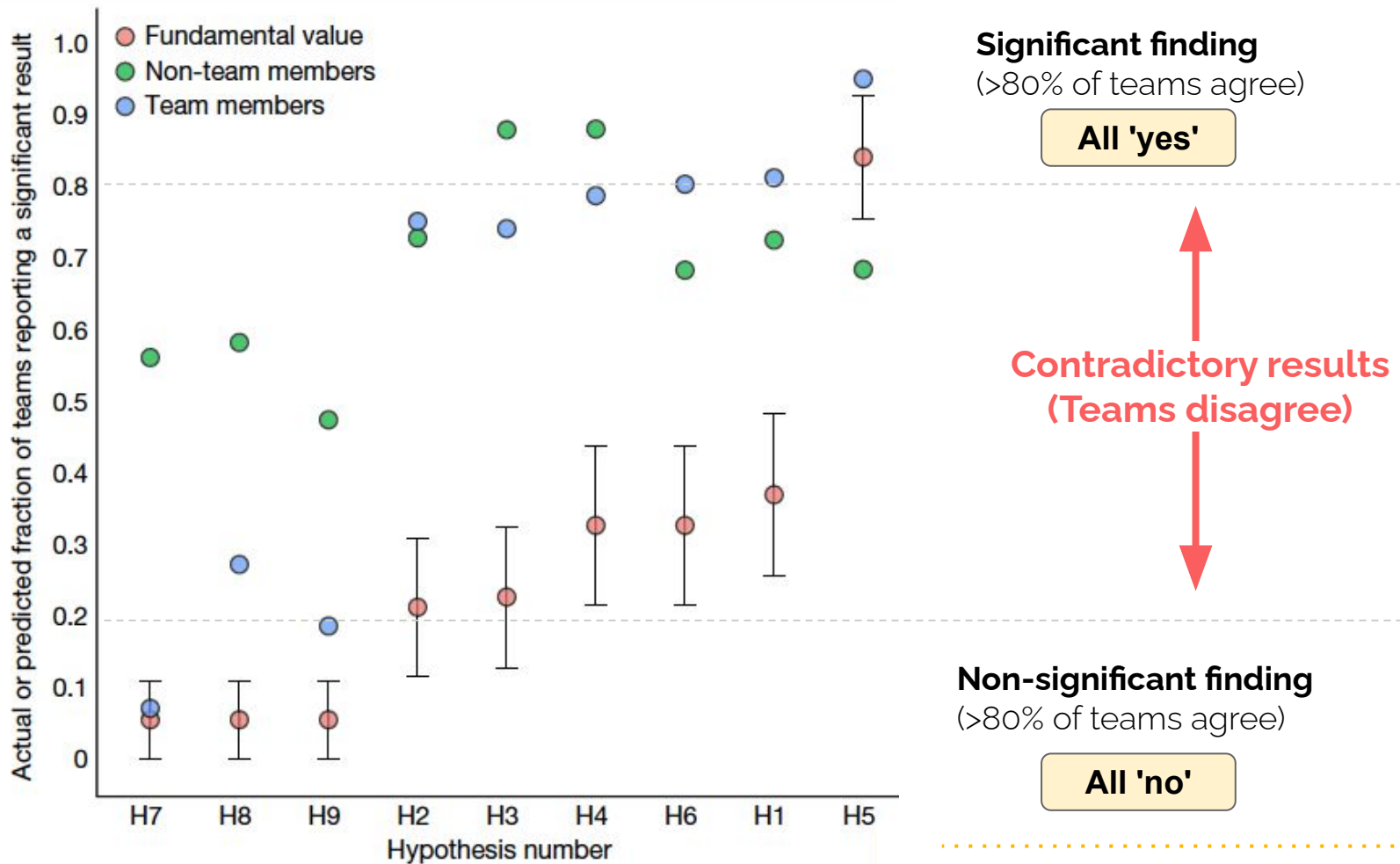


9 'yes/no' questions

[Botvinik-Nezer et. al, Nature 2020]

Many-analyst project : NARPS

Results over the **70** post-processing pipelines:



9 'yes/no' questions

[Botvinik-Nezer et. al, Nature 2020]

Different levels of variability in post-processing pipelines

- Different end-to-end pipelines [Carp 2012]
- Different software version [Gronenschild 2012]
- Different operating system [Glatard 2015]

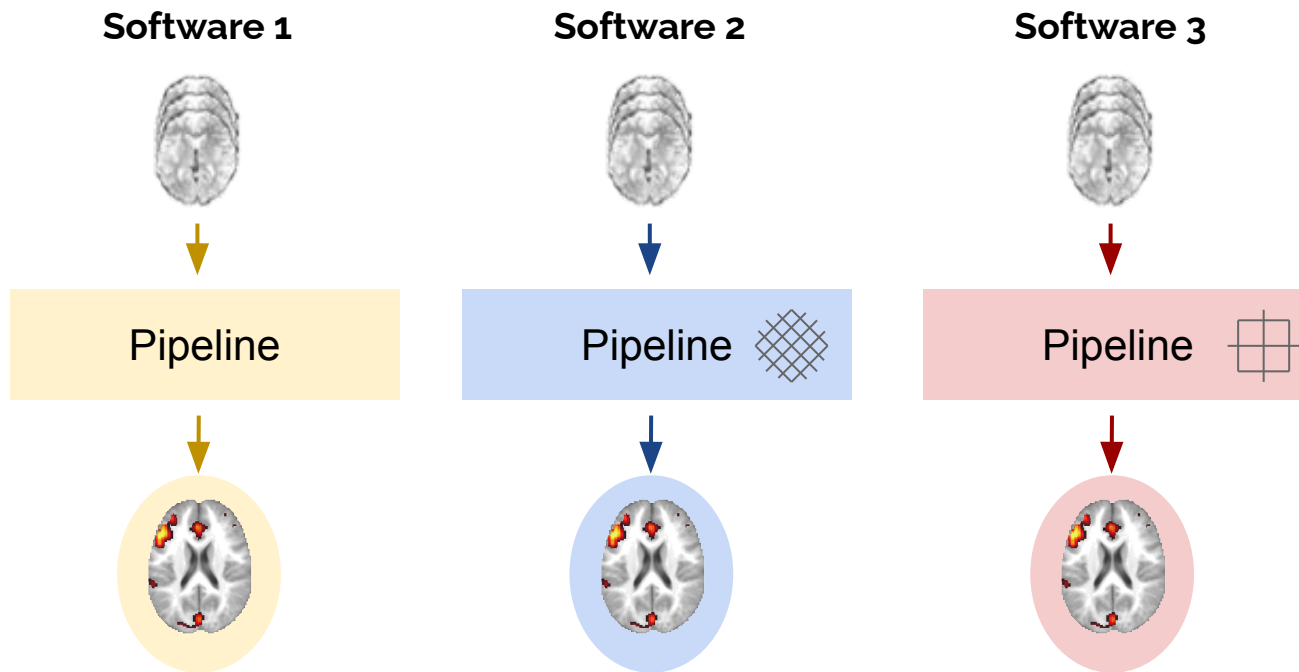
- Different **post-processing software packages** [Bowring 2019, 2021]

Variability across software

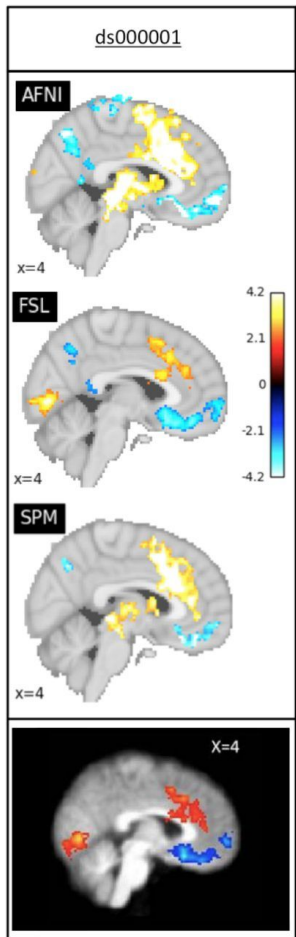
- Reproduced 3 published functional MRI studies
- Using 3 different software



