Systematic Reviews 101

An introduction to systematic reviews

Naomi Lee & Paul Whaley

23 May 2018

About me

- Digital Executive Editor at *The Lancet*, the leading independent general medical journal. Handles peer review and commissioning for *The Lancet* with a special interest in medical technology, health informatics, and surgery.
- Leading the digital transformation of *The Lancet* group, moving to a digital-first approach in all aspects of the business.
- Trained in surgery, specialising in urology and has worked in the United Kingdom, Argentina and Mexico. Studied medicine at Cambridge University and King’s College London.

What is a systematic review?

- A systematic review answers a pre-defined research question by collecting and summarizing all the evidence that fits into a pre-specified eligibility criteria
  - Can be qualitative or quantitative

- A meta-analysis uses statistical methods to summarize the results of the included studies
  - Is quantitative

Why are systematic reviews useful?

- To address unanswered questions without performing a new trial
- Efficient and ethical use of resources
- Can refine large amounts of information

What makes a good SR?

- Adds useful information
- Is registered e.g. with PROSPERO
- Has a clear search strategy
- Quality of evidence is important
- Follows reporting guidelines
- Puts the research into context

Reporting standards - PRISMA

- Guidelines defined by the EQUATOR network
- Preferred Reporting Items for Systematic Reviews and Meta-analyses
- Aims to improve the transparent reporting of systematic reviews and meta-analyses

PRISMA

TRANSPARENT REPORTING OF SYSTEMATIC REVIEWS AND META-ANALYSES

equator network

Cochrane

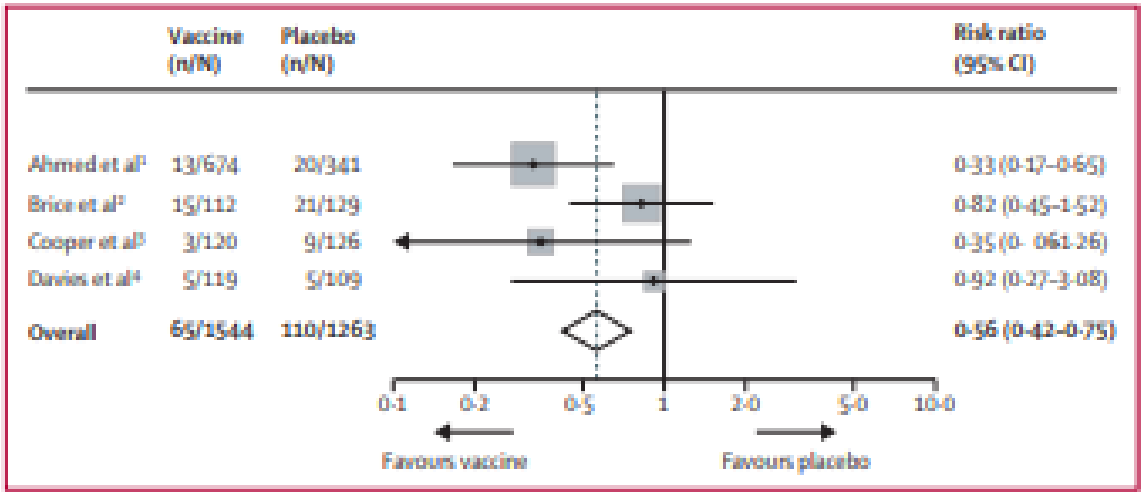
- Network of international collaborators
- Gather and summarize best health evidence to make informed health decisions
- Publishes systematic reviews, and provides support to authors
- Handbook for systematic reviews of Interventions

Meta-analysis

- Combines the results of the included studies to give an overall statistic

- Gives an estimate of the outcome of an intervention based on all the available evidence

- Results are summarized in a Forest plot



Example 1

Articles

Global epidemiology of yaws: a systematic review

Oriol Mitjà, Michael Marks, Diby J P Konan, Gilbert Ayelo, Camila Gonzalez-Beiras, Bernard Boua, Wendy Houinei, Yiragnima Kobara, Earnest N Tabah, Agana Nsiire, Damas Obvala, Fasiah Taleo, Rita Djupuri, Zhang Zaixing, Jürg Utzinger, Lasse S Vestergaard, Quique Bassat, Kingsley Asiedu

Summary

Background To achieve yaws eradication, the use of the new WHO strategy of initial mass treatment with azithromycin and surveillance twice a year needs to be extended everywhere the disease occurs. However, the geographic scope of the disease is unknown. We aimed to synthesise published and unpublished work to update the reported number of people with yaws at national and subnational levels and to estimate at-risk populations.

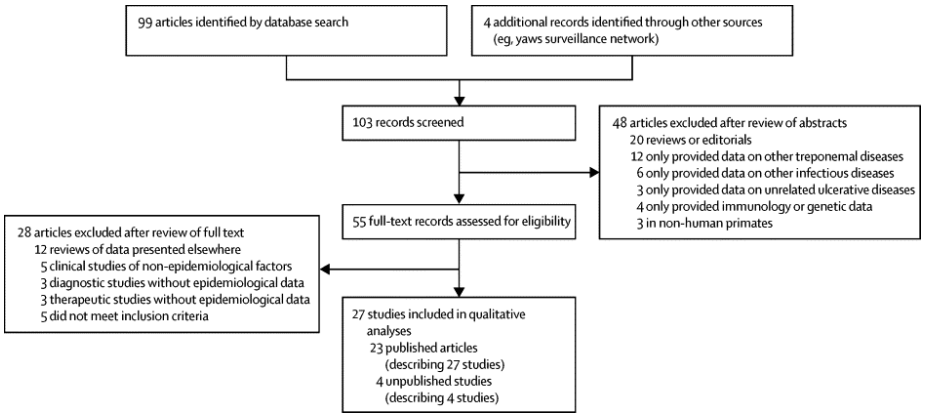
Methods We searched PubMed and WHO databases to identify published data for prevalence of active and latent yaws from Jan 1, 1990, to Dec 31, 2014. We also searched for ongoing or recently completed unpublished studies from the WHO yaws surveillance network. We estimated yaws prevalence (and 95% CIs). We collected yaws incidence data from official national surveillance programmes at the first administrative level from Jan 1, 2010, to Dec 31, 2013, and we used total population data at the second administrative level to estimate the size of at-risk populations.

Lancet Glob Health 2015; 3: e324–31

This online publication has been corrected. The corrected version first appeared at thelancet.com on June 8, 2015

See Comment page e300

Barcelona Institute for Global Health, Barcelona Centre for International Health Research, Hospital Clinic, University of Barcelona, Barcelona, Spain
(O Mitjà PhD,



Example 1

Characteristics and outcomes of the 24 included studies of active and latent yaws prevalence

	Year of study	Country	Location	Schoolchildren or community survey	Case ascertainment	Cases (sample size)	Prevalence, % (95% CI)
Africa							
Active yaws assessment							
Tabah et al (2012; Tabah EN, personal communication)	2012	Cameroon	Lomié, Zoubalot, Messok	Community	Clinical	97 (1075)	9·02 (7·38–10·90)
Herve et al (1992)[9]	1990	Central African Republic	Lobaye	School children	VDRL and TPHA	12 (213)	5·63 (2·94–9·63)
Boua et al (2012; Boua B, personal communication)	2012	Central African Republic	Lobaye, Sangha-Mbaeré	School children	Clinical	230 (2030)	11·33 (9·98–12·79)
Coldiron et al (2013)[10]	2012	Republic of Congo	Bétou, Ebyellé	Community	RDT	183 (6215)	2·94 (2·54–3·40)
Konan et al (2007)[33]	2004	Côte d'Ivoire	Adzopé	Community	RPR	11 (2182)	0·50 (0·25–0·90)
Gerstl et al (2009)[34]	2005	Democratic Republic of the Congo	Wasolo	Community	RPR and TPHA	56 (1176)	4·76 (3·62–6·14)
Nsiire et al (2011; Nsiire A, personal communication)	2011	Ghana	Volta Region	School c			
Akogun (1999)[35]	1998	Nigeria	Garkida	Commu			
Latent yaws assessment							
Ayelo et al (2012; Ayelo G, personal communication)	2012	Benin	Toffo, Zé, Allada	School c			
Herve et al (1992)[9]	1990	Central African Republic	Lobaye	School c			

