Reporting Whole Brain Data

Jack Van Horn
Laboratory of Neuro Imaging
UCLA

Themes of what to report:

- How much detail about my study methods should one provide?
- Does how one did their study really make any difference as long as the results are significant?
- Do the choices one makes really influence the outcome of the study?
- How big should one's study sample size be?
- Should one consider submitting their article as soon as they get a "publishable" (e.g. significant) result?
- How can one share their processing methods with others?

In this talk, Jack could….

- Impress upon you his vast and superior knowledge of neuroimaging research
- Tell you “how” to report your results
- Tell you “what” stuff to report
- Where to report it
- Give you a top ten list of “thou shalts”

- Or

- Illustrate quantitatively from the literature and from tools specifically designed to aid the reporting of study methods.

Three elements of this talk

- Meta-analysis of clinical neuroimaging study results
- Assessing one common type of bias in fMRI activation studies
- Data processing provenance

Meta-analysis

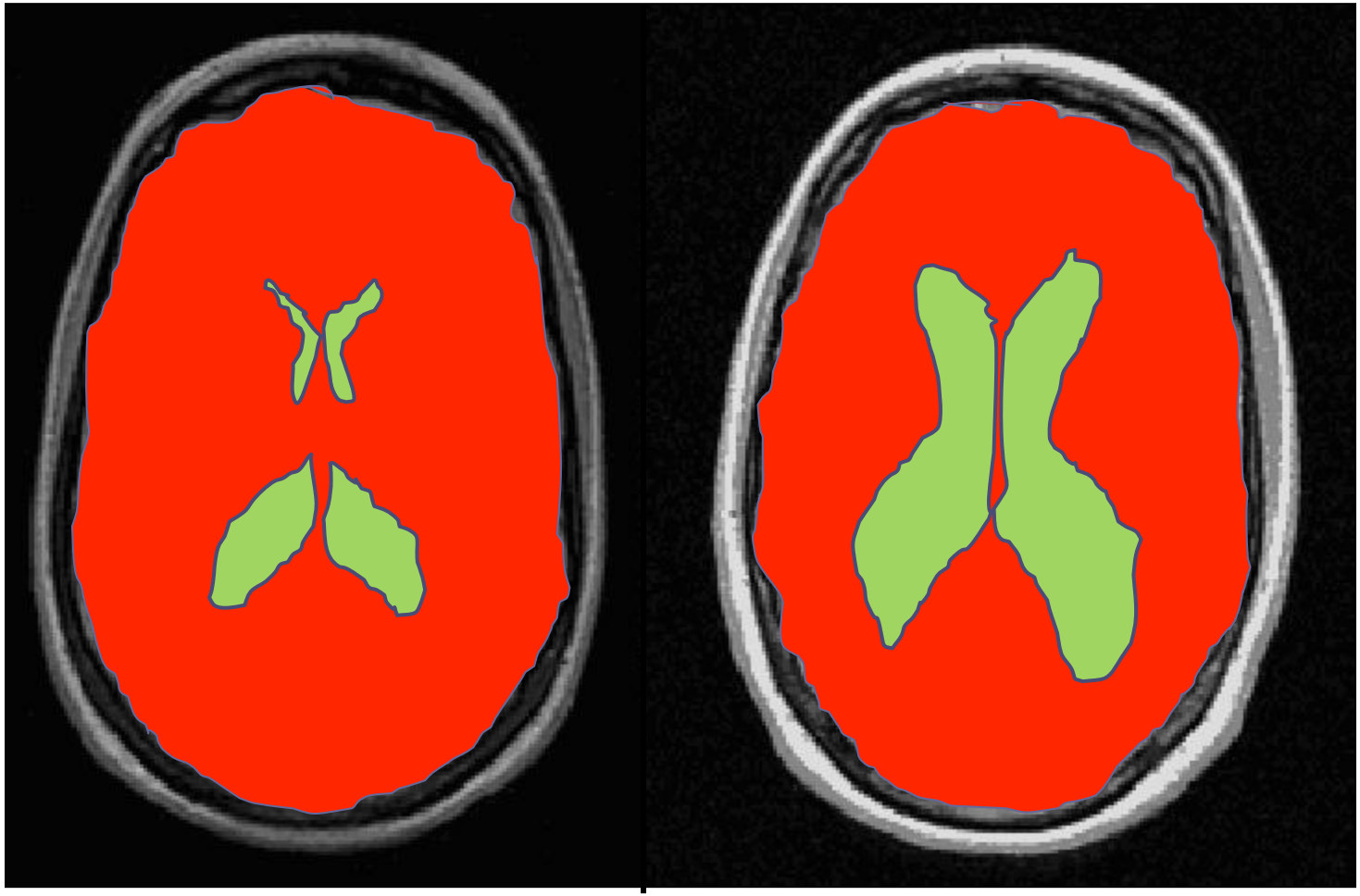
Meta-Analysis in Neuroimaging (1)

- Meta-analyses have become increasingly popular in neuroimaging.
- The development of highly detailed neuroimaging databases of published results has made quantitative assessment of the available research much easier.
  - i.e. BrainMap

Meta-Analysis in Neuroimaging (2)

- Meta-analyses of neuroimaging studies provide invaluable insights.
- However, published results in fMRI research are primarily small-study effects.
  - Recruitment of subjects is demanding.
  - Large samples are costly.
- Permit one to systematically examine the effects of commonly reported variables on published study effect sizes