Alex Bowring Tom Nichols



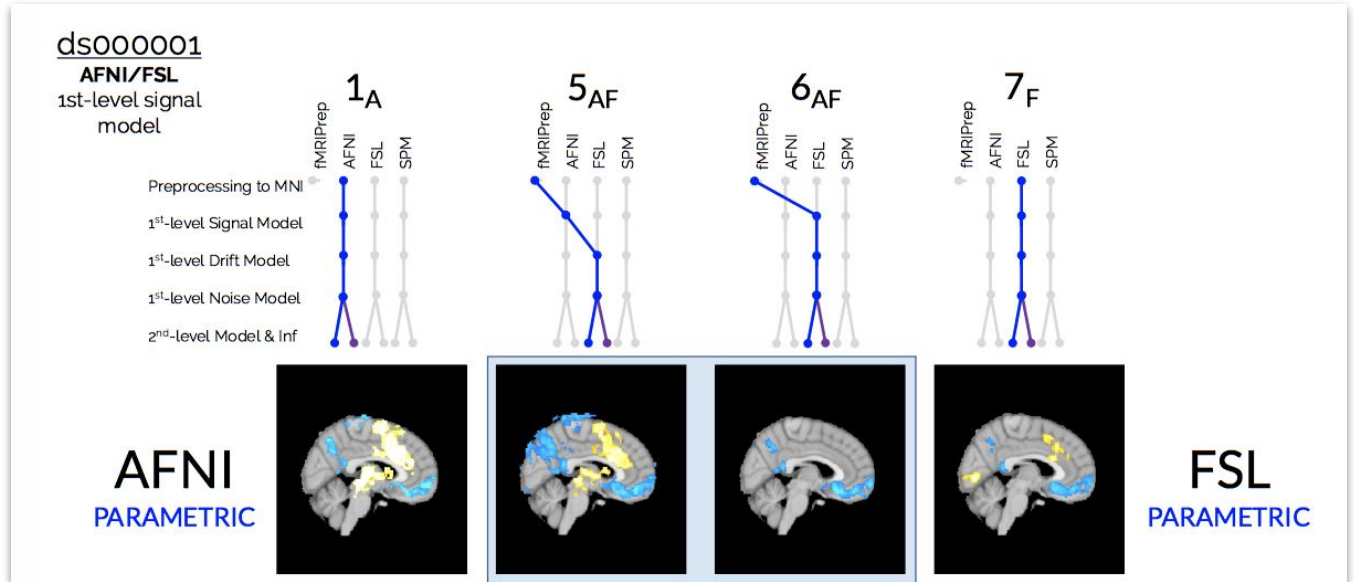
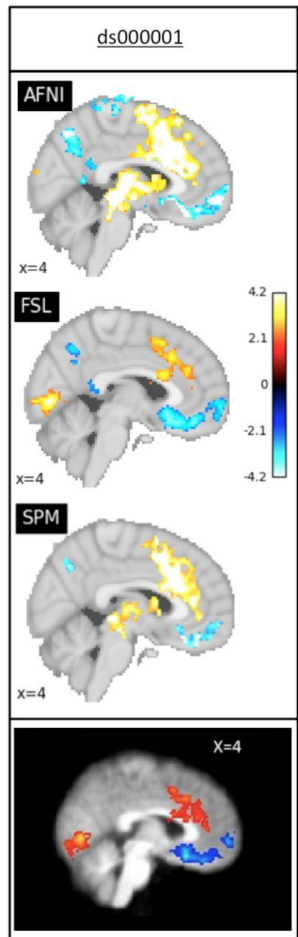
Variability across software



Thresholded maps in AFNI,
FSL, SPM + original

[Bowring et. al, HBM 2019]

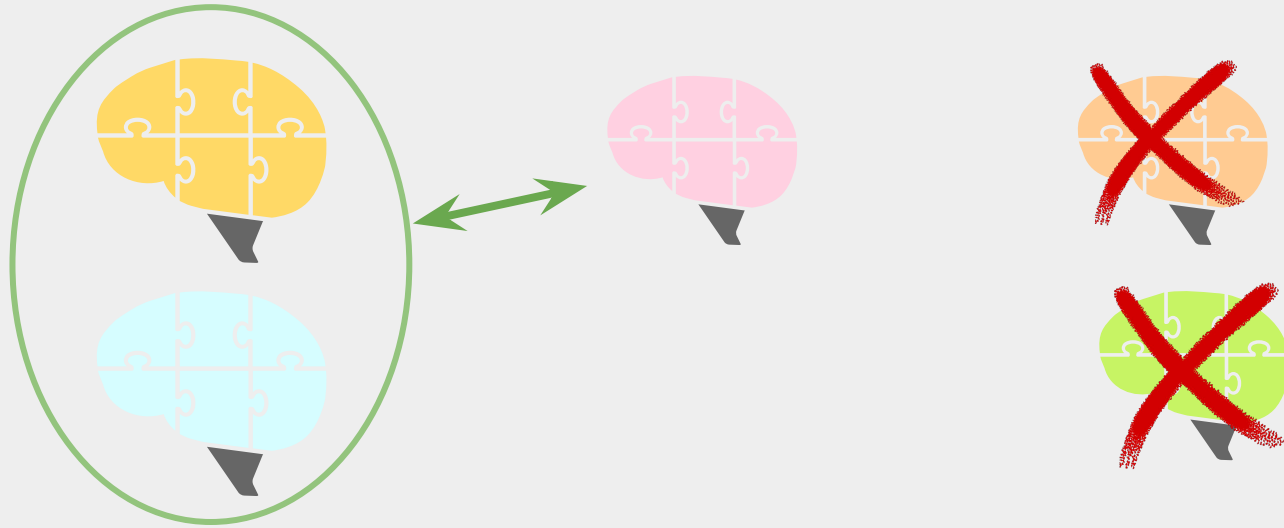
Variability across software



[Bowring et. al, HBM 2021]

Thresholded maps in AFNI,
FSL, SPM + original

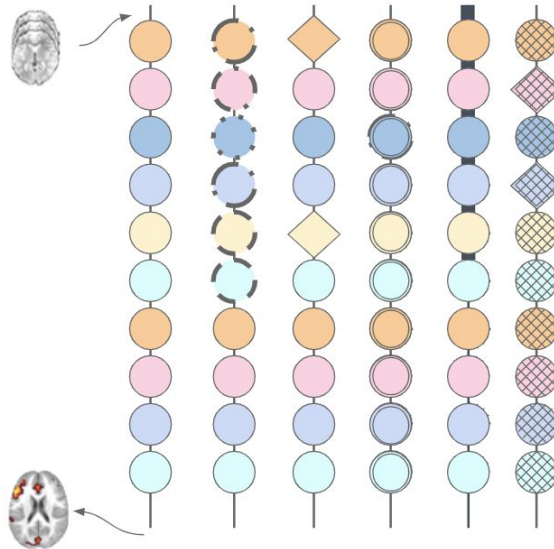
[Bowring et. al, HBM 2019]



Understanding
analytical variability

Understanding analytical variability...

?



Understanding analytical variability...

