



From SAPs to SHEAPs: Perspectives from economics and elsewhere

Will Hollingworth

School of Social and Community Medicine, Bristol

MRC

Hubs for Trials
Methodology Research

ConDuCT-II Hub

Pre-specified [NOT] – presentation plan

Primary outcome

Summarise recent debate in mainstream economics about the pros and cons of “pre-analysis plans”

Secondary outcome

Identify any initiatives on HEAPs in health economics in other countries

Pre-specified [NOT] – presentation plan

Data collection

- ‘Filing cabinet’ review of “pre-analysis plans” in the economics literature [i.e. the JEP symposium in Sarah Wordsworth’s filing cabinet]
- ‘Snowballing’ to other relevant articles [i.e. those that backed up the points I wanted to make]
- Survey of health economists (n=3) in other English speaking countries [i.e. HE friends in Australia, Canada, USA]

Pre-specified [**NOT**] – presentation plan

Data analysis

- Selective narrative summary of the economics literature
- Survey data, not informative [i.e. interesting], **so not reported in detail**

JEP Symposium

Olken, B. [Promises and perils of pre-analysis plans](#). Journal of Economic Perspectives: 29; 3. p61–80.

Coffman, L & Niederle, M. [Pre-analysis Plans Have Limited Upside, Especially Where Replications Are Feasible](#). Journal of Economic Perspectives: 29; 3. p81-98.

Casey et al. [Reshaping Institutions: Evidence on Aid Impacts Using a Pre-Analysis Plan](#). The Quarterly Journal of Economics (2012)

doi: 10.1093/qje/qje027

What's the problem?

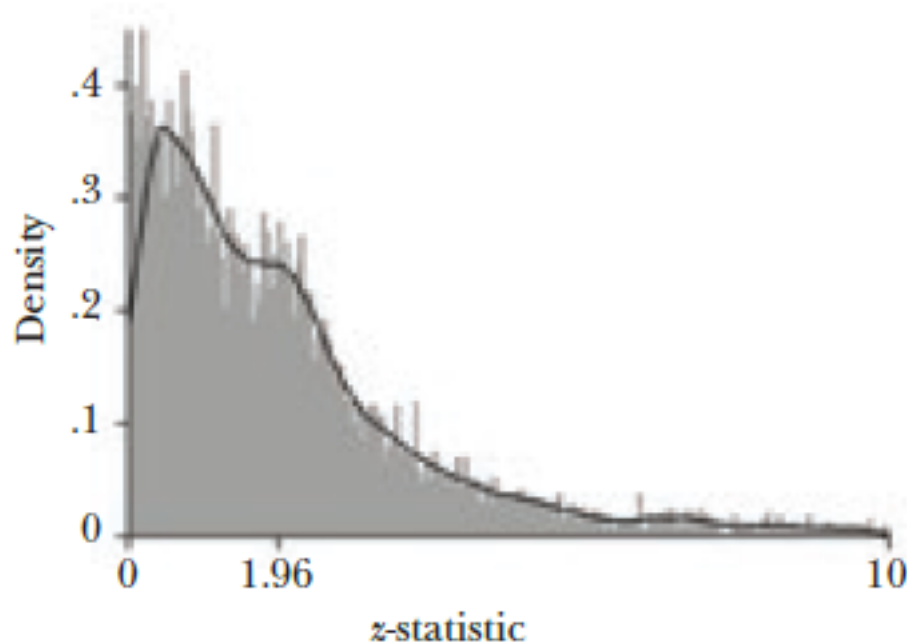
- The “nefarious researcher” – cherry picking
 - Which results to report
 - Which regressions to run (functional form, covariates)
 - Excluding outliers

Olken, B. [Promises and perils of pre-analysis plans](#). *Journal of Economic Perspectives*: 29; 3. p61–80.