Ventricular Enlargement in Schizophrenia



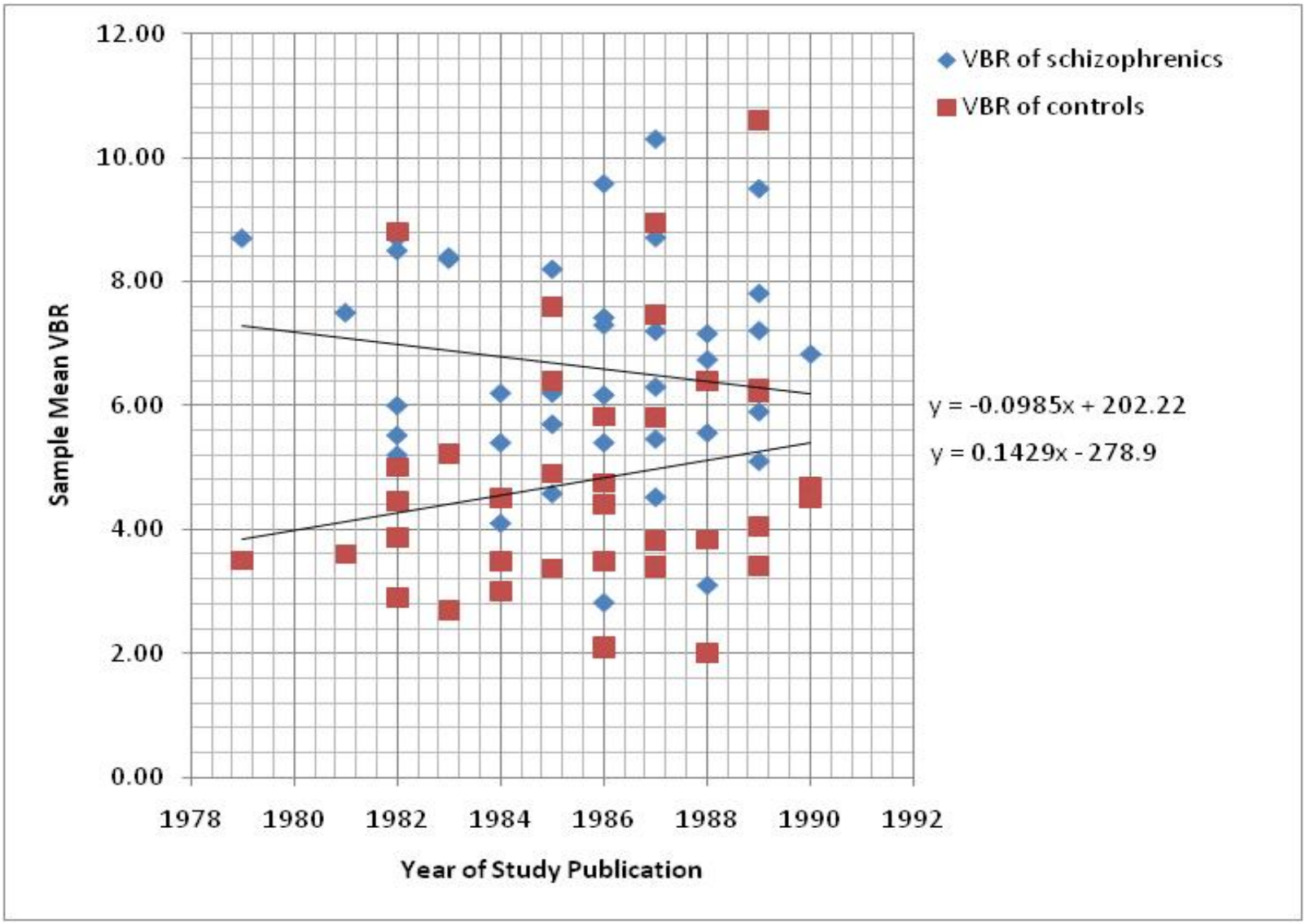
Normal Control Subject

Patient with Schizophrenia

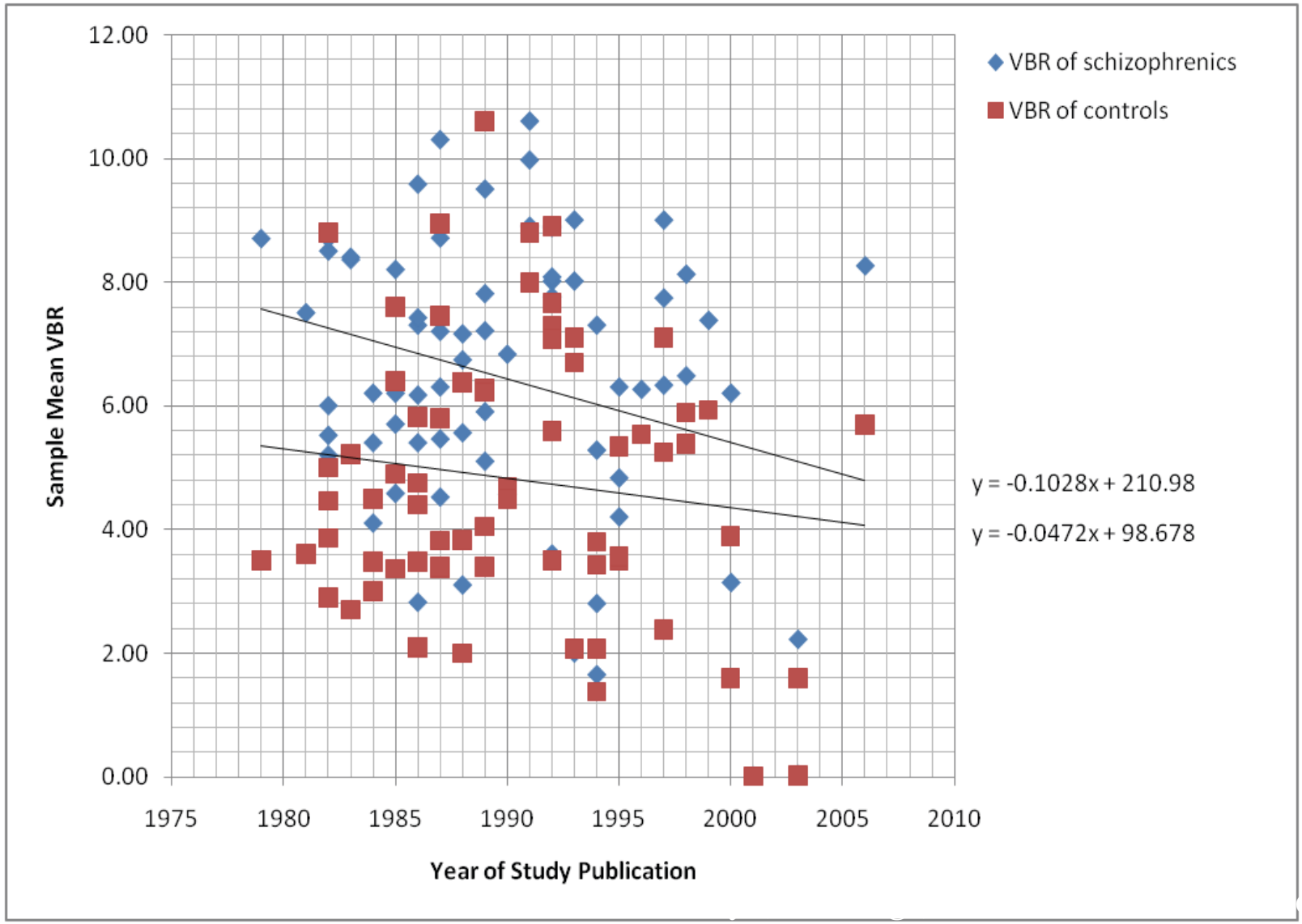
Meta-Analysis

- A follow-up study from 20 years ago
- 78 total studies of the VBR in schizophrenic patients and control subjects between 1978 and 2010
- Factors considered included:
  - Subject Age, Gender
  - Patient types
  - Diagnostic criteria
  - Year of study
  - Etc
- Within and between subject VBRs subjected to multivariate regression against these factors

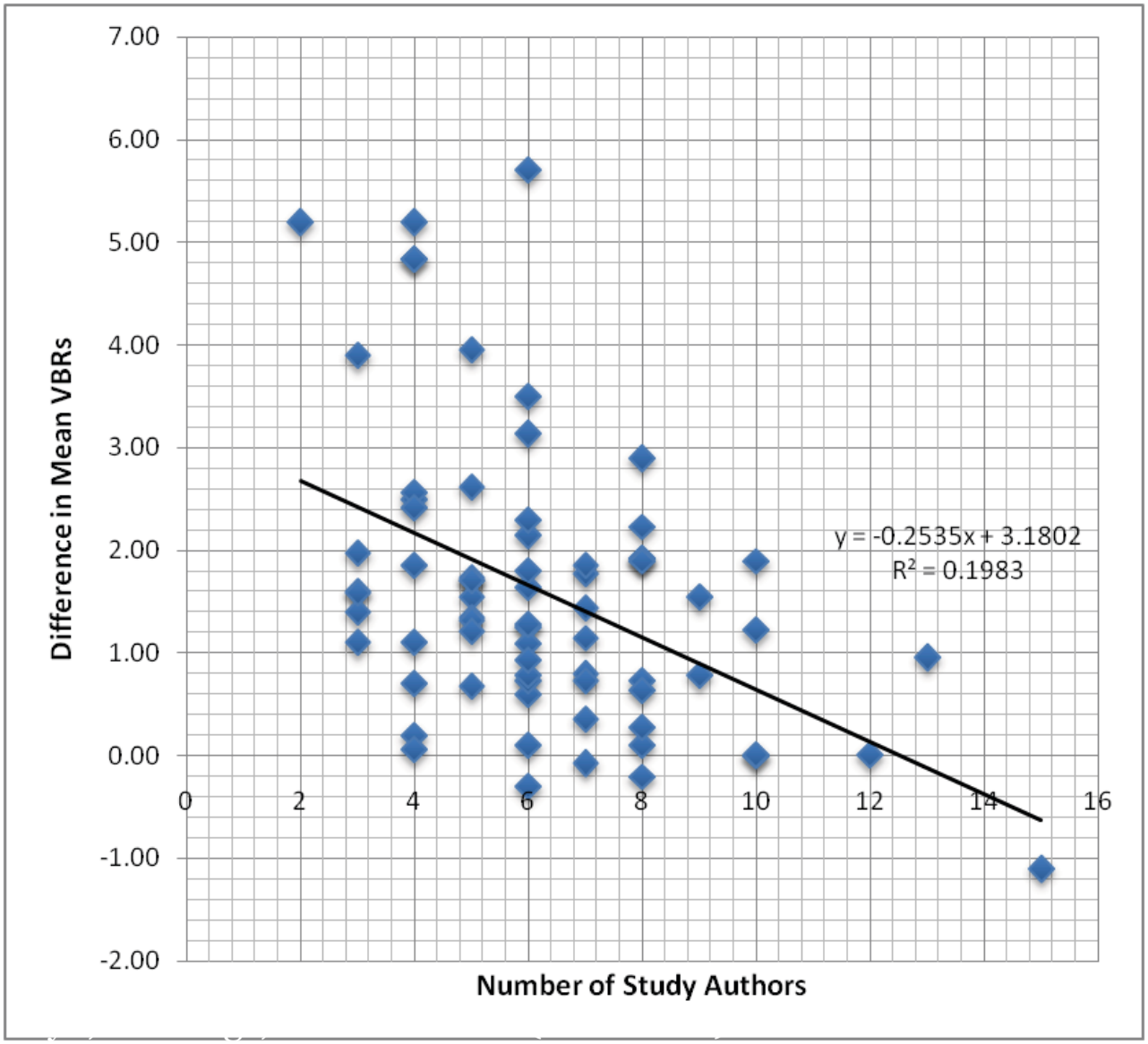
Year of Study Effect (1991)