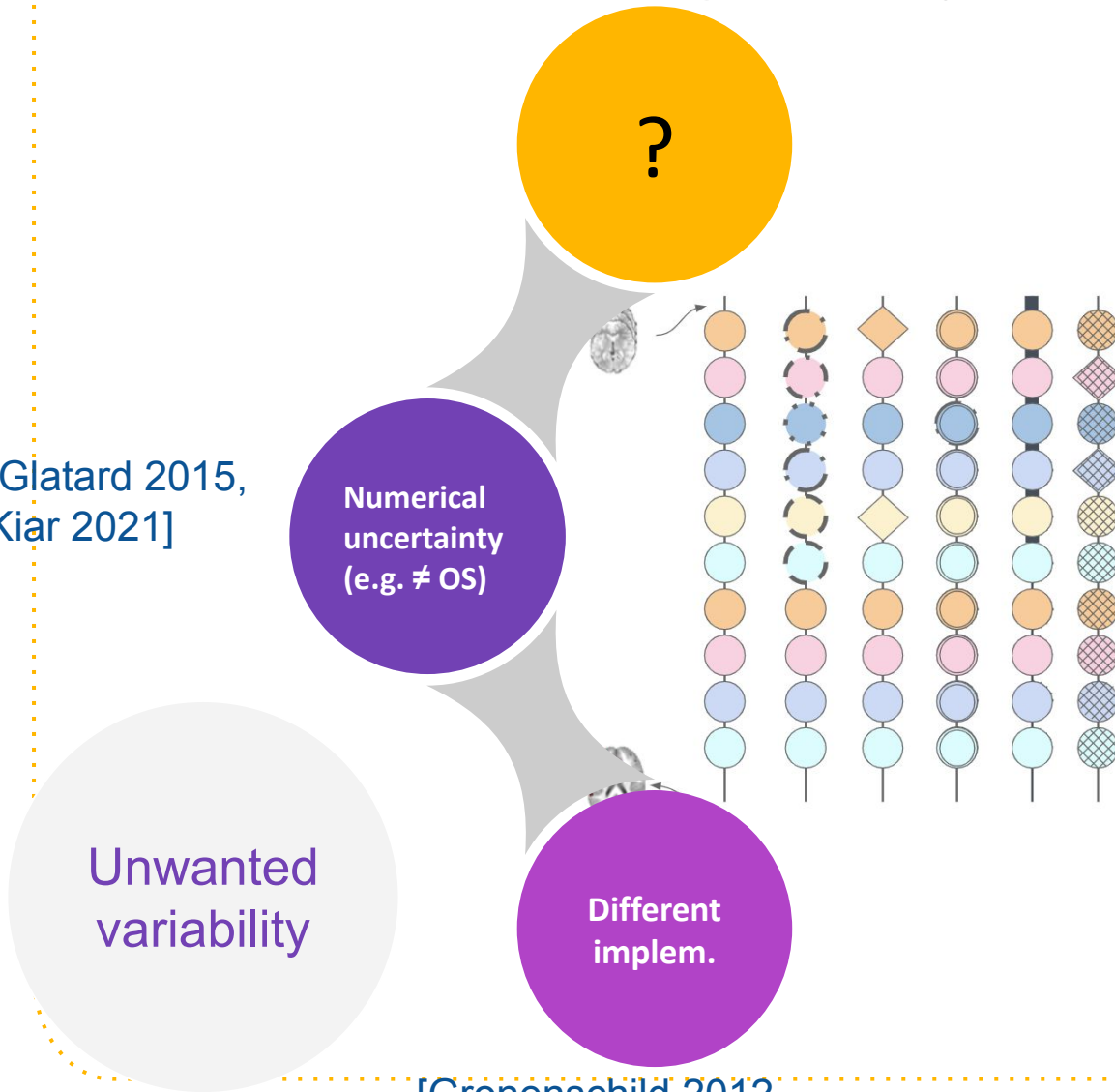
[Glatard 2015,
Kiar 2021]

Numerical
uncertainty
(e.g. \neq OS)

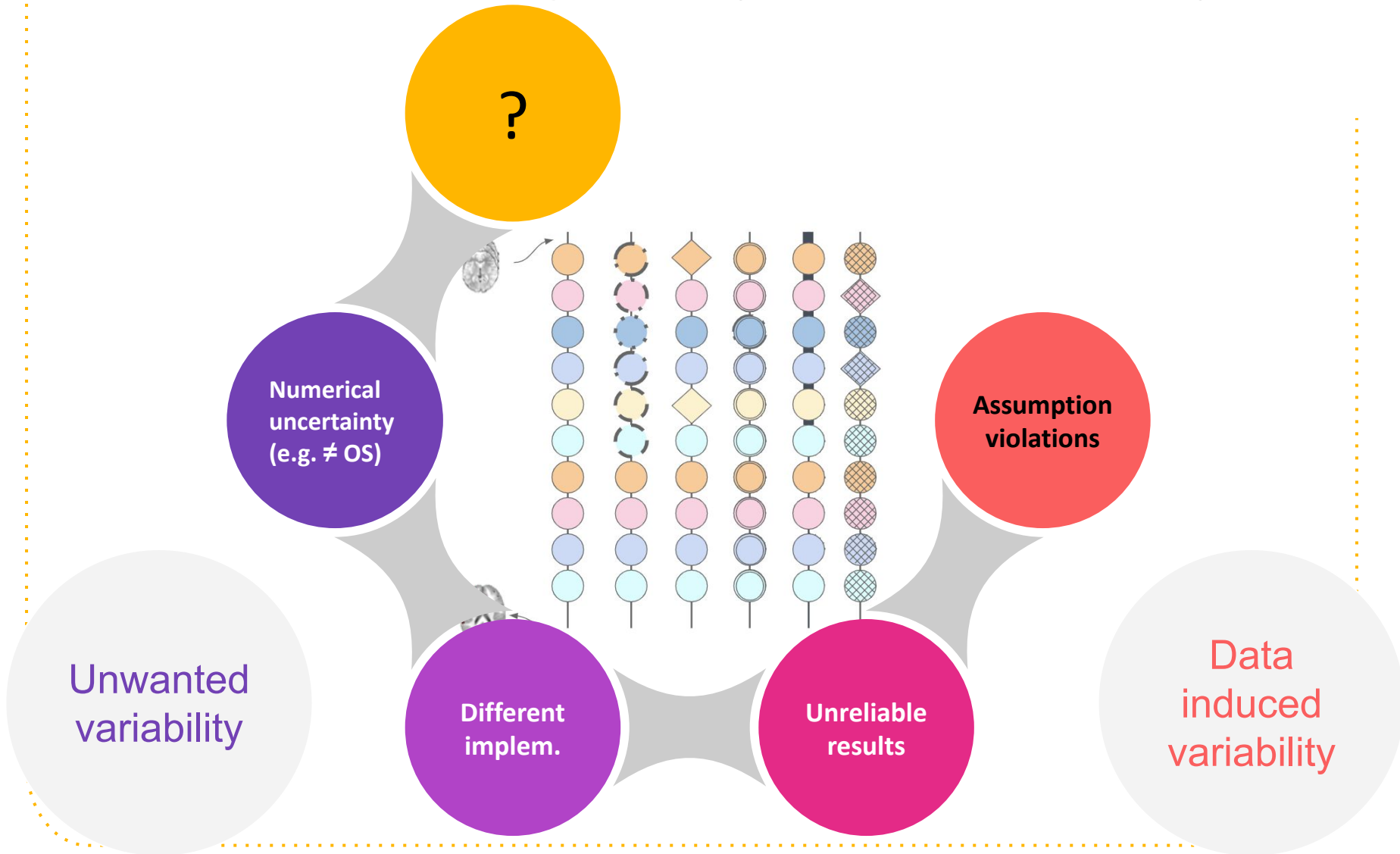
Unwanted
variability

Different
implem.

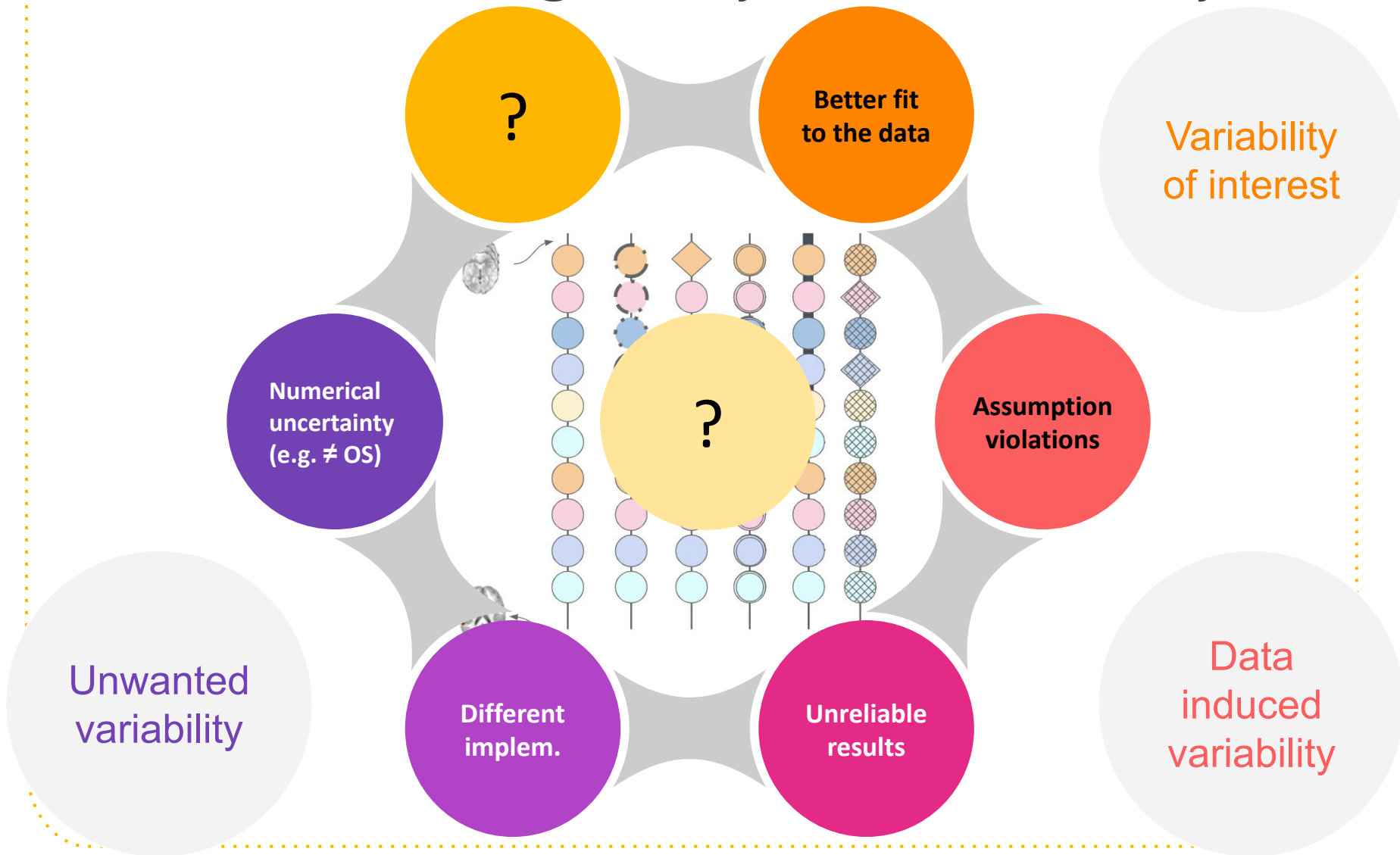
[Gronenschild 2012,
Bowring 2019]