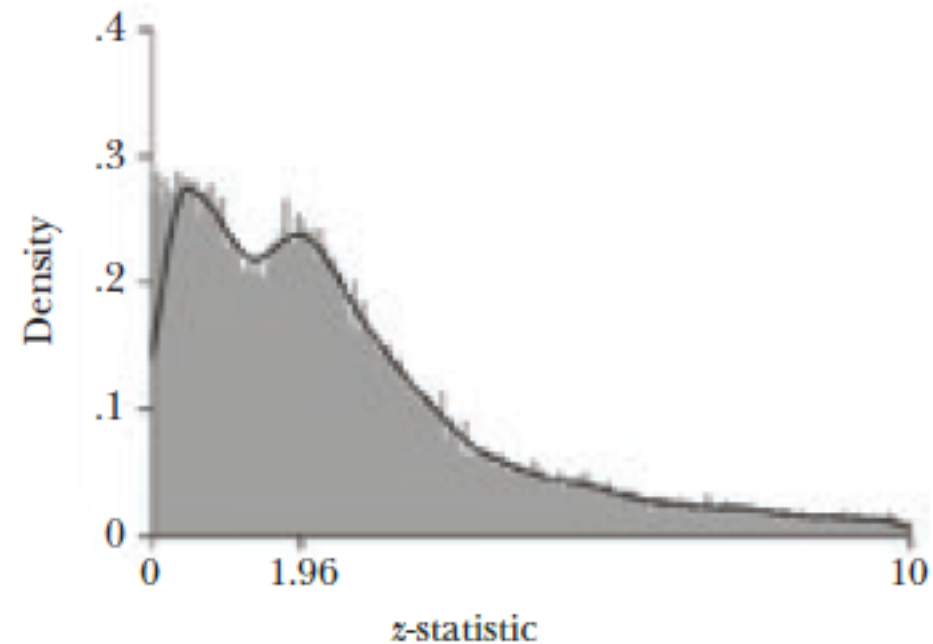
Not a problem for RCTs?

From: Olken, B. Promises and perils of pre-analysis plans. *Journal of Economic Perspectives*: 29; 3. p61–80.