Year of Study Effect (2011)



Number of Authors Effect



Sayo, Jennings, & Van Horn, in press, *NeuroImage*

Additional Predictive Factors…

- Diagnostic criteria
  - RDC
  - DSM
  - DSM+RDC
    - Combined criteria had smallest effect sizes
- Chronicity
  - Subject's labeled as acute were reported to have smaller ventricles than chronic schizophrenic patients
- The type of control subjects used
  - "Healthy" neurological patients
  - Normal volunteer subjects
    - Volunteers were less variable than patients
- BUT a number of things were not consistently reported which would have been really helpful
  - Age of onset, duration of illness, numbers of males/females, medication histories, drug use, concurrent diagnoses, etc

Ventricular Enlargement in Schizophrenia

- Conclusions
  - Ventricular enlargement exists and likely the most prominent biomarker present in the disease
  - However, study-based factors can significantly influence the size of reported differences between patients and controls
  - Selection of appropriate experimental controls and “sociological factors” affect group differences
  - If we had “complete” meta-data from such studies, direct assessment of ventricular size and potential biological precursors would be possible

Publishing and potential for bias

Introduction to Publication Bias (1)

- Publication bias occurs when the publication of research results depends on their nature and direction.
  - Statistically significant results are 3x more likely to be reported than results affirming a null hypothesis.

Introduction to Publication Bias (2)

- Such bias occurs despite the fact that studies with significant results do not appear to be superior to studies having null results with respect to quality of design.
- This leads to bias in the overall published literature toward only those effects considered to be statistically significant.

Relation to the "File-Drawer" Effect

- The "file-drawer" effect is the failure to report non-significant findings in any field of study.
  - Unpublished studies are "left in the file drawer".
- Hypothetically, the literature might contain only the 5% of studies that obtained significant $p<0.05$ values by chance alone, with the remaining 95% of non-significant studies unavailable for meta-analytic consideration.

Implications for Meta-analysis

- Not reporting negative effects can bias true average statistical effect sizes and mask particular trends present across studies over time.

- Publication bias present in the literature can severely hamper subsequent meta-analytic assessments from neuroimaging summary data archives containing reported statistical effects.

- Once detected, the influence of publication bias can be properly accounted for.

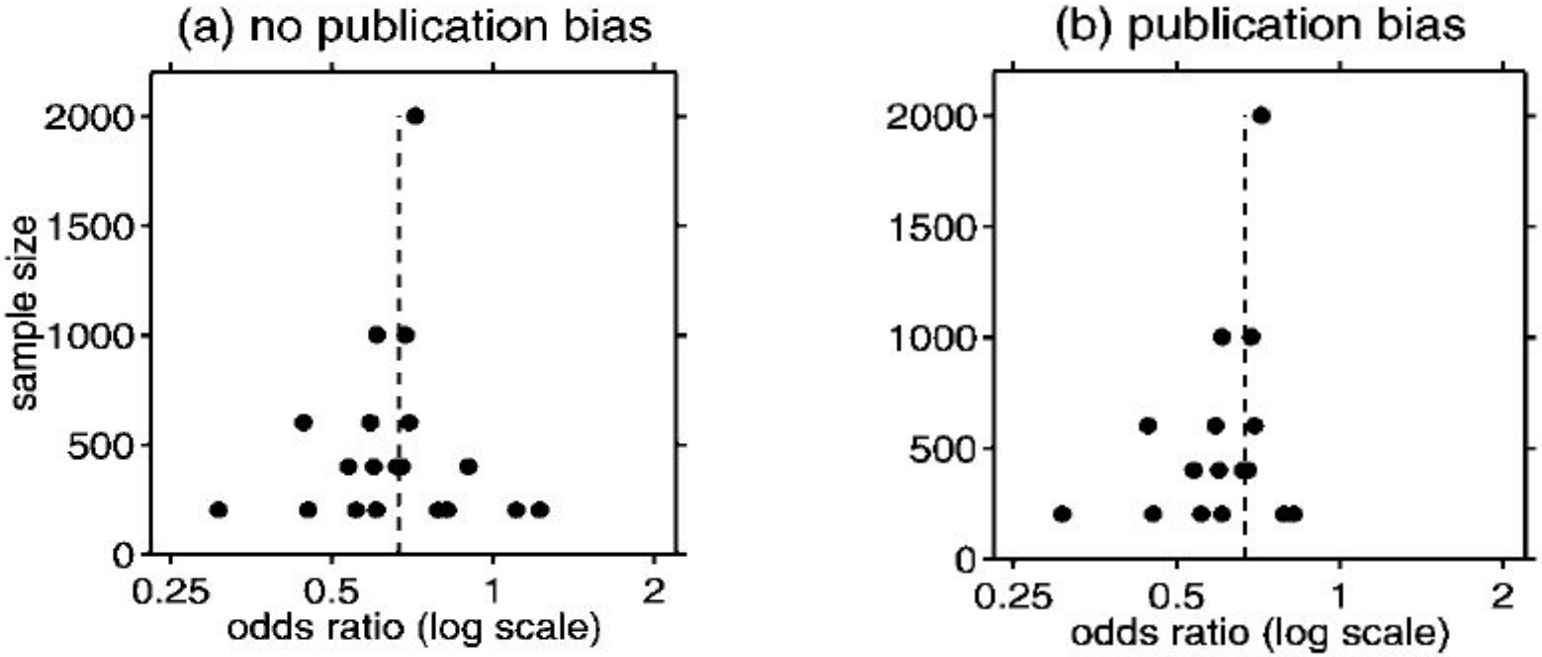
Methods for Detecting Publication Bias

(1) The funnel plot
(2) Macaskill's regression
(aka funnel plot regression)
(3) Egger's regression
(4) Trim and fill method

Jennings & Van Horn, 2011, *Neuroinformatics*

The Funnel Plot

- Publication bias can be examined by visual inspection of a funnel plot.
  - (Light & Pillmer, 1984)
- Plot of effect size (x) by sample size (y)
  - Existence of publication bias determined by symmetry, or lack thereof, in the generated graph.
- Drawback: subjective test.



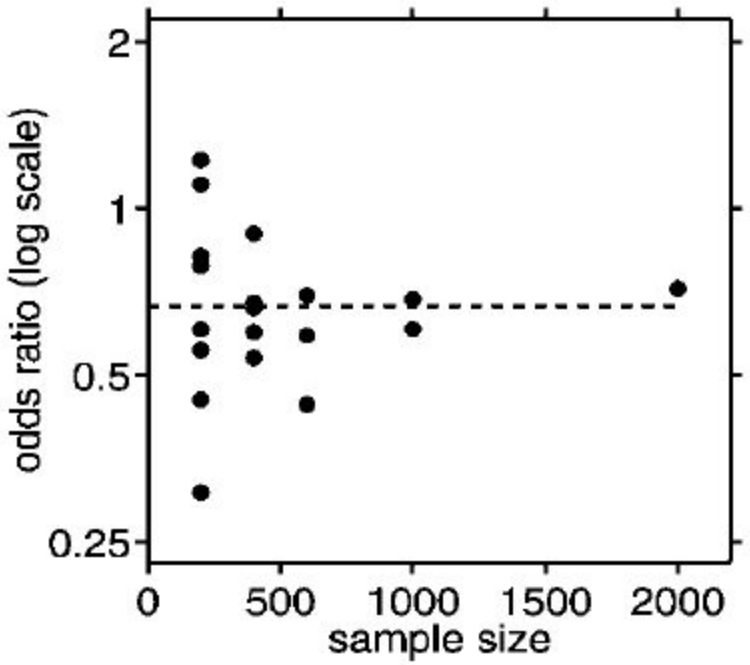
(a) Points form a symmetrical inverted funnel around the overall estimate of the effect, with results from smaller studies scattered more widely about the mean at the bottom of the graph.

(b) Graph is asymmetrical/skewed, in this example skewed to the left.

A 'classic' asymmetry involves non-publication of insignificant studies which causes gaps in the bottom left-hand corner of the graph and leaves the plot skewed to the right.