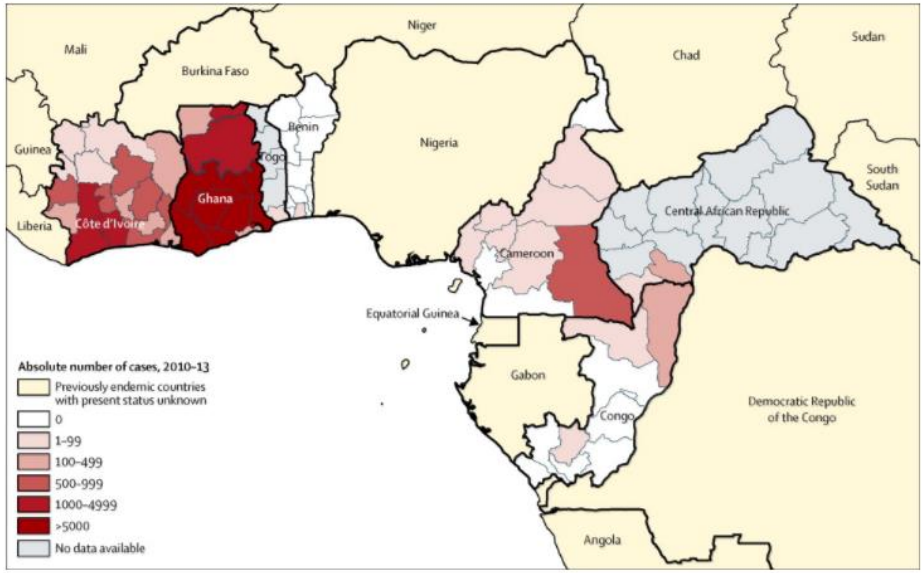


Figure 3

Cumulative number of yaws cases by subnational regions in the WHO Africa region

Example 2

Articles

Atraumatic versus conventional lumbar puncture needles: a systematic review and meta-analysis

Siddharth Nath, Alex Koziarz, Jetan H Badhiwala, Waleed Alhazzani, Roman Jaeschke, Sunjay Sharma, Laura Banfield, Ashkan Shoamanesh, Sheila Singh, Farshad Nassiri, Wieslaw Oczkowski, Emilie Belley-Côté, Ray Truant, Kesava Reddy, Maureen O Meade, Forough Farrokhyar, Malgorzata M Bala, Fayez Alshamsi, Mette Krag, Itziar Etxeandia-Ikobaltzeta, Regina Kunz, Osamu Nishida, Charles Matouk, Magdy Selim, Andrew Rhodes, Gregory Hawryluk, Saleh A Almenawer

Summary

Background Atraumatic needles have been proposed to lower complication rates after lumbar puncture. However, several surveys indicate that clinical adoption of these needles remains poor. We did a systematic review and meta-analysis to compare patient outcomes after lumbar puncture with atraumatic needles and conventional needles.

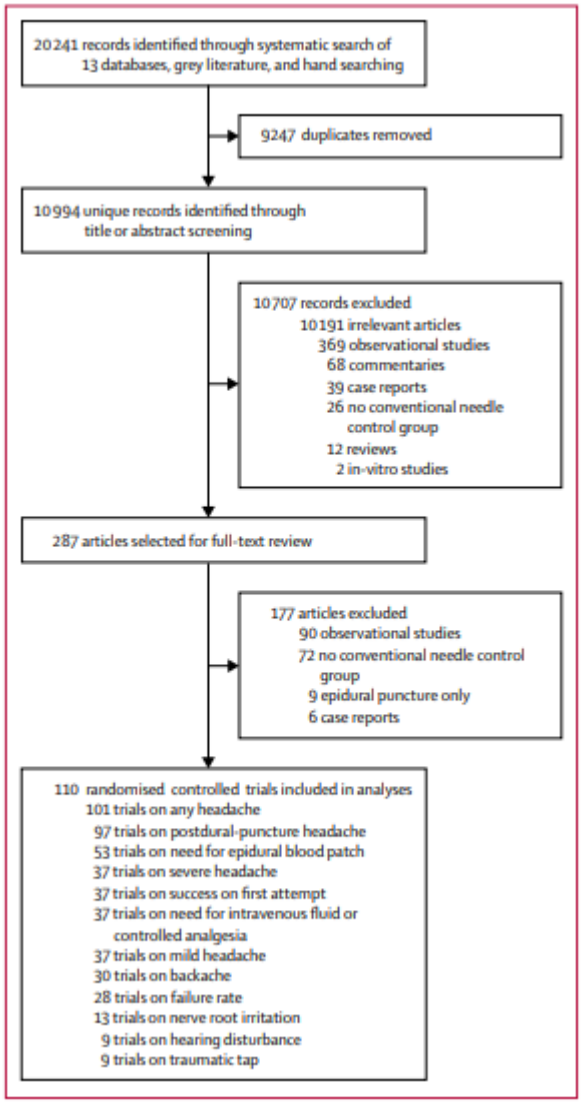
Methods In this systematic review and meta-analysis, we independently searched 13 databases with no language restrictions from inception to Aug 15, 2017, for randomised controlled trials comparing the use of atraumatic needles and conventional needles for any lumbar puncture indication. Randomised trials comparing atraumatic and conventional needles in which no dural puncture was done (epidural injections) or without a conventional needle control group were excluded. We screened studies and extracted data from published reports independently. The primary outcome of postdural-puncture headache incidence and additional safety and efficacy outcomes were assessed by random-effects and fixed-effects meta-analysis. This study is registered with the International Prospective Register of Systematic Reviews, number CRD42016047546.

Lancet 2018; 391: 1197–204

Published Online December 6, 2017 http://dx.doi.org/10.1016/S0140-6736(17)32451-0

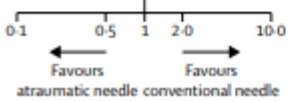
See Comment page 1128

Division of Neurosurgery (S Nath BSc, A Koziarz BSc, S Sharma MD, S Singh MD, Prof K Reddy MBBS, S A Almenawer MD), **Division of Critical Care** (W Alhazzani MD, Prof R Jaeschke MD, E Belley-Côté MD,



Example 2

	Atraumatic needle (n/N)	Conventional needle (n/N)	RR (95% CI)	p	I^2 (%)
Postdural-puncture headache	494/12358	1228/12543	0·40 (0·34–0·47)	<0·0001	45·4
Any headache	851/12649	1666/12900	0·50 (0·43–0·57)	<0·0001	54·9
Mild headache	78/3784	166/3740	0·52 (0·38–0·70)	<0·0001	10·5
Severe headache	32/2651	106/2527	0·41 (0·28–0·59)	<0·0001	0
Need for intravenous fluid or controlled analgesia	60/3652	138/3531	0·44 (0·29–0·64)	<0·0001	24·0
Need for epidural blood patch	38/3770	77/3168	0·50 (0·33–0·75)	0·001	0
Nerve root irritation	65/707	99/789	0·71 (0·54–0·92)	0·011	0
Hearing disturbance	3/531	30/568	0·25 (0·11–0·60)	0·002	0
Traumatic tap	41/755	54/830	0·87 (0·53–1·42)	0·57	18·5
Backache	393/2567	475/2864	0·96 (0·84–1·17)	0·656	31·7
Success on first attempt	3757/4365	3871/4417	0·99 (0·96–1·02)	0·484	47·8
Failure rate	94/2974	97/2540	0·86 (0·58–1·27)	0·44	31·7