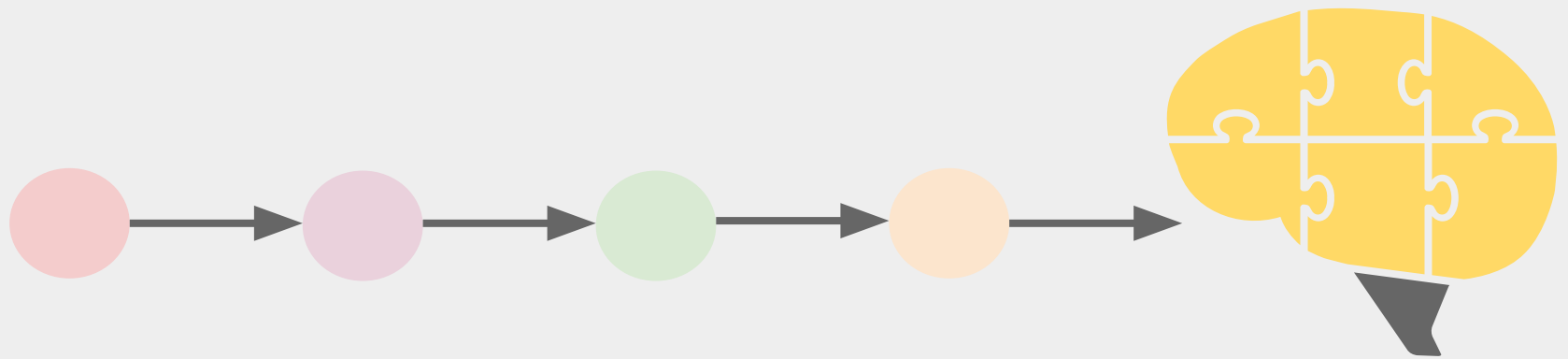


Understanding analytical variability...



Understanding analytical variability...