Macaskill's Regression

- Draws on the idea of the funnel plot more systematically (Macaskill, Walter *et al.* 2001).
- Linear regression model with the effect size as the dependent variable, and sample size as the independent variable.
  - WLS regression approach is taken with the effect size weighted by its inverse variance.
- Benefit: low false-positive rate



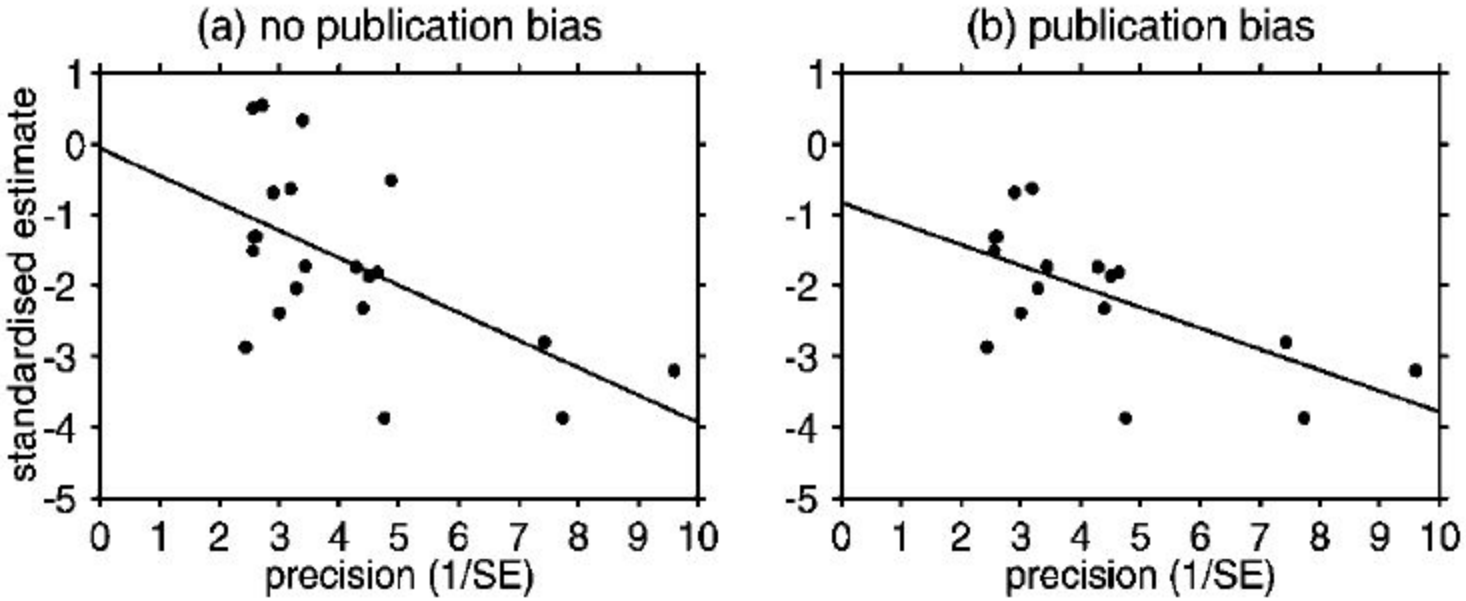
• If the funnel plot is symmetrical, i.e. no publication bias, the regression slope has an expected value of zero [Ho], shown above.

• A non-zero slope would suggest an association between effect size and sample size, potentially due to publication bias [Ha].

Egger's Regression

- Linear regression model with the effect size standardized by its standard error and regressed against its precision, defined as the inverse of the standard error (Egger, Davey Smith *et al.* 1997).

- Benefits: highly sensitive, strong statistical power.
- Drawback: high false-positive rate.



(a) No publication bias: the intercept will have an expected value of zero, and the slope is an unbiased estimate of the true (underlying) effect.

(b) Publication bias: the fitted line does not pass through the origin (intercept ≠ 0).

Trim & Fill method

- Non-parametric approach.
- Assumes that in addition to the number of published studies (n), there are another k_o studies which have not been reported due to non-statistical significance.
- Method ranks studies based on the absolute values of their deviations from the mean effect size.
  - ranks of studies with effect sizes smaller than the mean are given a negative sign, and those greater than the mean retain a positive sign.

Trim & Fill method mathematically…

$r_i = \text{rank}(|a_i - \bar{a}|)$
with a negative sign given to the r_i where $a_i < \bar{a}$
(individual effect size, a_i, and mean effect size $\bar{a}$)

γ^*, the length of the rightmost run of ranks associated with positive values of the observed ri, is defined as:

$\gamma^* = n - rh$

where n= number of studies in the meta-analysis and rh is the largest negative rank in the sample.
And k0 is estimated by R0, the "rightmost run" estimator, where:

$R0 = \gamma^* - 1$

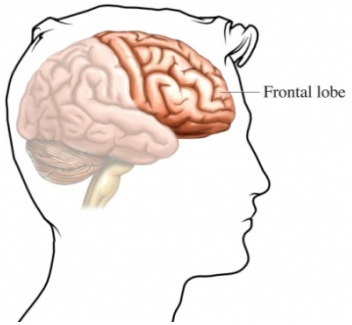
R0 is the sample estimate of k0, the number of studies which have not been reported due to publication bias.

Subsequently, publication bias is evident when R0>3 , as outlined by Duval and Tweedie (2000).

Methods in fMRI research

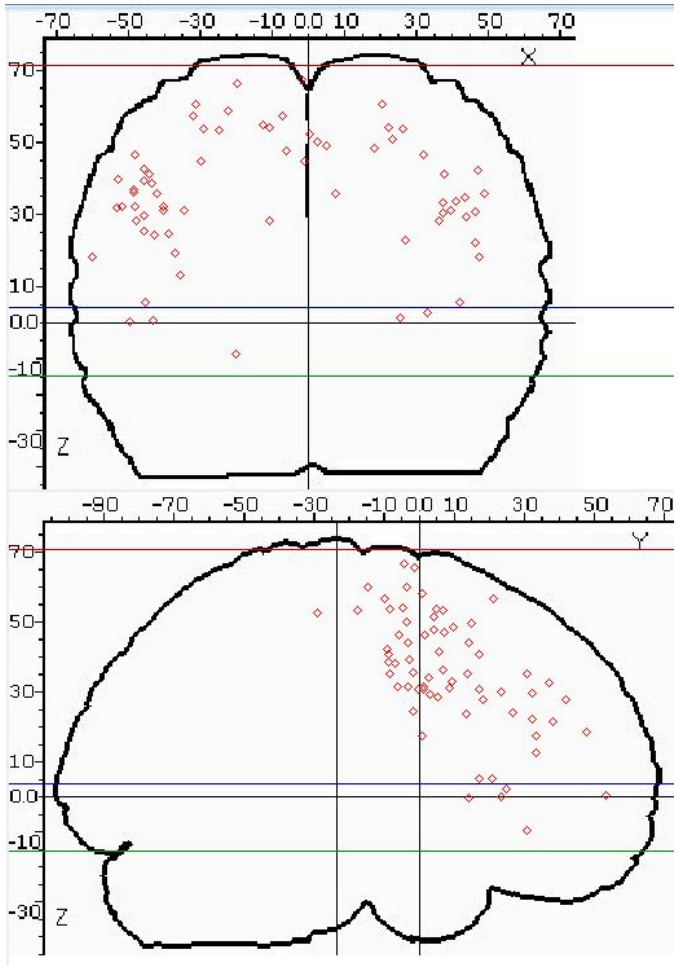
- Studies were included for analysis using the Sleuth program in BrainMap.
  - An application used to search for papers of interest and read their corresponding meta data.

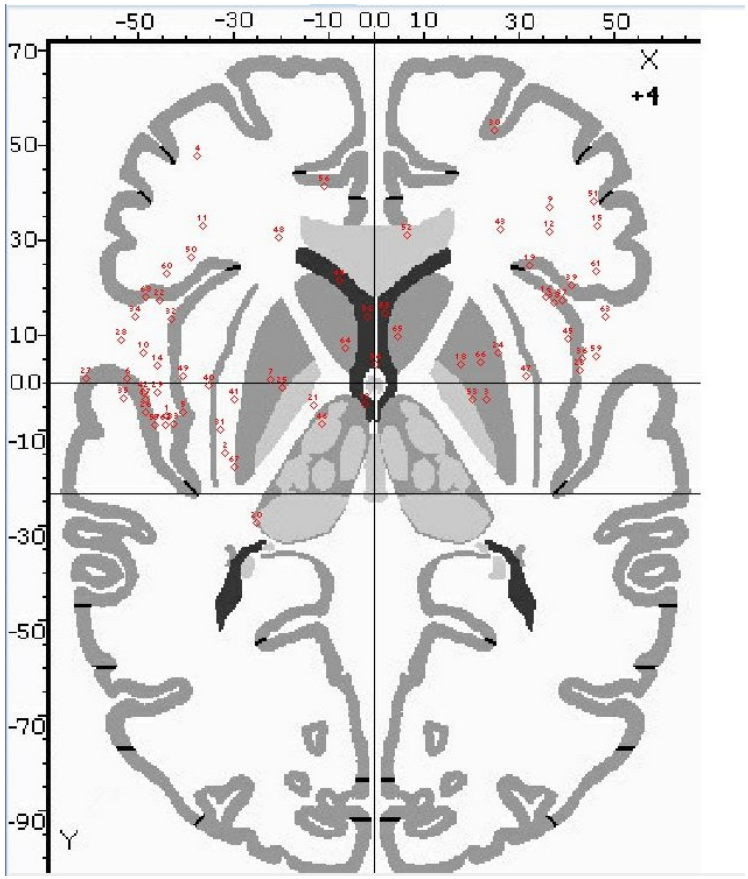
- Examined studies of *working memory* in a control population identifying *activation in the frontal lobes*.
  - Narrow focus
  - Long history of research in this area
  - Well-tested task paradigms



Methods in fMRI research (cont.)