A: Laboratory experiments or randomized control trials data



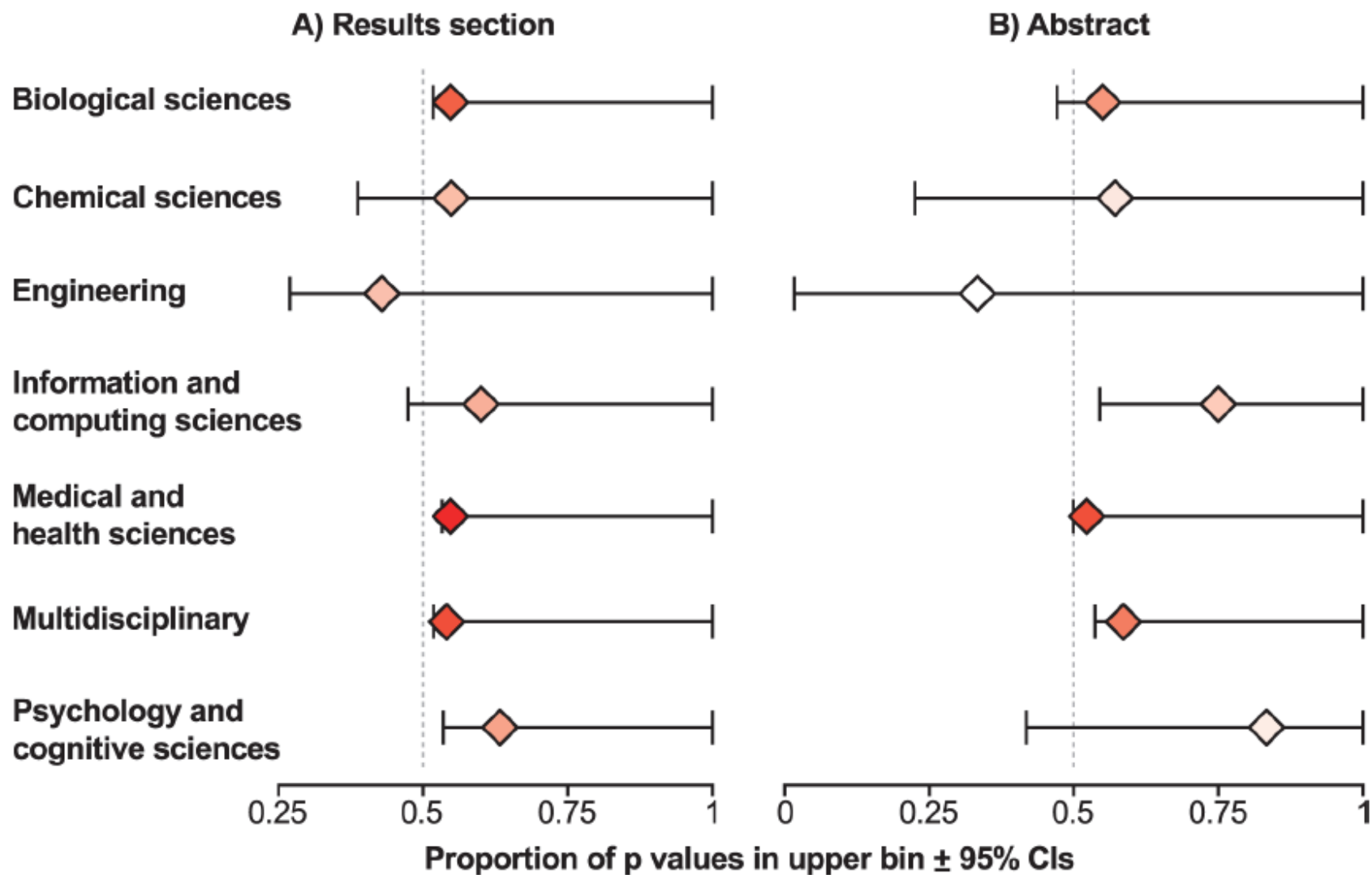
B: Other [nonexperimental] data



Source: Figures 6e and f from Brodeur, Lé, Sangnier, and Zylberberg (forthcoming).

Notes: Displays distribution of z-statistics reported in all papers appearing in either the *American Economic Review*, *Journal of Political Economy*, or *Quarterly Journal of Economics* between 2005 and 2011. Experiments, both lab and field, are in the left panel; all other papers in the right panel.

(Health) economics particularly vulnerable?

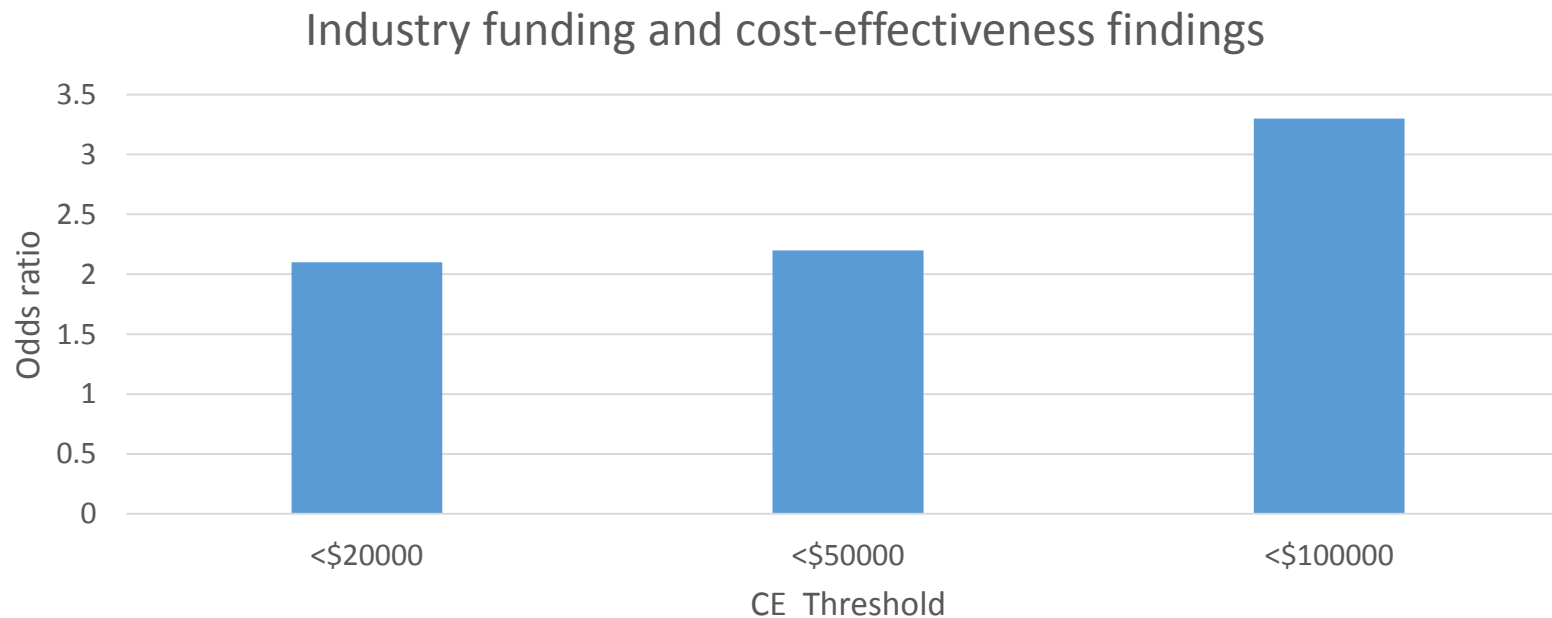


Head et al. PLoS
Biol 13(3): e
1002106.