**Describing the
neuroimaging
pipeline space**



Boris Clenet

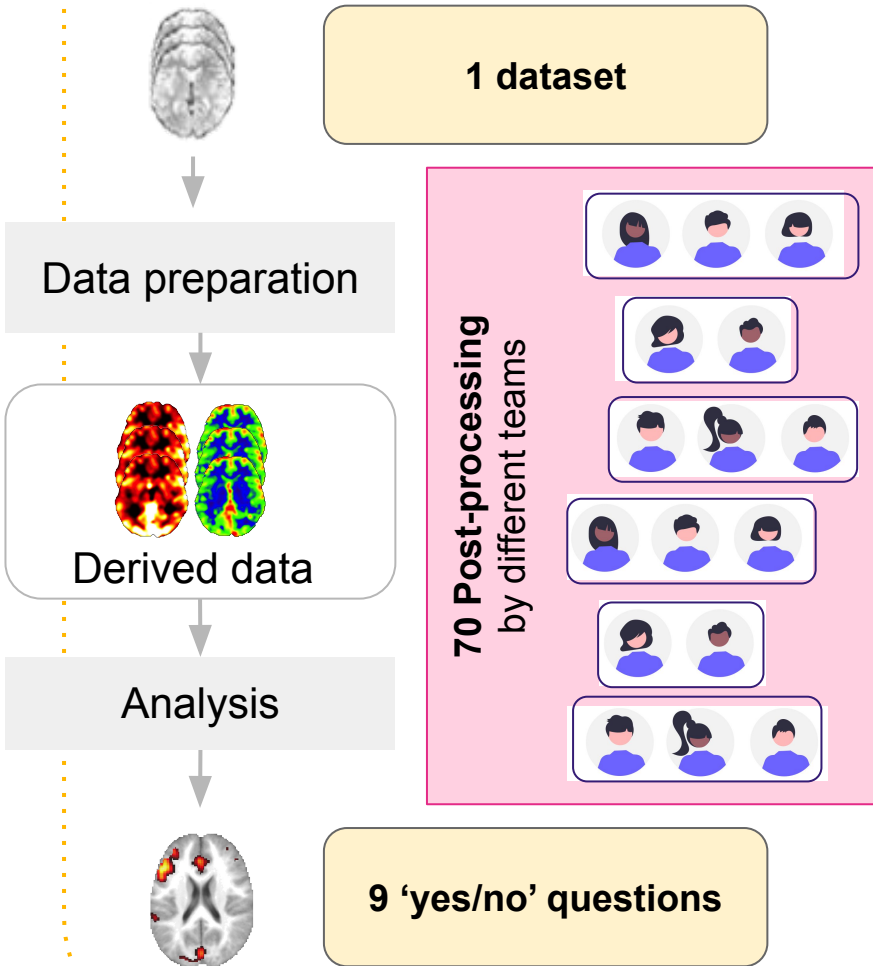


Elodie Germani



JérémY Lefort-Besnard

NARPS open pipelines





Boris Clenet

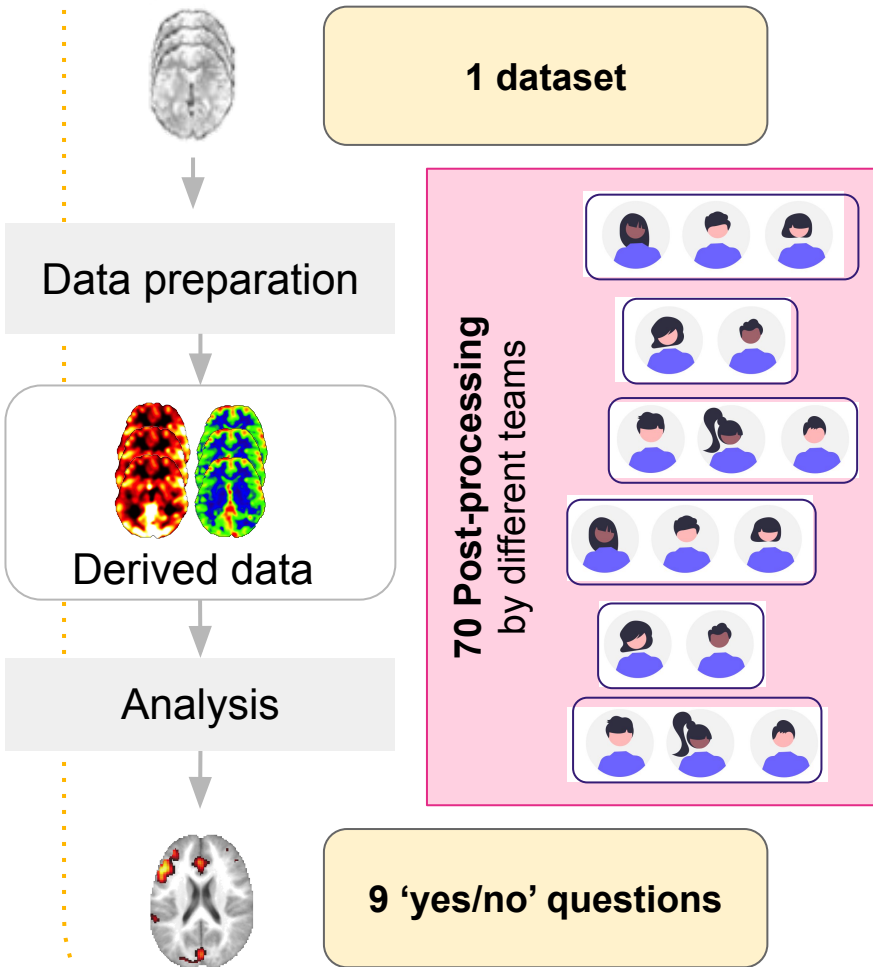


Elodie Germani



Jérémie Lefort-Besnard

NARPS open pipelines



NARPS open pipelines

- (Re)-coding the 70 pipelines
- Collaborative effort



Project NARPS open pipelines

Join us at the Brainhack Global!

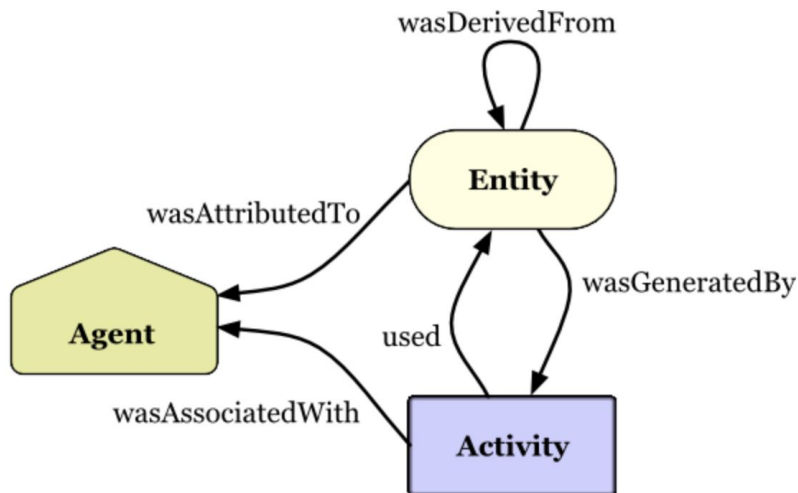


https://github.com/Inria-Empenn/narps_open_pipelines/

Provenance

“**Provenance** describe how an artifact or set of artifacts was produced” [Wikipedia](#)

“**Provenance** is information about entities, activities, and people involved in producing a piece of data [...]” [W3C PROV.](#)



W3C PROV model

<https://www.w3.org/TR/2013/NOTE-prov-primer-20130430/>

Provenance Use Case 1 **An unambiguous description of the methods**



Provenance Use Case 1 **An unambiguous description of the methods**



Which release of SPM was used?

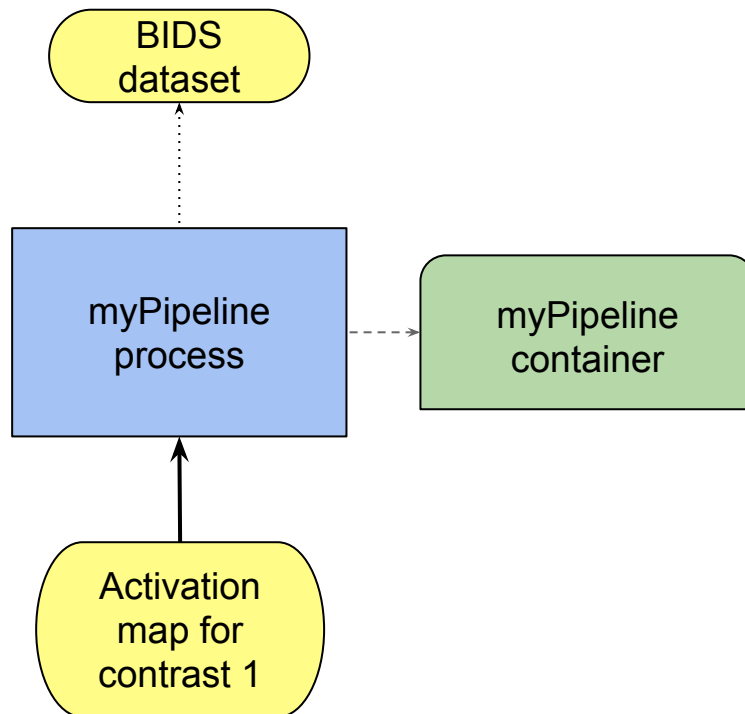
Which registration algorithm was selected?

How was the brain mask computed?

Provenance Use Case 1 **An unambiguous
description of the methods**

Can we
reproduce
exactly the same
results
(same data, same
methods)?

Provenance Use Case 1 **An unambiguous description of the methods**



Can we
**reproduce
exactly the same
results**
(same data, same
methods)?

Provenance Use Case 2 **Dealing with
unwanted variability**

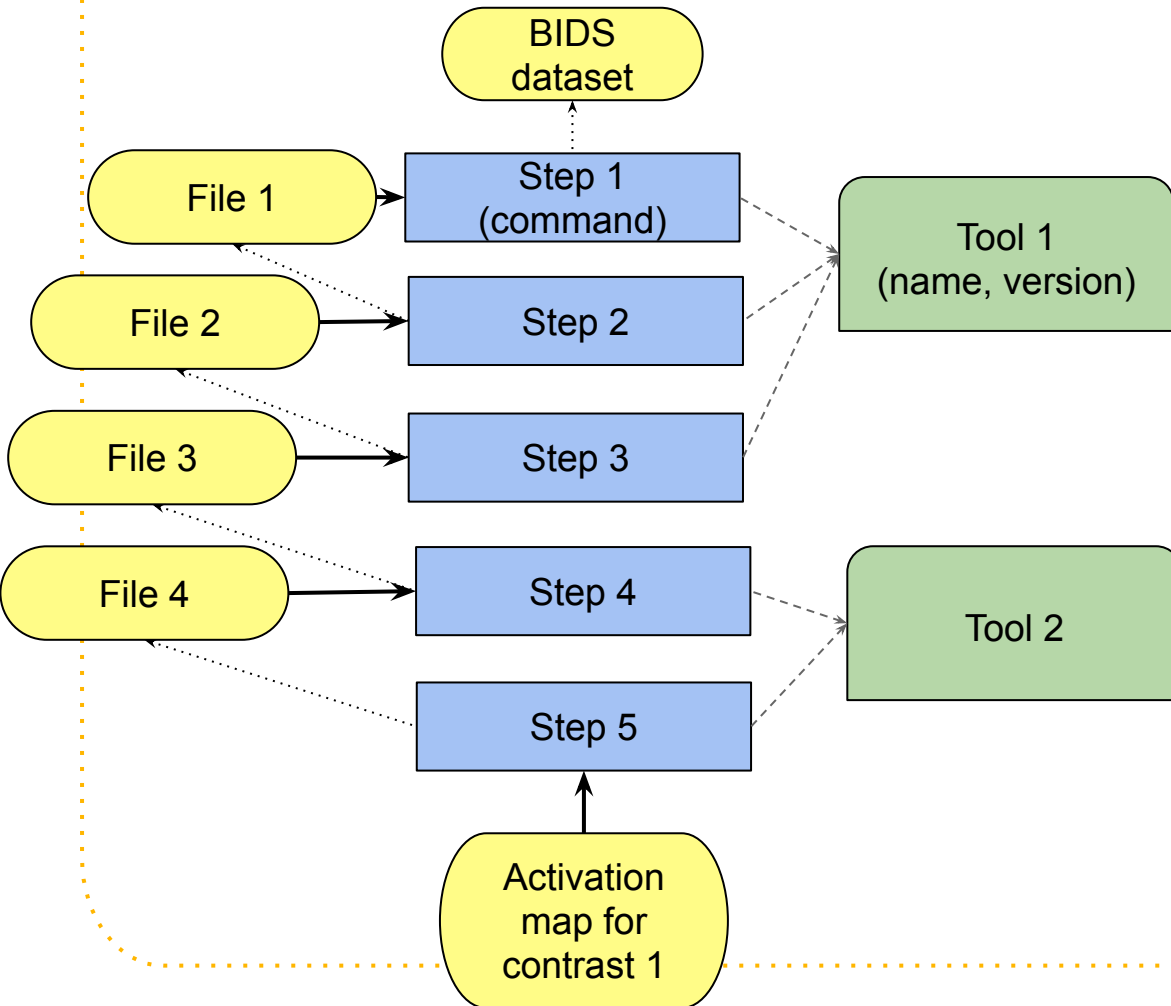
A previously
widely-used
**parameter is no
longer considered
as best practice**

Provenance Use Case 2 **Dealing with unwanted variability**

A previously widely-used **parameter is no longer considered as best practice**

A bug was found in a given tool **version**, can we retrieve all corresponding results?