ELSEVIER

Raising the standard of published systematic reviews

A case study from chemical risk research

Paul Whaley

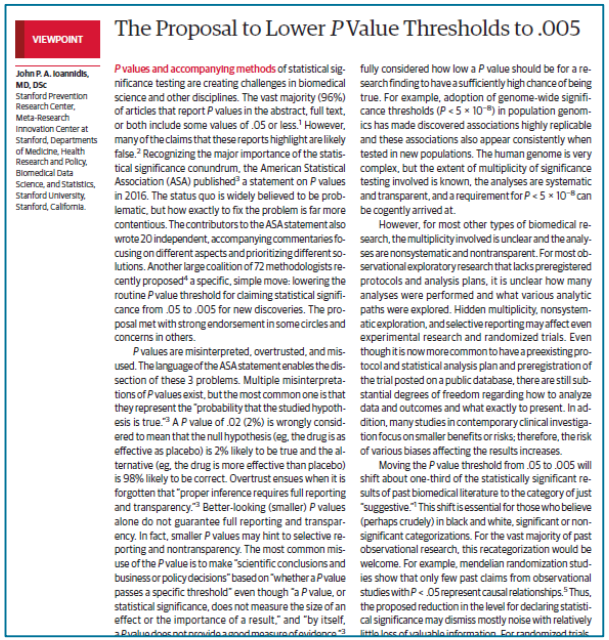
About me

- Researcher at Lancaster University and the Evidence-Based Toxicology Collaboration at Johns Hopkins BSPH
- Background in environmental health advocacy and science communication
- Introduced to systematic reviews as gold-standard approach to evidence synthesis in early 2010
- Associate Editor for Systematic Reviews at *Environment International* (IF 7.088) – first specialist EH SR editor
- The “frameworks guy”: systematic approaches to evidence surveillance and synthesis; critical appraisal tools; codes of practice; research quality management

Today's presentation

- Reproducibility issues in chemical risk assessment as a driver of interest in systematic review methods
- Uptake of SR methods
- Challenges we are seeing (poor quality SRs)
- How we are addressing these challenges at *Environment International*
- Implications for you as potential submitting authors and conductors of systematic reviews

A “reproducibility crisis” in primary research



VIEWPOINT

The Proposal to Lower *P* Value Thresholds to .005

John P. A. Ioannidis, MD, DSc
Stanford Prevention Research Center, Meta-Research Innovation Center at Stanford, Departments of Medicine, Health Research and Policy, Biomedical Data Science, and Statistics, Stanford University, Stanford, California.

P values and accompanying methods of statistical significance testing are creating challenges in biomedical science and other disciplines. The vast majority (96%) of articles that report *P* values in the abstract, full text, or both include some values of .05 or less.[1] However, many of the claims that these reports highlight are likely false.[2] Recognizing the major importance of the statistical significance conundrum, the American Statistical Association (ASA) published[3] a statement on *P* values in 2016. The status quo is widely believed to be problematic, but how exactly to fix the problem is far more contentious. The contributors to the ASA statement also wrote 20 independent, accompanying commentaries focusing on different aspects and prioritizing different solutions. Another large coalition of 72 methodologists recently proposed[4] a specific, simple move: lowering the routine *P* value threshold for claiming statistical significance from .05 to .005 for new discoveries. The proposal met with strong endorsement in some circles and concerns in others.

P values are misinterpreted, overtrusted, and misused. The language of the ASA statement enables the dissection of these 3 problems. Multiple misinterpretations of *P* values exist, but the most common one is that they represent the “probability that the studied hypothesis is true.”[3] A *P* value of .02 (2%) is wrongly considered to mean that the null hypothesis (eg, the drug is as effective as placebo) is 2% likely to be true and the alternative (eg, the drug is more effective than placebo) is 98% likely to be correct. Overtrust ensues when it is forgotten that “proper inference requires full reporting and transparency.”[3] Better-looking (smaller) *P* values alone do not guarantee full reporting and transparency. In fact, smaller *P* values may hint to selective reporting and nontransparency. The most common misuse of the *P* value is to make “scientific conclusions and business or policy decisions” based on “whether a *P* value passes a specific threshold” even though “a *P* value, or statistical significance, does not measure the size of an effect or the importance of a result,” and “by itself, a *P* value does not provide a good measure of evidence.”[3] fully considered how low a *P* value should be for a research finding to have a sufficiently high chance of being true. For example, adoption of genome-wide significance thresholds ($P < 5 \times 10^{-8}$) in population genomics has made discovered associations highly replicable and these associations also appear consistently when tested in new populations. The human genome is very complex, but the extent of multiplicity of significance testing involved is known, the analyses are systematic and transparent, and a requirement for $P < 5 \times 10^{-8}$ can be cogently arrived at.

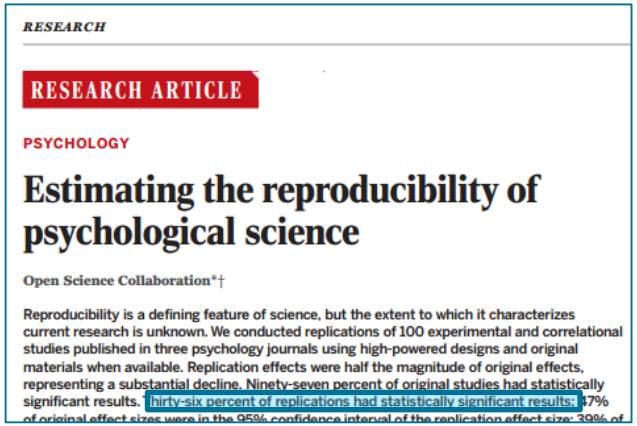
However, for most other types of biomedical research, the multiplicity involved is unclear and the analyses are nonsystematic and nontransparent. For most observational exploratory research that lacks preregistered protocols and analysis plans, it is unclear how many analyses were performed and what various analytic paths were explored. Hidden multiplicity, nonsystematic exploration, and selective reporting may affect even experimental research and randomized trials. Even though it is now more common to have a preexisting protocol and statistical analysis plan and preregistration of the trial posted on a public database, there are still substantial degrees of freedom regarding how to analyze data and outcomes and what exactly to present. In addition, many studies in contemporary clinical investigation focus on smaller benefits or risks; therefore, the risk of various biases affecting the results increases.

Moving the *P* value threshold from .05 to .005 will shift about one-third of the statistically significant results of past biomedical literature to the category of just “suggestive.”[5] This shift is essential for those who believe (perhaps crudely) in black and white, significant or nonsignificant categorizations. For the vast majority of past observational research, this recategorization would be welcome. For example, mendelian randomization studies show that only few past claims from observational studies with $P < .05$ represent causal relationships.[5] Thus, the proposed reduction in the level for declaring statistical significance may dismiss mostly noise with relatively

REPRODUCIBILITY PROJECT
Cancer Biology

The Reproducibility Project: Cancer Biology is a collaboration between the Center for Open Science and Science Exchange to independently replicate selected results from a substantial