Fig 3. Evidence for p-hacking across scientific disciplines. A) Evidence for p-hacking from p -values obtained from Results sections.

(Health) economics particularly vulnerable?

- Financial pressures sometimes high on health economic analyses



(Health) economics particularly vulnerable?

- Short journal word counts create pressures on transparency
- Tradition of presenting CIs and CEACs may protect us from over obsession with p-values

What's the problem?

“Even researchers who have the noblest of intentions may end up succumbing to the same sorts of biases when... ..[making] sense of a complex set of results”

Olken, B. [Promises and perils of pre-analysis plans](#). *Journal of Economic Perspectives*: 29; 3. p61–80.

If S[HE]APs are the answer: 1) Content

Item	HE Example
Primary outcome	NHS & PSS cost per QALY gained at trial end
Secondary outcomes	Societal cost per QALY; cost per DFS, etc. etc.
Variable definitions	Unit costs to be applied
Inclusion/exclusion rules	Missing values, outliers.
Model specification	Costs (e.g.), QALYs
Covariates	e.g. key baseline variables
Subgroup analyses	By key clinical / demographic characteristics
Other issues	e.g. interim analyses

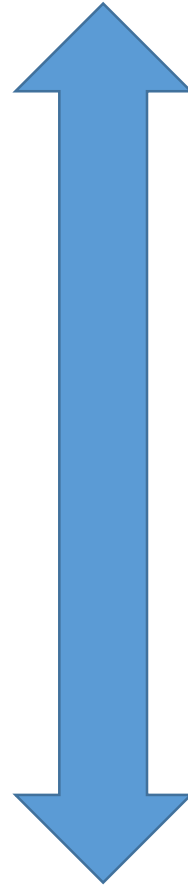
Olken, B. [Promises and perils of pre-analysis plans](#). Journal of Economic Perspectives: 29; 3. p61–80.

If S[HE]APs are the answer: 2) When

Before the project begins

At end of recruitment

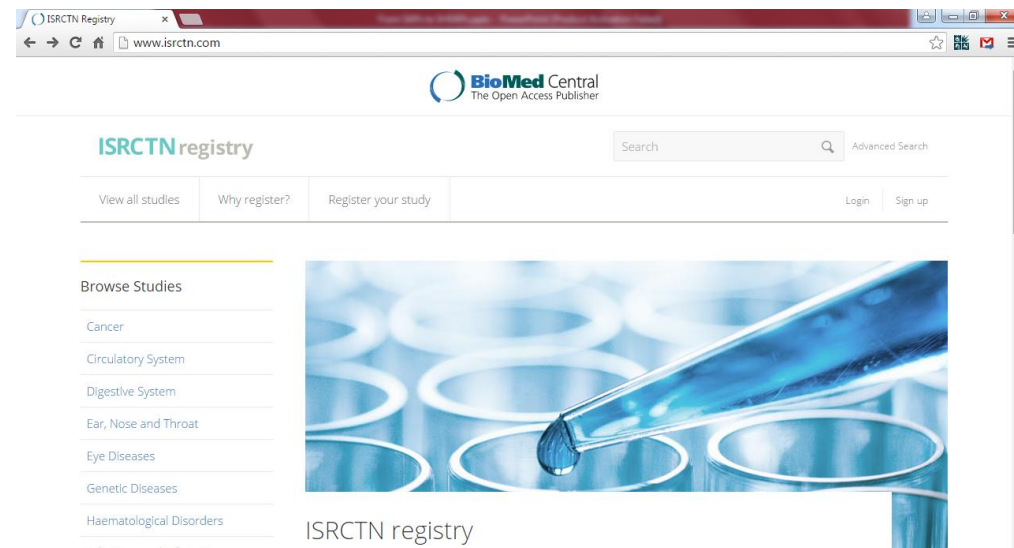
Before unblinding data



If S[HE]APs are the answer: 3) Where

Trial registries

- Time stamping
- Version control / Post-unblinding changes
- Intellectual property
- DMC / TSC oversight?