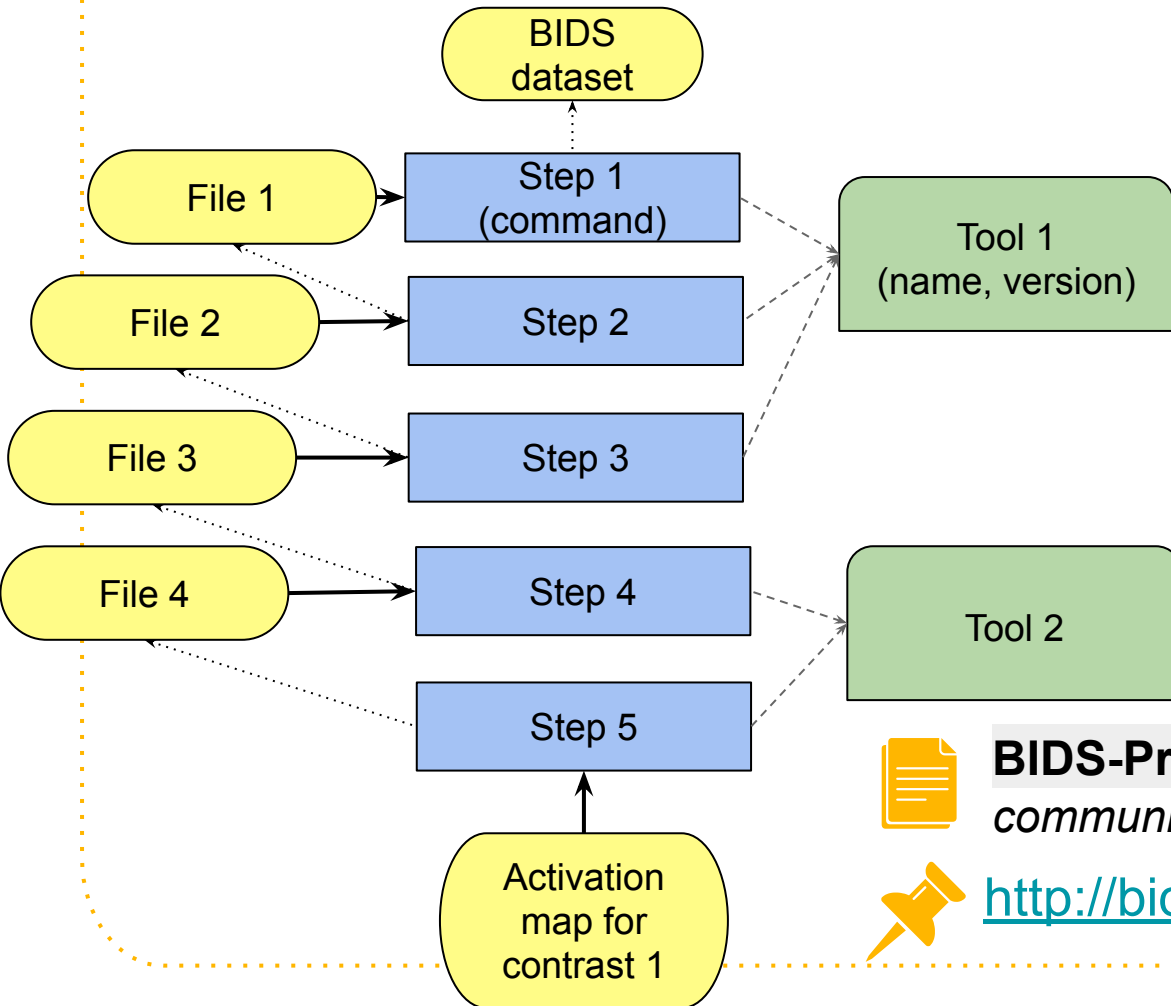
Provenance Use Case 2 **Dealing with unwanted variability**



A previously widely-used **parameter is no longer considered as best practice**

A bug was found in a given tool version, can we retrieve all corresponding results?

Provenance Use Case 3 **Exploring factors impacting analytical variability**



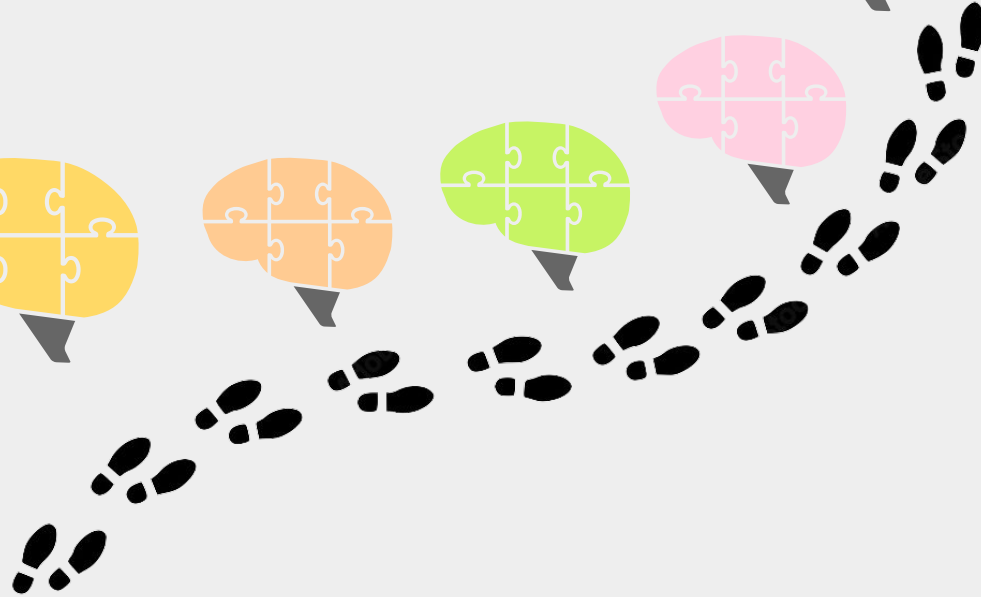
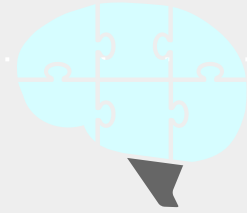
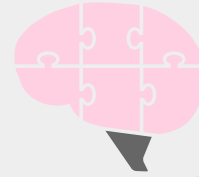
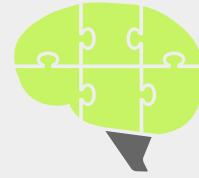
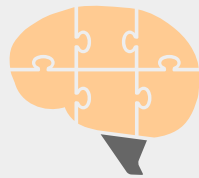
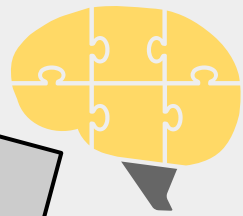
Are my conclusions dependant on the **size of the smoothing kernel** used at Step 5?



BIDS-Prov will soon be open for community feedback!