EDITORIAL



RESEARCH

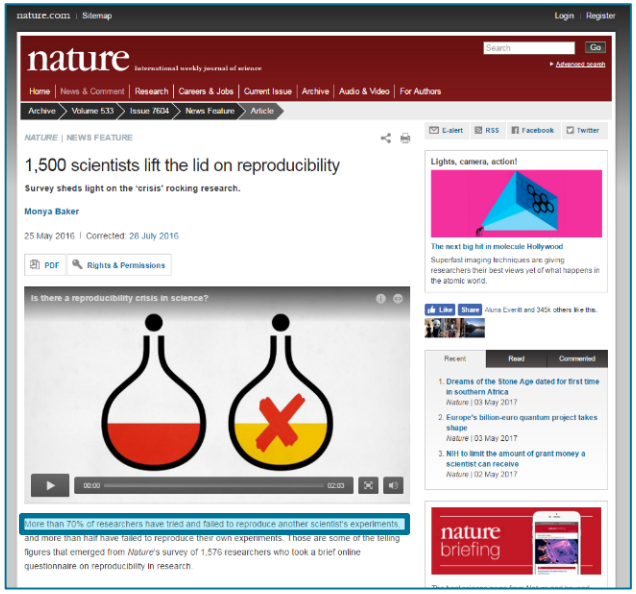
RESEARCH ARTICLE

PSYCHOLOGY

Estimating the reproducibility of psychological science

Open Science Collaboration*†

Reproducibility is a defining feature of science, but the extent to which it characterizes current research is unknown. We conducted replications of 100 experimental and correlational studies published in three psychology journals using high-powered designs and original materials when available. Replication effects were half the magnitude of original effects, representing a substantial decline. Ninety-seven percent of original studies had statistically significant results. Thirty-six percent of replications had statistically significant results; 47% of original effect sizes were in the 95% confidence interval of the replication effect size; 39% of



nature.com | Sitemap

Login | Register

nature
international weekly journal of science

Search | Go | Advanced search

Home | News & Comment | Research | Careers & Jobs | Current Issue | Archive | Audio & Video | For Authors

Archive > Volume 533 > Issue 7604 > News Feature > Article

NATURE | NEWS FEATURE

E-alert | RSS | Facebook | Twitter

1,500 scientists lift the lid on reproducibility

Survey sheds light on the ‘crisis’ rocking research.

Monya Baker

25 May 2016 | Corrected: 28 July 2016

PDF | Rights & Permissions

Is there a reproducibility crisis in science?

More than 70% of researchers have tried and failed to reproduce another scientist’s experiments, and more than half have failed to reproduce their own experiments. Those are some of the telling figures that emerged from *Nature*’s survey of 1,576 researchers who took a brief online questionnaire on reproducibility in research.

Lights, camera, action!

The next big hit in molecule Hollywood

Superfast imaging techniques are giving researchers their best views yet of what happens in the atomic world.

Like | Share | Nuna Everitt and 345k others like this.

Recent | Read | Commented

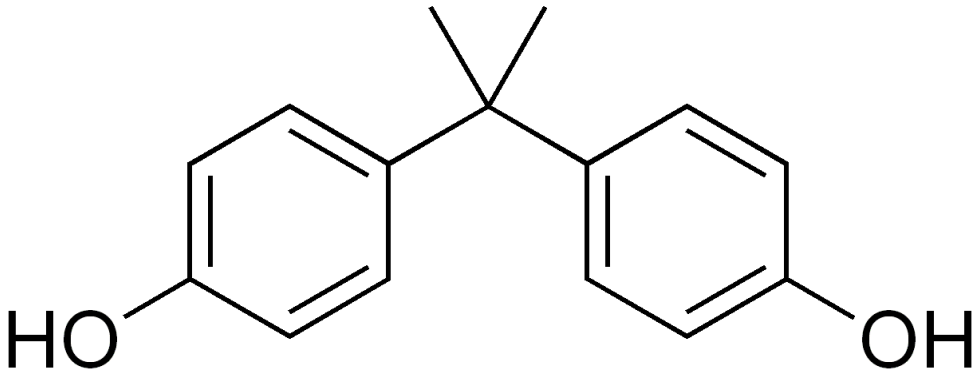
1. Dreams of the Stone Age dated for first time in southern Africa
Nature | 03 May 2017
2. Europe’s billion-euro quantum project takes shape
Nature | 03 May 2017
3. NIH to limit the amount of grant money a scientist can receive
Nature | 02 May 2017

nature briefing

Chemical risk assessment

- Making sense of complex and contradictory evidence about health risks posed by exposure to chemical substances

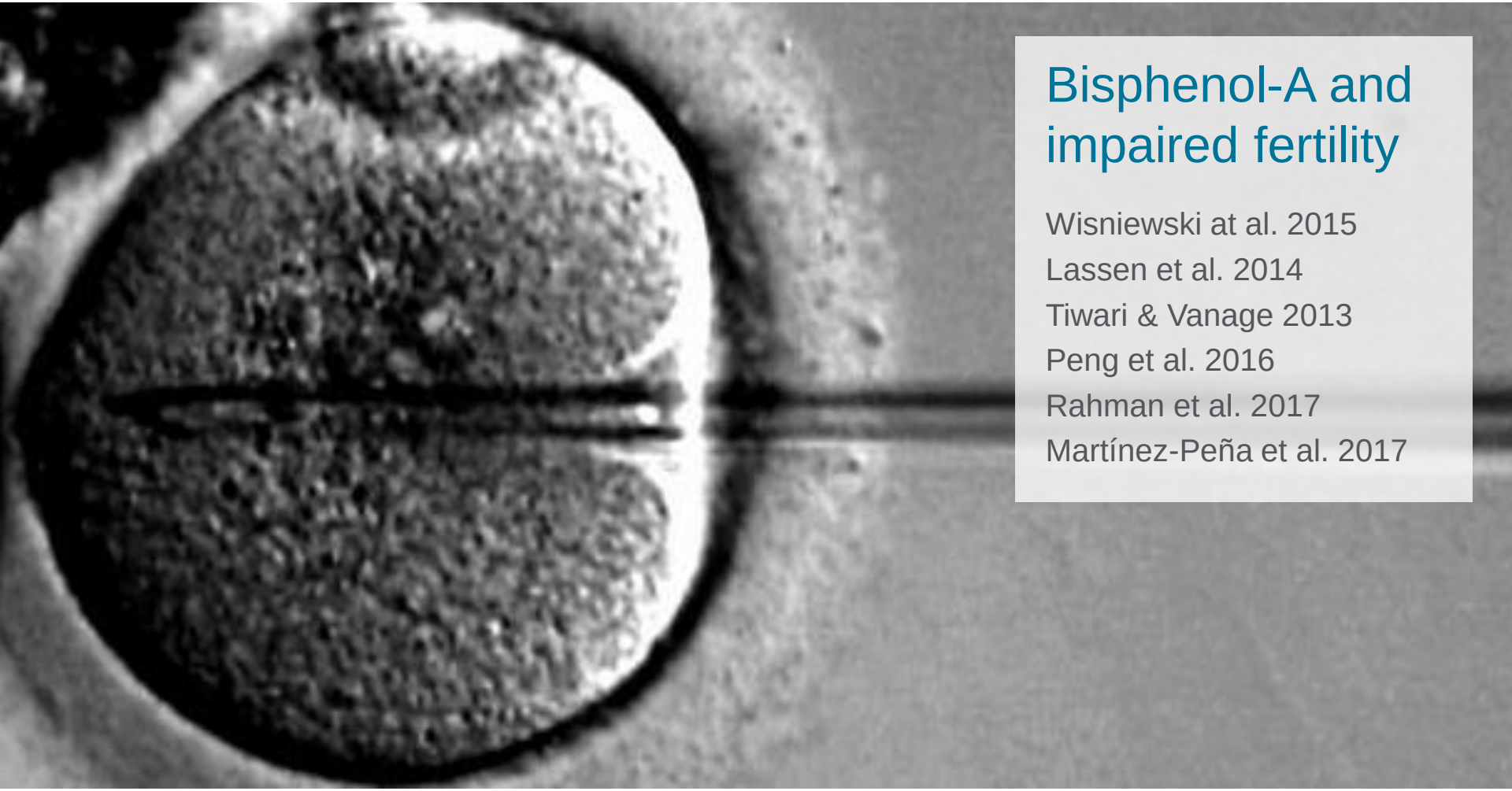
Reproducibility crisis in chemical risk assessment