- 162 papers were found in the database, of which 77 were selected that gave statistical information to calculate effect size (Cohen's d, Pearson's r).
- 74 studies were examined in the final cohort (3 were excluded since they did not include a control population).
- 68 out of the 74 papers reported a significant effect in the frontal lobe.
  - Studies which were classified as using a working memory task that did not report activation in the frontal lobe had their effect size set to zero.





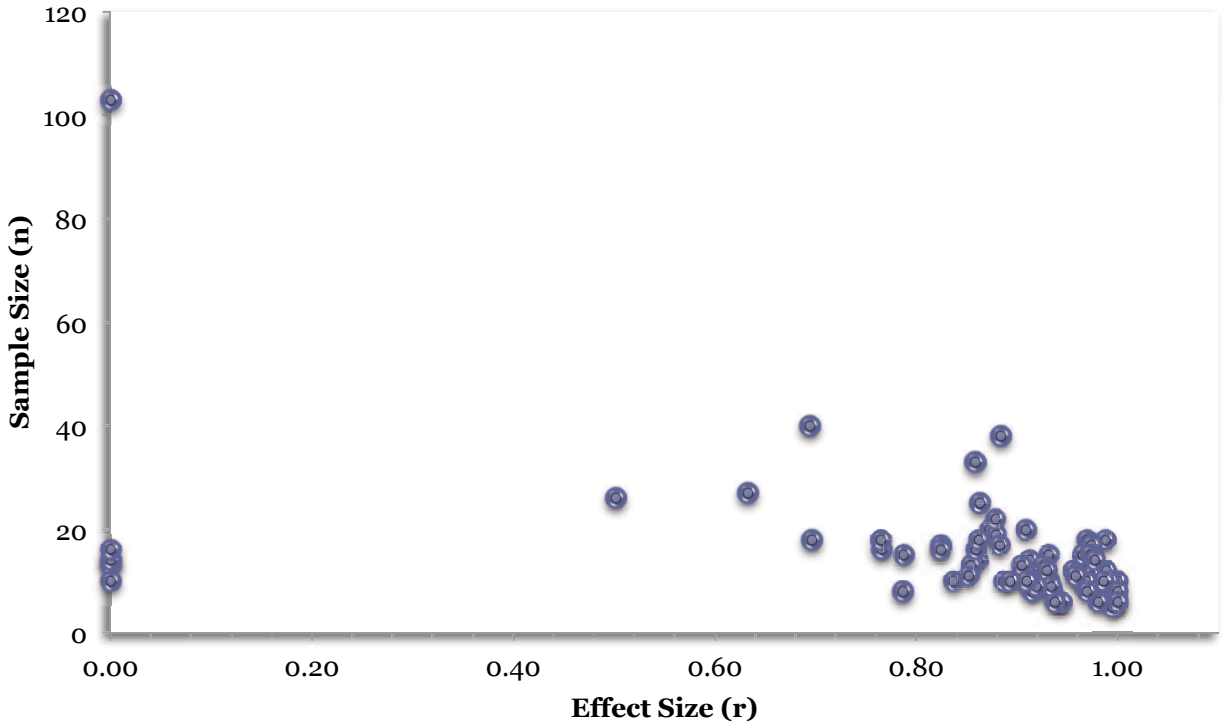
Results for studies, plotted on a standard glass brain in Talairach space using BrainMap, showing each reported study local maxima located in the frontal lobes (n=68)

Results (1)

- Effect sizes were generally very large, with a number of extreme/outlier values.
- Pearson's r was plotted against sample size for the best visual assessment of the funnel plot.
- Cohen's d effect size was used in the formal evaluation of publication bias – Macaskill's, Egger's regression & the Trim and Fill method.

Funnel Plot

- Pearson's r – 'classic' funnel plot asymmetry:



Results (2)

- Cohen's d
  - Analysis performed both with and without outlier values.
  - Outliers defined as values greater than the 95% percentile (n=4).

Results with outliers (n = 74)

- Macaskill's regression:
  - F = 12.07, p=0.0009
- Egger's regression:
  - F = 6.7, p =0.01
- Trim and Fill Method:
  - Publication bias found to be present in both tails, R0 >3

Results without outliers (n = 70)

- Macaskill's regression:
  - F = 9.92, p=0.002
- Egger's regression:
  - F = 8.17, p =0.006
- Trim and Fill Method:
  - Publication bias found to be present in the right tail, R0 >3

Additional analyses

- Each study reported structural domains (Brodmann Areas, BA) as well as behavioral domains for the task paradigm.

- Publication bias was examined by structural and cognitive domain to determine if one region or one behavioral domain specifically studied in fMRI was driving the overall presence of publication bias.

Results of additional analyses (1)

- Region (BA)
  - Analyses of publication bias were performed individually in BA6 (n=33), BA9 (n=15), and 'other' BAs (n=26), due to low sample size.
    - All were found to be statistically significant ($p<0.05$). Thus findings of publication bias were not regionally dependent.
  - A multivariate analog of Macaskill's regression was performed across sub-regions
    - An overall finding of publication bias was still found with and without outliers, $p<0.0001$

Results of additional analyses (2)

- Cognitive Domain
  - Analyses of publication bias were performed individually by tasks of *strict working memory* (n=55) compared to 'other' domains (n=19)
    - Sub-domains of working memory still found to be statistically significant
  - Multivariate Macaskill's regression found the presence of publication bias across sub-domains.
    - P<0.0001 with and without outliers.

Discussion (1)

- Presence of publication bias found in fMRI of working memory in the frontal lobe.
  - While we focused on this region and functional domain, they were chosen arbitrarily, and these findings are likely to generalize to fMRI literature as a whole.
    - With such small sample sizes per study (common in fMRI), expect that there should exist a large number of negative findings due to lack of power alone
    - Analysis was not restricted by age (min age 7, max age >61). Patterns of activation in children vary much more widely than adults, and therefore we might expect more studies with non-significant findings.

Discussion (2)

- The importance of openly accessible data archives such as BrainMap should not be overshadowed by this outcome.
  - Universal coordinate databases and a space to share meta-data for analysis are invaluable in the imaging field.

- Due to inherent but unintended publication bias in archives of study summary data, appropriate methods can detect such bias.
  - Trim & Fill
  - "Fail-safe File drawer" estimate

Limitations

- Assessment of publication bias was limited due to the use of BrainMap.
  - Not all studies are indexed in this database, though there is no evidence that this archive is systematically biased compared to the literature as a whole.
  - In focusing on BrainMap we sought to note what would likely be present in any particular archive of published results (and which may be particularly used for meta-analyses).