<http://bids.neuroimaging.io/bep028>



Which solutions **for** **practitioners?**

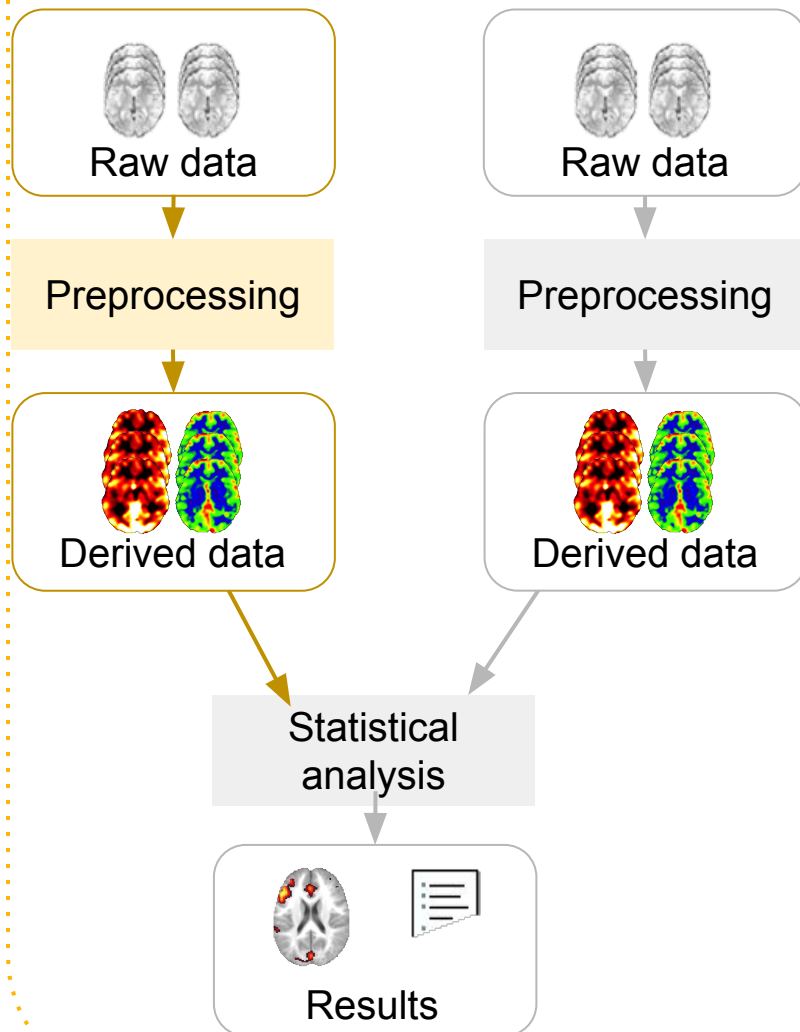


Xavier Rolland



Pierre Maurel

Data re-use



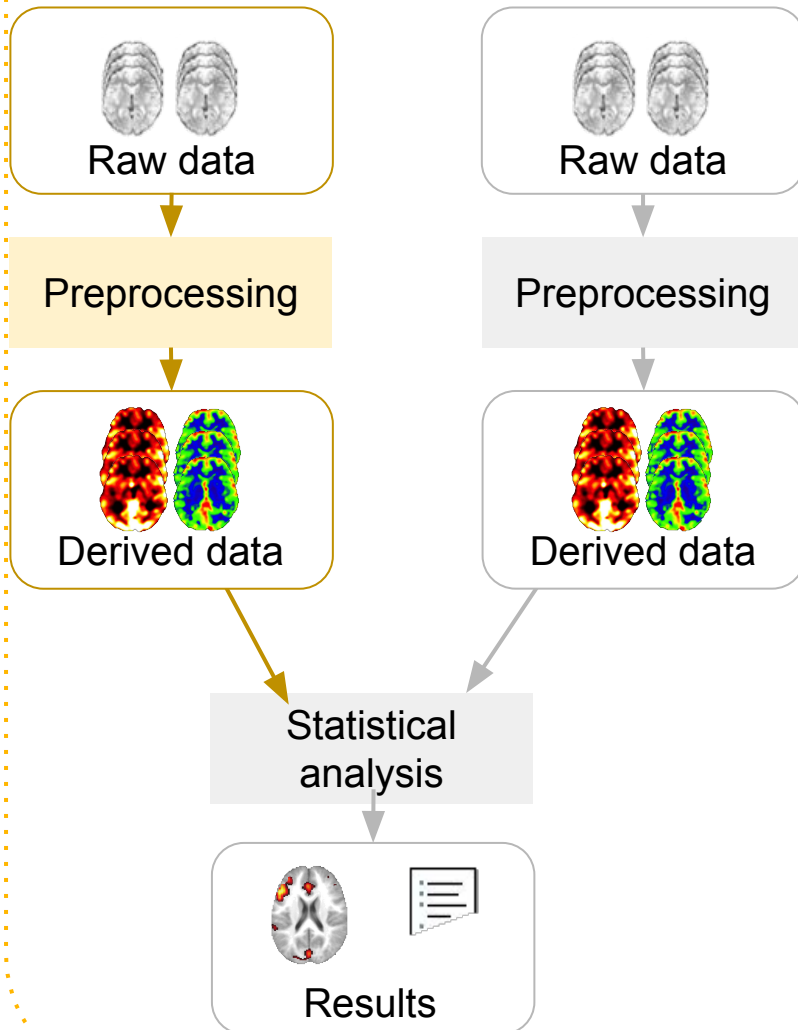
[Rolland et. al, ISBI 2022]

Data re-use

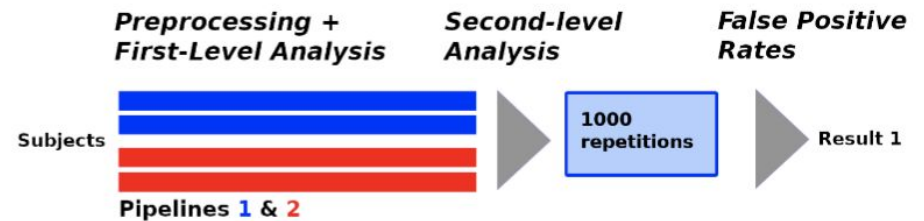


Xavier Rolland

Pierre Maurel



- Using HCP data under the null
- Comparisons of pipelines 2-by-2
- 3 variations
 - Smoothing: 5 or 8
 - HRF derivative: yes or no
 - Motion regressors: 0, 6, 24
- “Default” : 5mm, 24 regressors, no temp. deriv.