Bisphenol-A

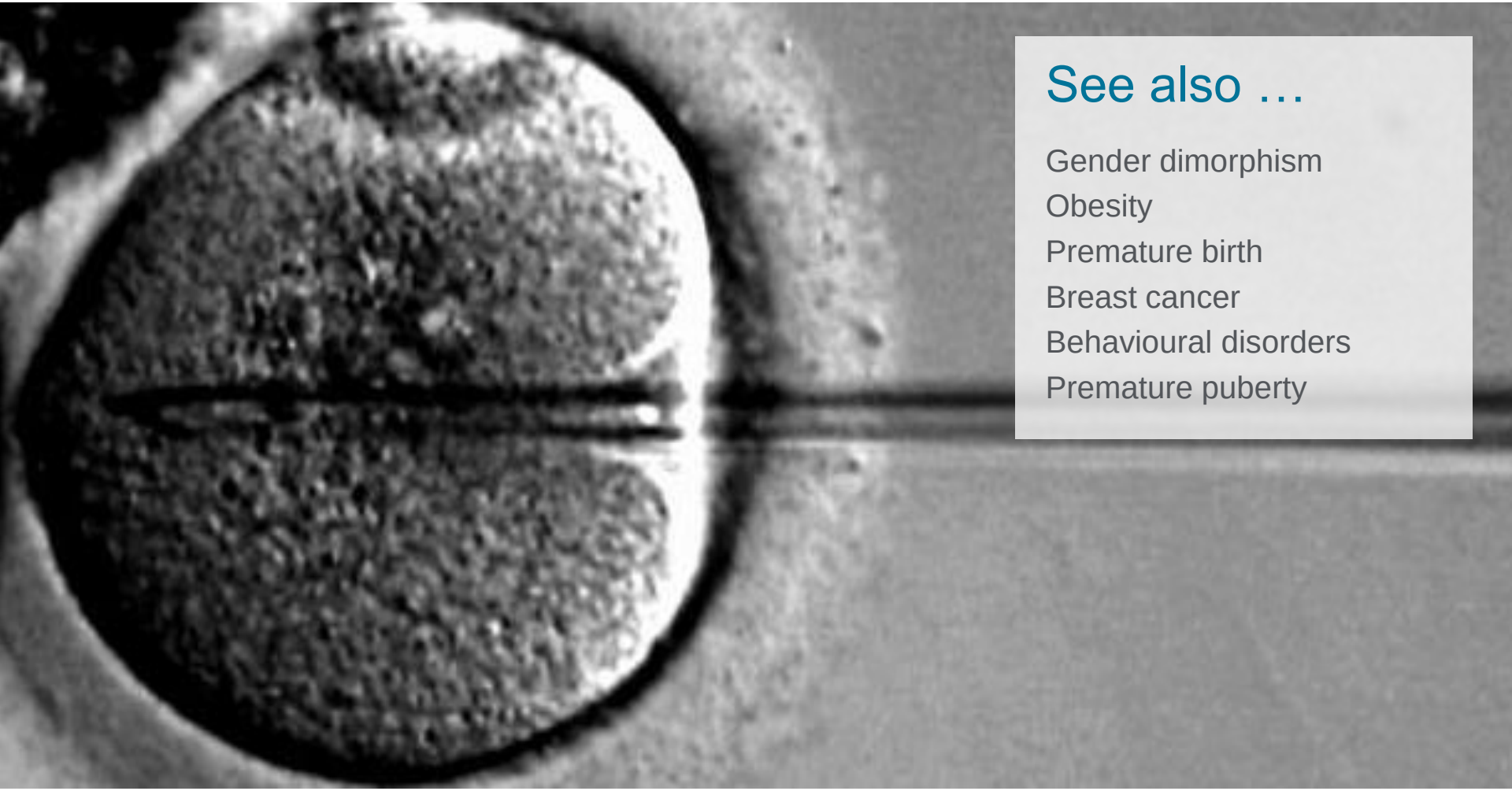

Shark Fin Soup
Ligo
SQUID


@KoenigJ: Im Foyer ist es so kalt, dass die Empf
angsdamen durch Roboter ersetzt wurden. sic! (08
:58:24)
@bommel/Jog1: die Zeit ist mir wieder einmal wei
voraus. (08:22:08)
#LARRR: Habe jetzt gelernt einen Text in me
Muttersprache zu verfassen. (07:57:28)
#LARRR: Gruß an den Mann, der gerade
Phone gafft. (07:08:19)
POWER



Bisphenol-A and impaired fertility

Wisniewski at al. 2015
Lassen et al. 2014
Tiwari & Vanage 2013
Peng et al. 2016
Rahman et al. 2017
Martínez-Peña et al. 2017



See also …

Gender dimorphism
Obesity
Premature birth
Breast cancer
Behavioural disorders
Premature puberty

efsa
European Food Safety Authority

KAROLINSKA INSTITUTET
ANNO 1810
Karolinska Institutet

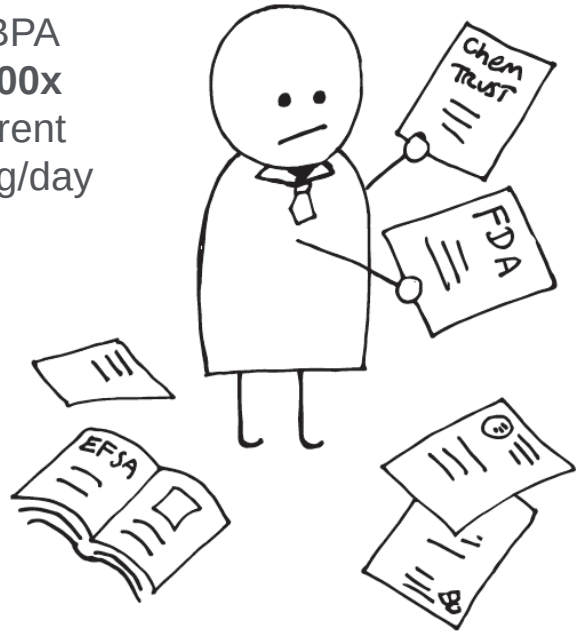
Public Health England

International Agency Research on Cancer

World Health Organization

FDA

…effects have been demonstrated for BPA [at] levels **10–10,000x lower** than the current LOAEL of 50 mg/kg/day
Vandenberg et al. 2014