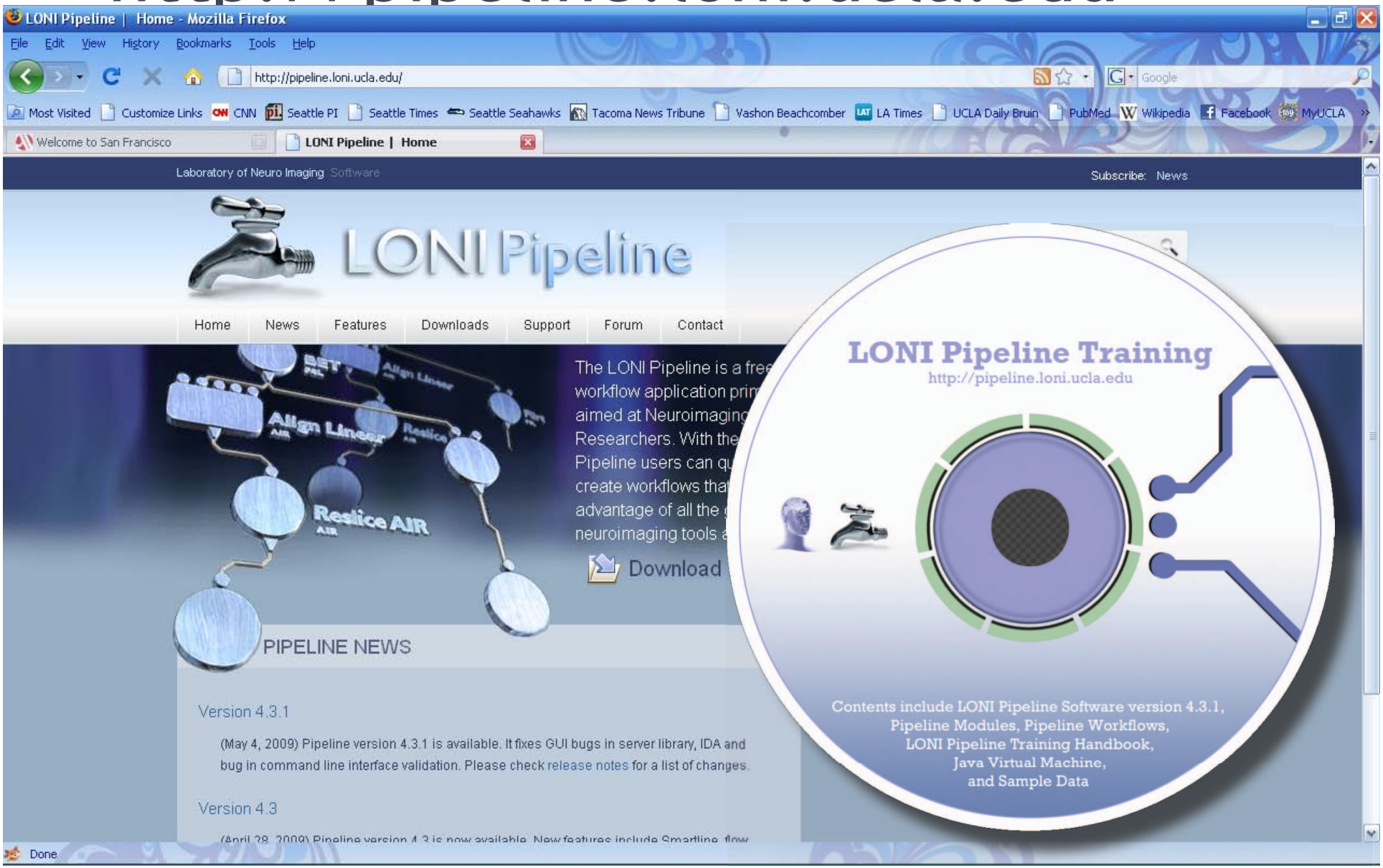
- Low sample size and power can affect the reported correlations and effect sizes, inflating the results.
  - The so-called “voodoo” correlations
  - This type of bias is closely linked to publication bias, and it is impossible to specifically distinguish between them.

Provenance

Provenance

- *Provenance* is the record or history of things
  - In scientific data, it refers to the history of what operations were done to data, how they were done, in what order, etc
    - e.g. “Raw data were aligned using FSL FLIRT, a mean EPI image was then computed, this was used to…..”
    - e.g. Computer-based approaches to recording the history of processing steps