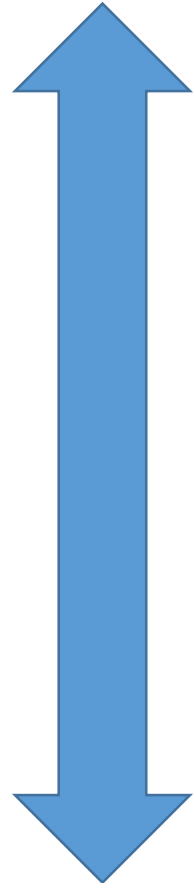


If S[HE]APs are the answer: 4) Level of detail

Vague waffle



Painted into a corner



S[HE]APs: The costs

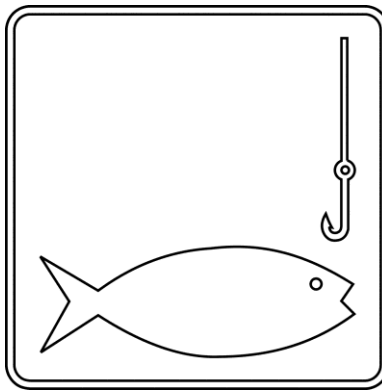
Complexity – researcher time

- Multiple arms
- Conditional analyses
- Desire to explore inside the black box
- Every TSC / DMC needs an economist?
- Blinded and unblinded research teams?

S[HE]APs: The costs

Danger that post-hoc analyses become unpublishable no matter how potentially enlightening or strong the evidence.

- Use SHEAPs as a working document to list and justify all post-hoc analyses done?

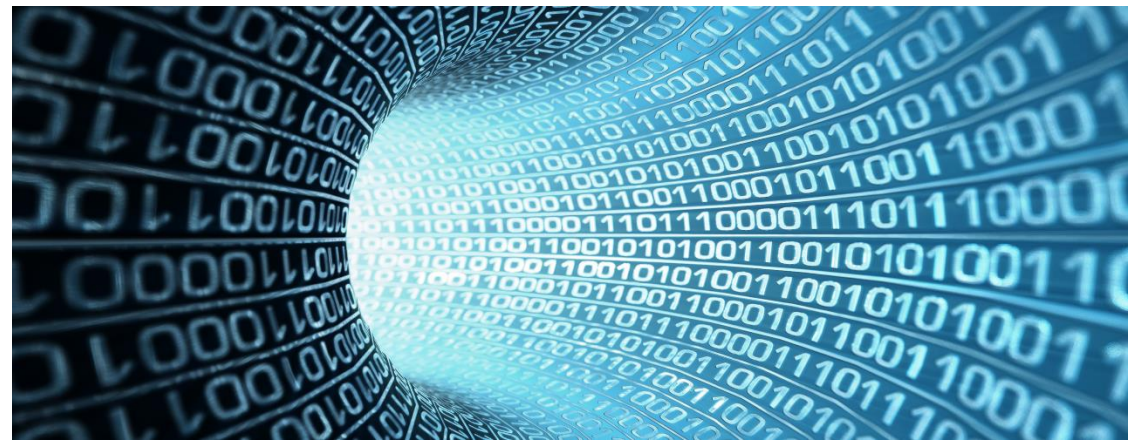


S[HE]APs: The fringe benefits

- Dress rehearsal [with fake IDs and randomisation?]
- Protection from overzealous research partners
- Protection from OTT reviewers

S[HE]APs: Alternatives suggested in mainstream economics

- Hypotheses registries
- Replication studies
- Open data / code



Some concluding thoughts

- SHEAPs not SAP + HEAP
- SHEAPs for public/private funded full-scale RCTs
- Streamlined, specific SHEAPs not exhaustive

Some concluding thoughts

- Publicly available before unblinding
- Doubt TSC overview needed
- Not slavish adherence – “Let the data speak”

Christopher Columbus: The SAP

Null hypothesis: Sailing west from Iberia to reach SE Asia takes as long as going east around Africa

Primary outcome: Time to reach SE Asia

Analytical method: Unpaired t-test

Olken, B. **Promises and perils of pre-analysis plans.** *Journal of Economic Perspectives*: 29; 3. p61–80.



Christopher Columbus: The report

Key finding: SE Asia not found; cannot reject the null hypothesis

Footnote: The observation that this route leads to a “new world” should be interpreted with caution and requires replication studies