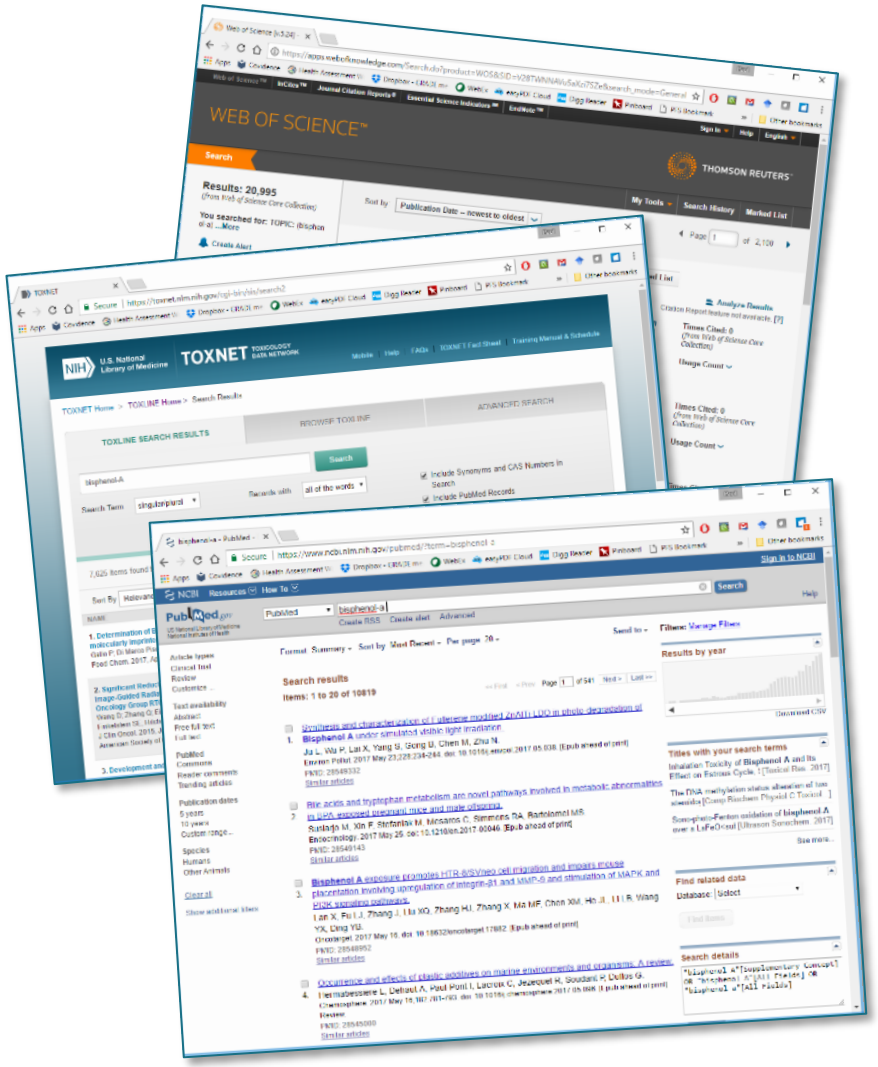


…**no health concern** for any age group from dietary exposure
EFSA 2015

…a TDI for BPA has to be **0.7 µg/kg bw/day** or lower to be sufficiently protective
National Food Institute, Denmark 2015

…a **potential risk to the unborn children** of exposed pregnant women [relating to] a change in the structure of the mammary gland
ANSES 2013

Same evidence, different conclusions



?

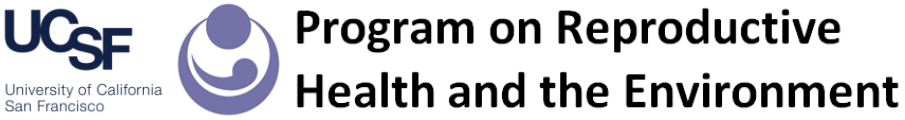
…**no health concern** for any age group from dietary exposure

…a TDI for BPA has to be **0.7 µg/kg bw/day** or lower to be sufficiently protective
National Food Institute

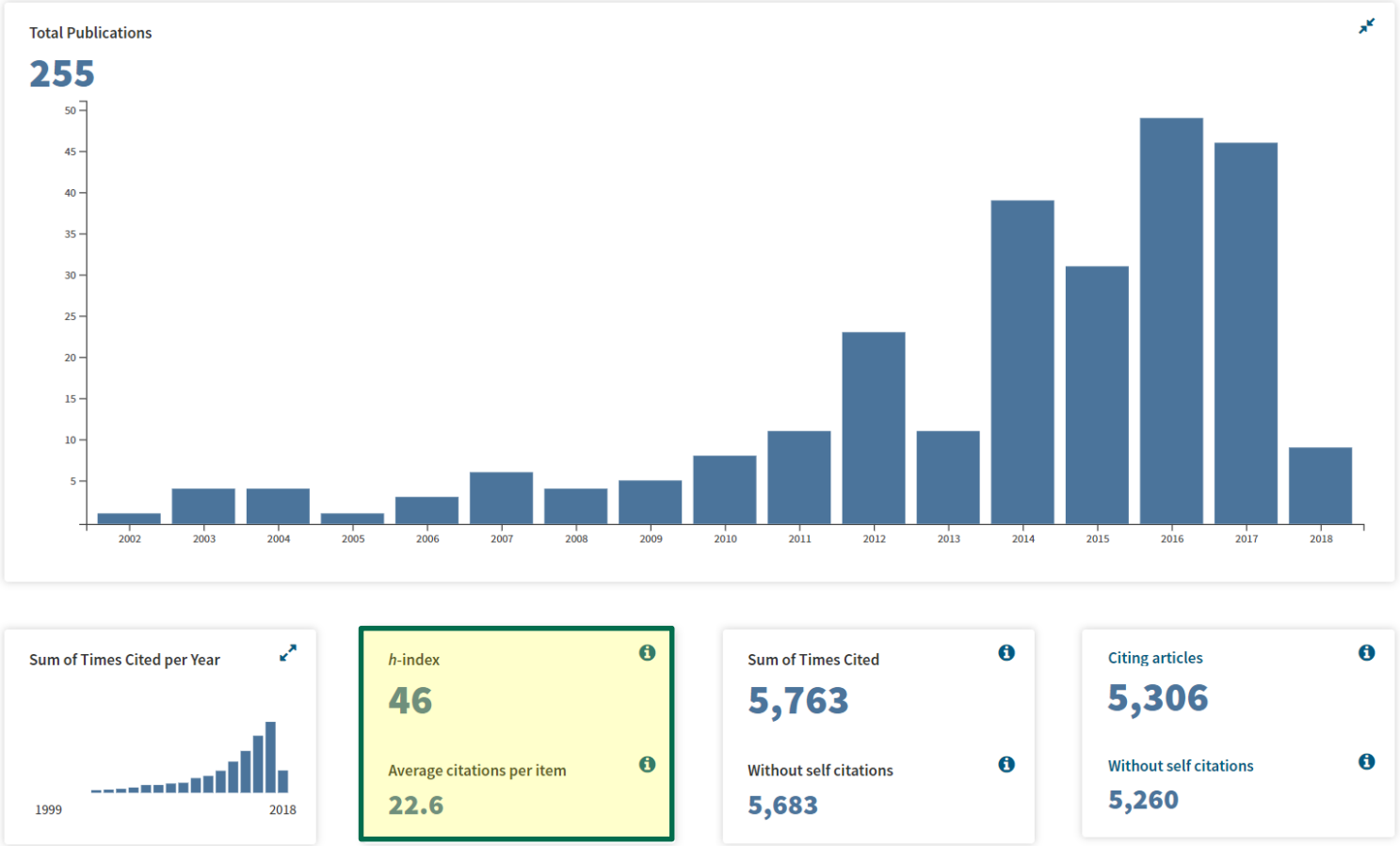
…effects have been demonstrated for BPA [at] levels **10–10,000x lower** than the current LOAEL of 50 mg/kg/day
Vandenberg et al. 2014

Solving the problem with systematic review methods

- Accelerating uptake since I started working on this in 2010



Rapid growth in publication of SRs



TITLE: ("systematic review"); Refined by: WEB OF SCIENCE CATEGORIES: (TOXICOLOGY) AND [excluding] WEB OF SCIENCE CATEGORIES: (PHARMACOLOGY PHARMACY); Timespan: All years. Indexes: SCI-EXPANDED, SSCI, A&HCI, CPCI-S, CPCI-SSH, ESCI, IC.