http://pipeline.loni.ucla.edu

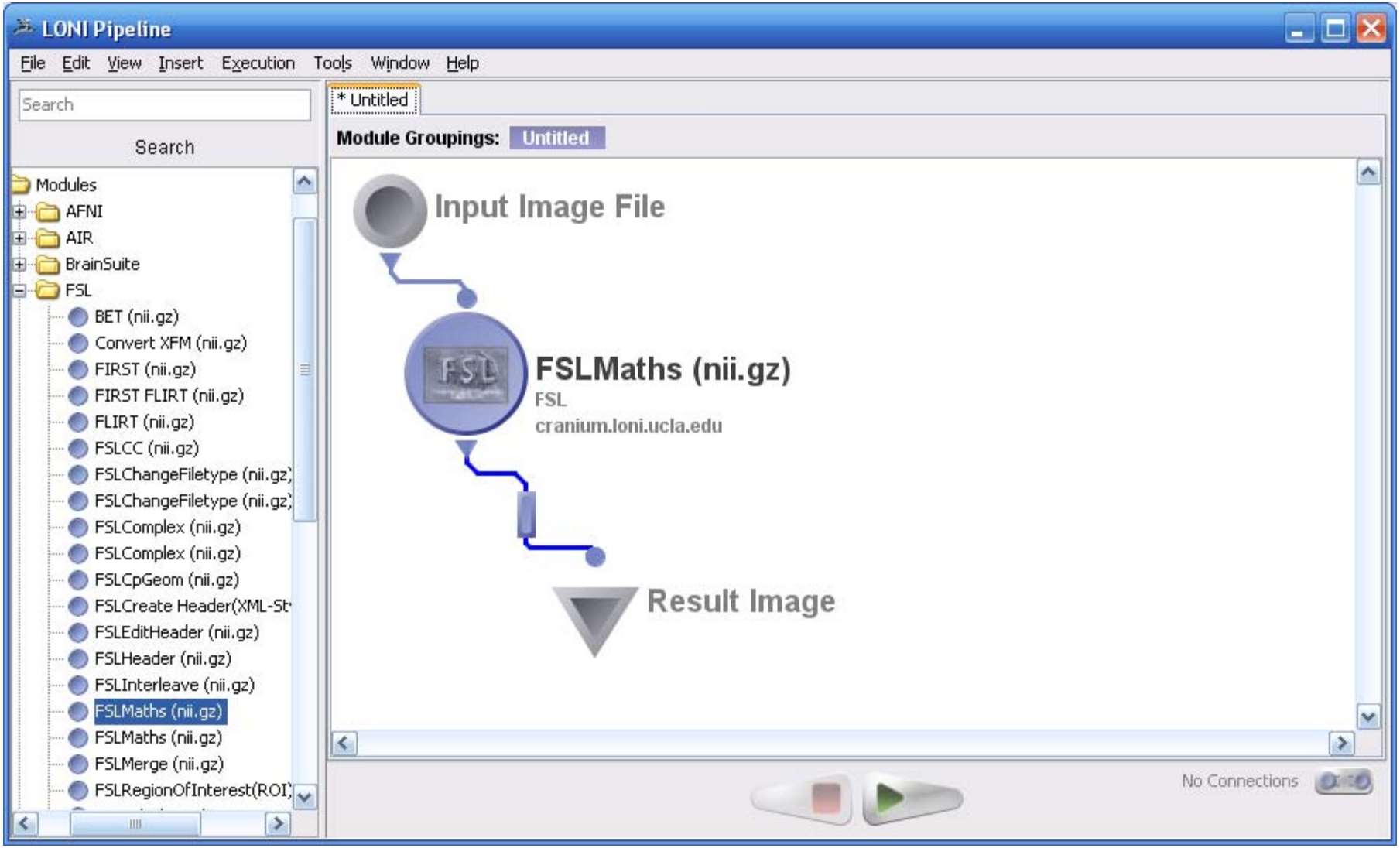


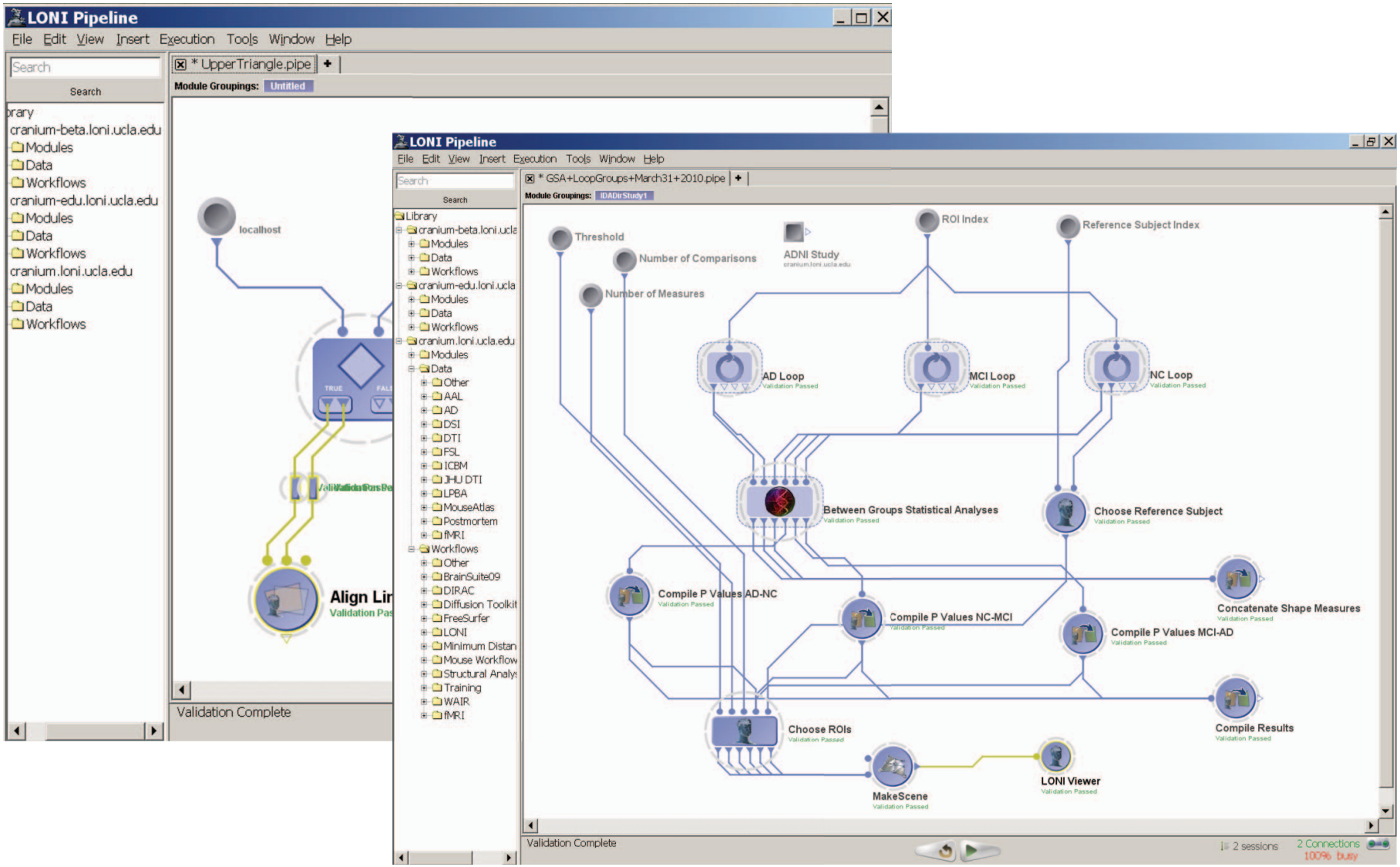
LONI Pipeline Background

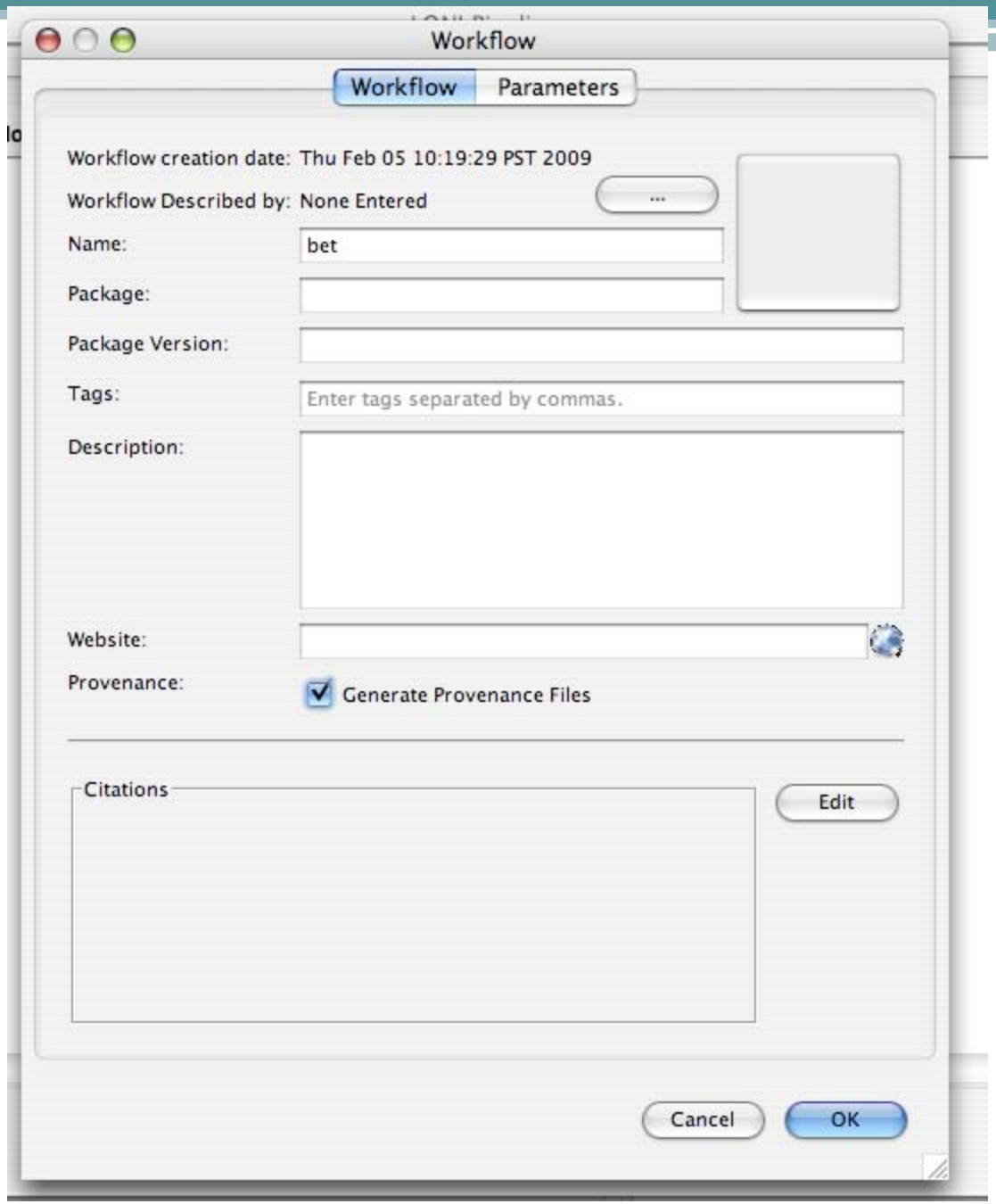
- Recognition that most neuroimaging processing steps are input/output controlled by several parameters, options, and flags
- Neuroimage processing is often a connected set of inputs, outputs, which form the next set of inputs to processing executables
- People with the right knowhow "script" these steps (e.g. FSL, AFNI, FreeSurfer, homegrown) using various programming languages (e.g. Matlab, bash, Python, C++, etc) into "workflows"
- Variability in workflow construction within/ between laboratories

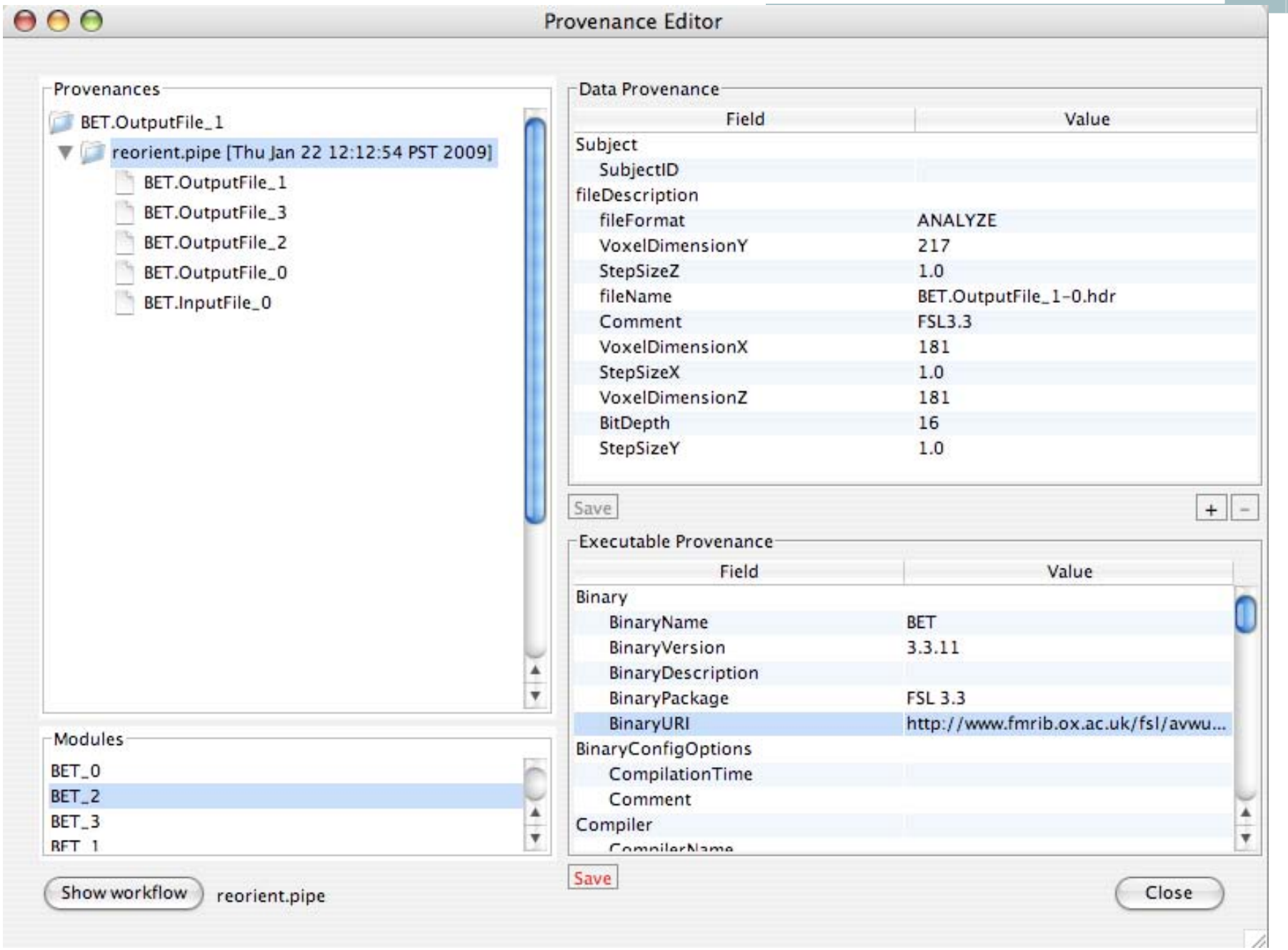
What is LONI Pipeline?