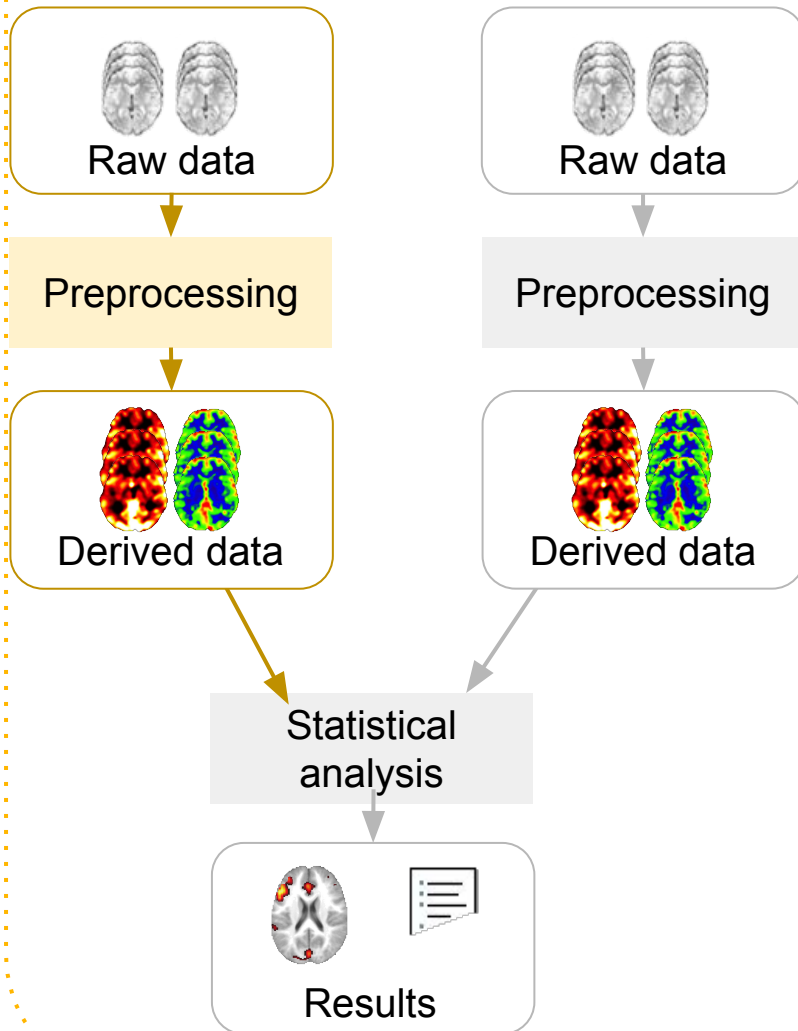


Data re-use



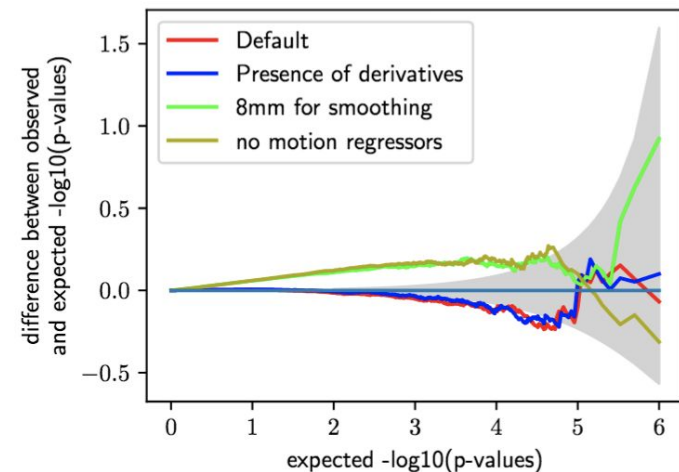
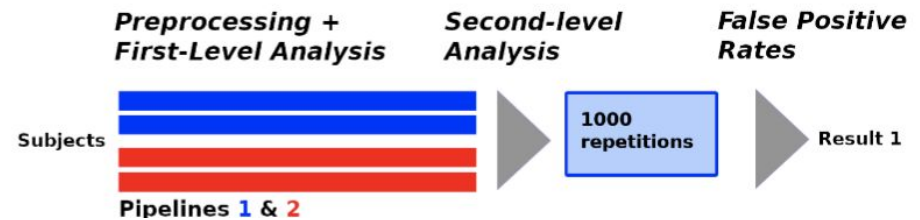
Xavier Rolland

Pierre Maurel



[Rolland et. al, ISBI 2022]

- Using HCP data under the null
- Comparisons of pipelines 2-by-2
- 3 variations
 - Smoothing: 5 or 8
 - HRF derivative: yes or no
 - Motion regressors: 0, 6, 24
- “Default” : 5mm, 24 regressors, no temp. deriv.



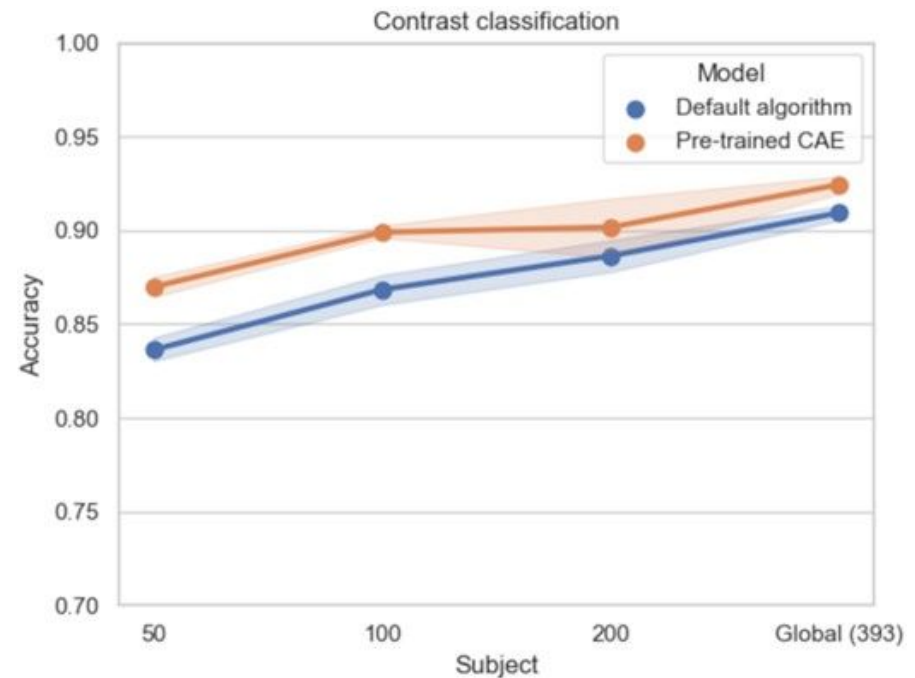
Building on large and more diverse samples



Elodie
Germani

Elisa
Fromont

- Latent representation of fMRI statistic maps based on NeuroVault
- Apply to a **decoding task**



[Germani et. al, Gigascience 2023]



<https://inria.hal.science/hal-03769993>

Same data meta-analysis



JérémY Lefort-Besnard
Tom Nichols

- A statistical approach to **combine the results of different pipelines** and obtain **a single (consensus) summary**



Poster *Systematic review and evaluation of meta-analysis methods for same data meta-analyses* JérémY Lefort-Besnard

Same data meta-analysis
Enter the multiverse...
Lefort-Besnard JeremY¹, Nichols Thomas^{2*}, Maumet Camille^{1*}
¹ Inria, Univ Rennes, CNRS, Inserm, IRISA UMR 6074, Emgenn ERL U 1228, Rennes, France
² Big Data Institute, Li Ka Shing Centre for Health Information and Discovery, Nuffield Department of Population Health, University of Oxford, Oxford, UK
* Sharing last authorship

Problem
Researchers using fMRI data have a wide range of analysis tools to model brain activity. This diversity of analytical approaches means there are many possible variations of the same imaging result (Bowring et al., 2019). Analyzing a dataset with a single approach can thus be misleading.
Alternatively, a multiverse analysis can be used, where multiple sets of results are obtained from running different pipelines on the same single dataset. Such a setting produces several test statistic maps. Meta-analysis approaches can then help to effectively extract valid and robust results from these maps.
A required assumption for traditional meta-analyses is the independence among input datasets (Fig 1a). This assumption is no longer true in multiverse setting (Fig 1b), thus treating dependent studies as independent may lead to invalidity and thus decrease the generalizability of the result.
Therefore, we present a variety of same data meta-analysis (SDMA) models. The validity and accuracy of these models were assessed in a set of different simulations as well as on a real world dataset from "NARPS", a multiverse analysis with 70 different statistic maps originating from the same data. We also contrasted the outcomes of these SDMA models with those of a conventional meta-analysis approach.

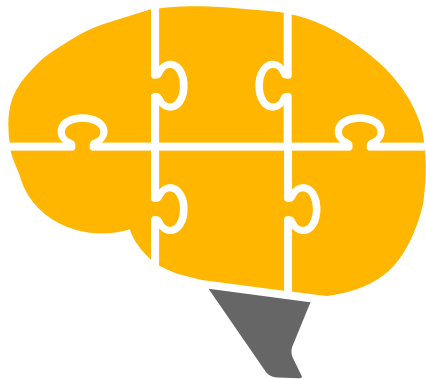
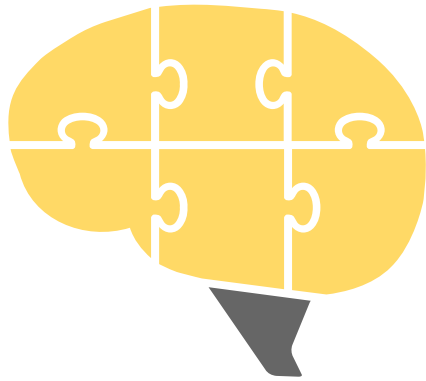
a Per study dataset | Per study pipeline | Per study result
Study 1
Study 2
...
Study K

b Unique dataset | Multiverse pipelines | Per pipeline result
Traditional meta-analysis result | Same-data meta-analysis result

Fig 1: Traditional meta-analysis setting (a) and multiverse setting (b). The traditional approach assumes independence among images (K studies) and thus K pipelines computed with N different participants each) when estimating summary effect. This underlying assumption is undermined in the multiverse setting where the outcomes are dependent over pipelines (One unique set of participants data for each pipeline).



<https://inserm.hal.science/inserm-04222730>



Towards reproducible neuroimaging across different analysis pipelines

Camille Maumet

Thank you!



We are hiring

<https://team.inria.fr/empenn/job-offers>

Credit: Presentation template and icons by [SlidesCarnival](https://www.slidescarnival.com/), adapted



@cmaumet@fediscience.org



@cmaumet