But we have a problem with quality

- 8989 PubMed records tagged by 2004 as “systematic review” yet actual number of stringently-defined SRs was ~2500 (Moher et al. 2007)
- Most published SRs have major flaws in conduct and reporting (Page et al. 2016)
- ~3% of manuscripts are “decent and clinically useful” (Ioannidis 2016)
- Our own pilot data shows serious omissions in reporting of 19 of 25 SRs published in the top environmental health journals through 2014-2015, before we even look at the validity of the actual methods used
- Fundamental errors mean a lot of effort is being put into projects which are not fit for purpose

My job as an editor

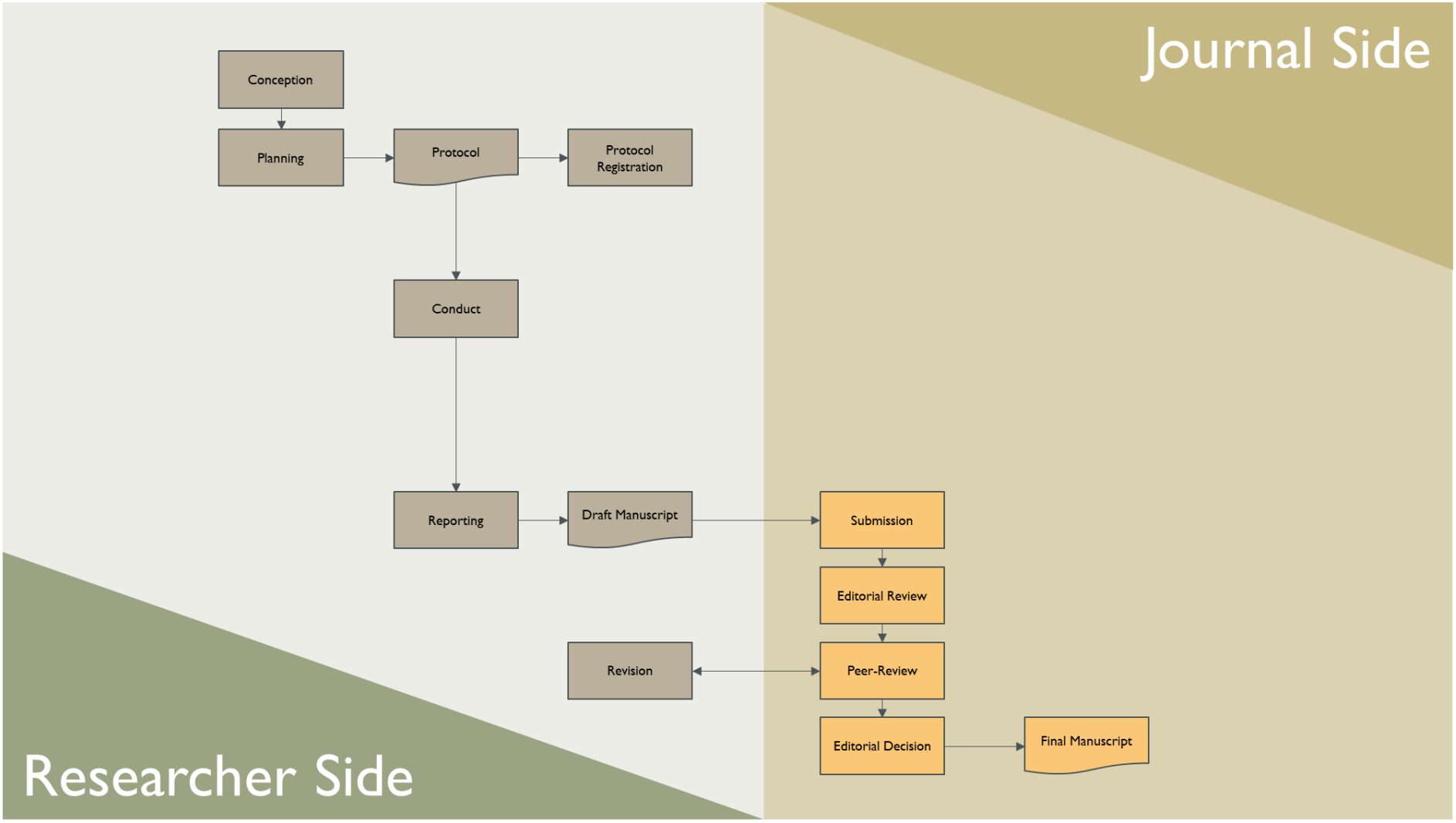
- What can I do at our journal to ensure each SR we publish is fit for purpose?
  - Asks an important question
  - Is truthful
  - Includes all information about methods and results, such that a reader can appraise the validity of the SR's findings and assess its relevance to their decision-making context
- Gatekeeper and midwife strategies for ensuring we publish high-quality research
- Implications for you as researchers

EDITOR AS GATEKEEPER

Enforcement of reporting standards

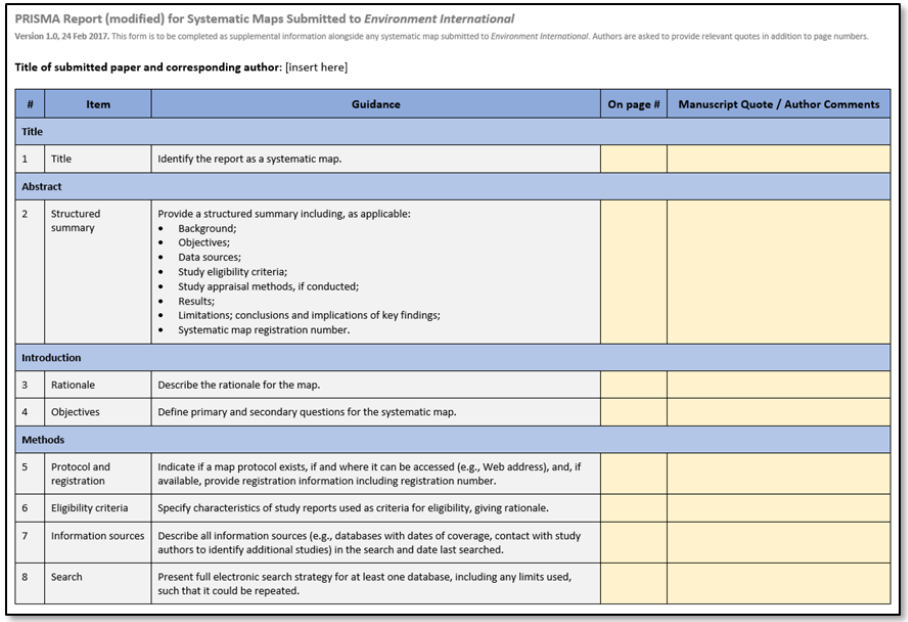
Editorial triage

Making best use of peer-review


Journal Side
Conception
Planning
Protocol
Protocol Registration
Conduct
Reporting
Draft Manuscript
Submission
Editorial Review
Revision
Peer-Review
Editorial Decision
Final Manuscript
Researcher Side

Enforcement of reporting standards

- Option of PRISMA (Moher et al. 2009) or ROSES (Haddaway et al. 2018)
- Submission of PRISMA or ROSES report as supplemental information is compulsory
- Useful quick check on basic standards

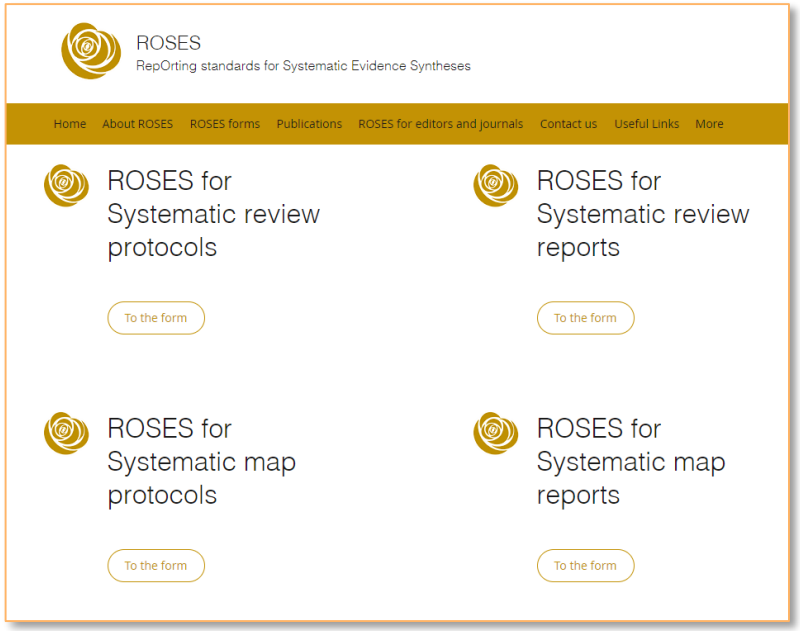


PRISMA Report (modified) for Systematic Maps Submitted to *Environment International*