- A graphical workflow design and runtime environment now Version 5.2
- “Modularizes” many common neuroimaging software tools
- Enables linking of modules to create data processing and analysis “pipelines”
- Permits workflows containing heterogeneous software elements
- Runs using the 1200+ CPU LONI grid


LONI Pipeline
File Edit View Insert Execution Tools Window Help
Search
Search
Modules
AFNI
AIR
BrainSuite
FSL
BET (nii.gz)
Convert XFM (nii.gz)
FIRST (nii.gz)
FIRST FLIRT (nii.gz)
FLIRT (nii.gz)
FSLCC (nii.gz)
FSLChangeFiletype (nii.gz)
FSLChangeFiletype (nii.gz)
FSLComplex (nii.gz)
FSLComplex (nii.gz)
FSLCpGeom (nii.gz)
FSLCreate Header(XML-St
FSLEditHeader (nii.gz)
FSLHeader (nii.gz)
FSLInterleave (nii.gz)
FSLMaths (nii.gz)
FSLMaths (nii.gz)
FSLMerge (nii.gz)
FSLRegionOfInterest(ROI)
* Untitled
Module Groupings: Untitled
Input Image File
FSL
FSLMaths (nii.gz)
FSL
cranium.loni.ucla.edu
Result Image
No Connections


LONI Pipeline
File Edit View Insert Execution Tools Window Help
Search
Search
brary
cranium-beta.loni.ucla.edu
Modules
Data
Workflows
cranium-edu.loni.ucla.edu
Modules
Data
Workflows
cranium.loni.ucla.edu
Modules
Data
Workflows
* UpperTriangle.pipe
Module Groupings: Untitled
localhost
Validation Complete
LONI Pipeline
File Edit View Insert Execution Tools Window Help
Search
Search
Library
cranium-beta.loni.ucla
Modules
Data
Workflows
cranium-edu.loni.ucla
Modules
Data
Workflows
cranium.loni.ucla.edu
Modules
Data
Other
AAL
AD
DSI
DTI
FSL
ICBM
JHU DTI
LPBA
MouseAtlas
Postmortem
fMRI
Workflows
Other
BrainSuite09
DIRAC
Diffusion Toolkit
FreeSurfer
LONI
Minimum Distan
Mouse Workflow
Structural Analys
Training
WAIR
fMRI
* GSA+LoopGroups+March31+2010.pipe
Module Groupings: IDADirStudy1
Threshold
Number of Comparisons
Number of Measures
ADNI Study
ROI Index
Reference Subject Index
AD Loop
MCI Loop
NC Loop
Between Groups Statistical Analyses
Choose Reference Subject
Compile P Values AD-NC
Compile P Values NC-MCI
Compile P Values MCI-AD
Concatenate Shape Measures
Choose ROIs
Compile Results
MakeScene
LONI Viewer
Validation Complete
2 sessions
2 Connections


Workflow
Workflow
Parameters
Workflow creation date: Thu Feb 05 10:19:29 PST 2009
Workflow Described by: None Entered
...
Name:
bet
Package:
Package Version:
Tags:
Enter tags separated by commas.
Description:
Website:
Provenance:
Generate Provenance Files
Citations
Edit
Cancel
OK


Provenance Editor
Provenances
BET.OutputFile_1
reorient.pipe [Thu Jan 22 12:12:54 PST 2009]
BET.OutputFile_1
BET.OutputFile_3
BET.OutputFile_2
BET.OutputFile_0
BET.InputFile_0
Modules
BET_0
BET_2
BET_3
Show workflow
reorient.pipe
Data Provenance
Field
Value
Subject
SubjectID
fileDescription
fileFormat
ANALYZE
VoxelDimensionY
217
StepSizeZ
1.0
fileName
BET.OutputFile_1-0.hdr
Comment
FSL3.3
VoxelDimensionX
181
StepSizeX
1.0
VoxelDimensionZ
181
BitDepth
16
StepSizeY
1.0
Save
+
-
Executable Provenance
Field
Value
Binary
BinaryName
BET
BinaryVersion
3.3.11
BinaryDescription
BinaryPackage
FSL 3.3
BinaryURI
http://www.fmrib.ox.ac.uk/fsl/avwu...
BinaryConfigOptions
CompilationTime
Comment
Compiler
Save
Close

Provenance Conclusions

- Tools such as LONI Pipeline are specifically designed for recording how data was processed
- These files can be shared with others or included as supplemental methods
- Other tools as well
  - E.g. Nipy, SPM batch scripts, etc
- Sharing these files
  - Reduced the level of detail you write about in your methods section
  - But provides more net detail overall
  - People can more accurately reproduce your findings

Talk Summary

- Meta-analyses are useful in neuroimaging and other fields of research, increasing sample size and power for detecting an effect.
- Factors related to “how” your study was done can significantly influence your results.
  - This may even relate to the size of your research team!
- Publication bias effects were found to be present in results contained in an fMRI data
  - May be characteristic of the functional imaging literature as a whole.
- Provenance tools exist to help you
  - keep track of the operations your data goes through, the parameters, and order of steps so that you may share it with others.
- Be thorough and thoughtful about reporting your brain structure and function results.

Three elements of this talk

- Meta-analysis of clinical neuroimaging study results
  - Can illuminate those study factors that most influence reported results
  - Your choices make a difference!!
- Assessing one common type of bias in fMRI activation studies
  - Underscore having proper study power.
  - Non-significant findings should be published, too!
- Data/methods processing provenance
  - Workflow tools simplify and make possible methods sharing, e.g. as supplemental materials

Recommendations

- When reporting your neuroimaging result, be…
  - Rigorous and thorough
  - Consider how your choices influence your results
  - Have a “big” sample size to have adequate power
  - Don’t settle for “significance”
  - Be exhaustive about the detail you provide on how your data were processed (provenance)
  - Provide supplemental materials where allowed
  - Don’t just report the significant stuff

References

- Jennings, R. and **Van Horn, J.D.** (2011) "Assessing publication bias in neuroimaging results archives", Neuroinformatics, PMID: 21643733.
- Sayo, A., Jennings, R., and **Van Horn, J.D.** (in press) "Ventricular Enlargement in Schizophrenia: A 20-year follow-up meta-analysis", NeuroImage.
- Duval, S. and Tweedie, R. (2000). "Trim and fill: A simple funnel-plot-based method of testing and adjusting for publication bias in meta-analysis." Biometrics **56**(2): 455-463.
- Egger, M., Davey Smith, G., Schneider, M. and Minder, C. (1997). "Bias in meta-analysis detected by a simple, graphical test." BMJ **315**(7109): 629-634, '2127453:' 2127453.
- Light, R. J. and Pillemer, D. B. (1984). Summing up. The science of reviewing research. Cambridge, MA, Harvard University Press.
- Macaskill, P., Walter, S. D. and Irwig, L. (2001). "A comparison of methods to detect publication bias in meta-analysis." Stat Med **20**(4): 641-654.
- www.wikipedia.com

...among others