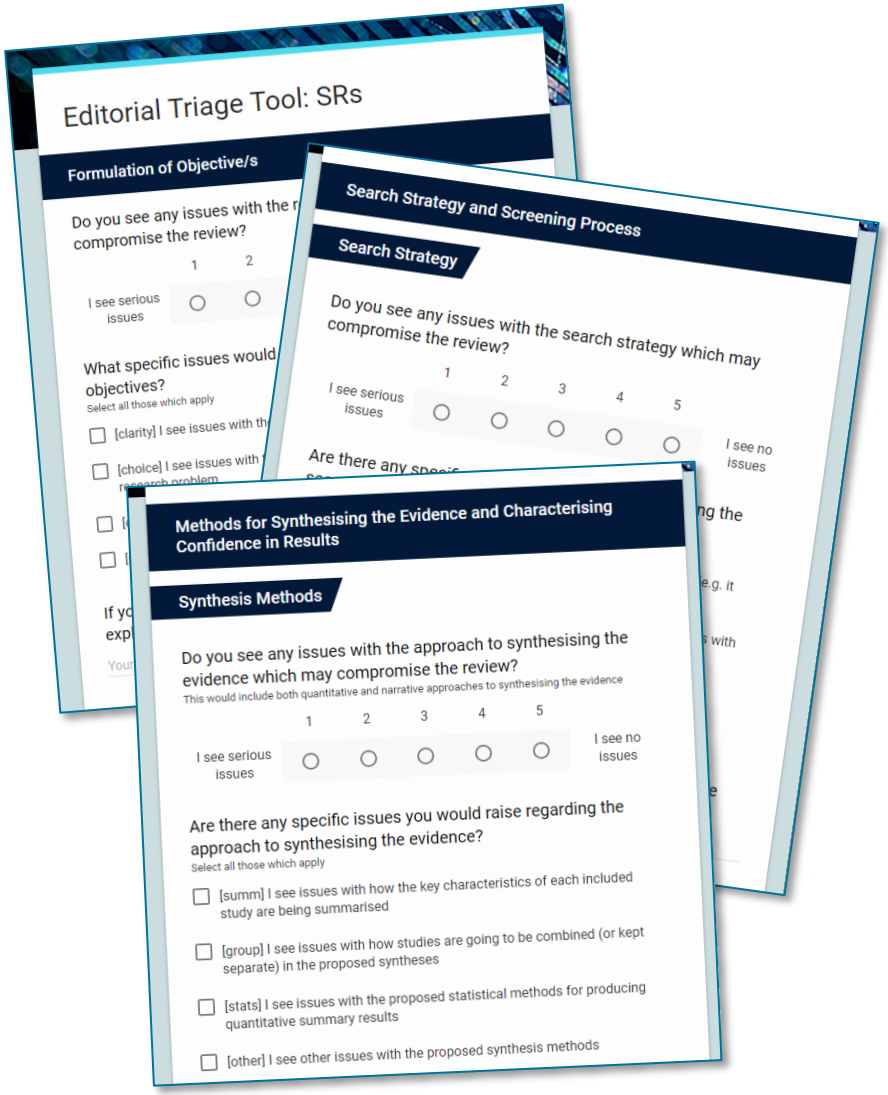
Version 1.0, 24 Feb 2017. This form is to be completed as supplemental information alongside any systematic map submitted to *Environment International*. Authors are asked to provide relevant quotes in addition to page numbers.

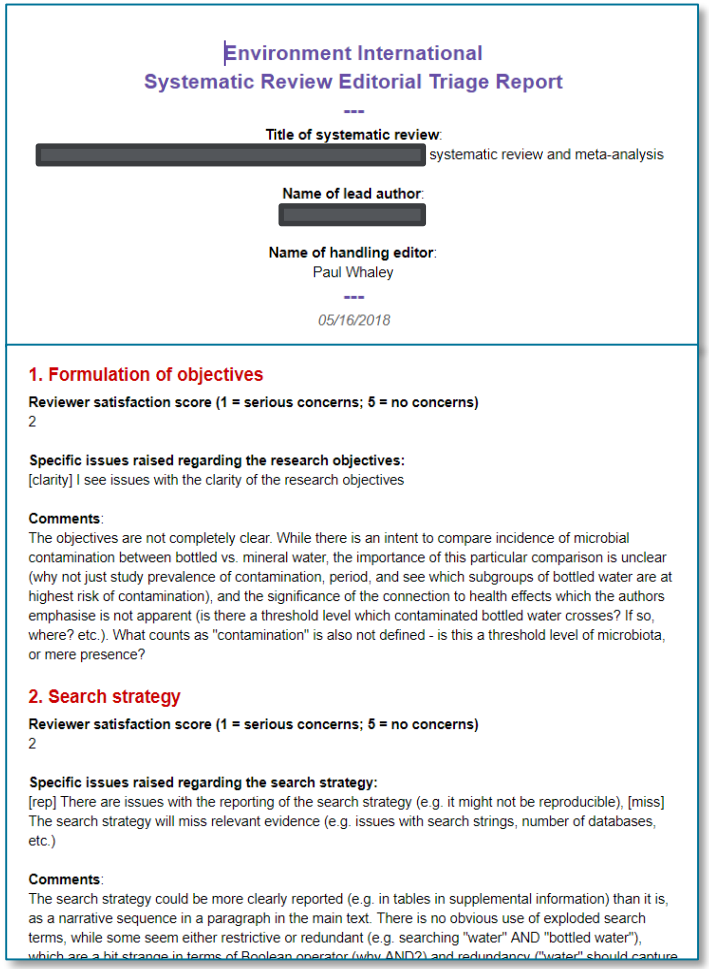
Title of submitted paper and corresponding author: [insert here]

#	Item	Guidance	On page #	Manuscript Quote / Author Comments
Title				
1	Title	Identify the report as a systematic map.		
Abstract				
2	Structured summary	Provide a structured summary including, as applicable: • Background; • Objectives; • Data sources; • Study eligibility criteria; • Study appraisal methods, if conducted; • Results; • Limitations; conclusions and implications of key findings; • Systematic map registration number.		
Introduction				
3	Rationale	Describe the rationale for the map.		
4	Objectives	Define primary and secondary questions for the systematic map.		
Methods				
5	Protocol and registration	Indicate if a map protocol exists, if and where it can be accessed (e.g., Web address), and, if available, provide registration information including registration number.		
6	Eligibility criteria	Specify characteristics of study reports used as criteria for eligibility, giving rationale.		
7	Information sources	Describe all information sources (e.g., databases with dates of coverage, contact with study authors to identify additional studies) in the search and date last searched.		
8	Search	Present full electronic search strategy for at least one database, including any limits used, such that it could be repeated.		



Editorial triage reports





Environment International
Systematic Review Editorial Triage Report

Title of systematic review:
[redacted] systematic review and meta-analysis

Name of lead author:
[redacted]

Name of handling editor:
Paul Whaley

05/16/2018

1. Formulation of objectives

Reviewer satisfaction score (1 = serious concerns; 5 = no concerns)
2

Specific issues raised regarding the research objectives:
[clarity] I see issues with the clarity of the research objectives

Comments:
The objectives are not completely clear. While there is an intent to compare incidence of microbial contamination between bottled vs. mineral water, the importance of this particular comparison is unclear (why not just study prevalence of contamination, period, and see which subgroups of bottled water are at highest risk of contamination), and the significance of the connection to health effects which the authors emphasise is not apparent (is there a threshold level which contaminated bottled water crosses? If so, where? etc.). What counts as "contamination" is also not defined - is this a threshold level of microbiota, or mere presence?

2. Search strategy

Reviewer satisfaction score (1 = serious concerns; 5 = no concerns)
2

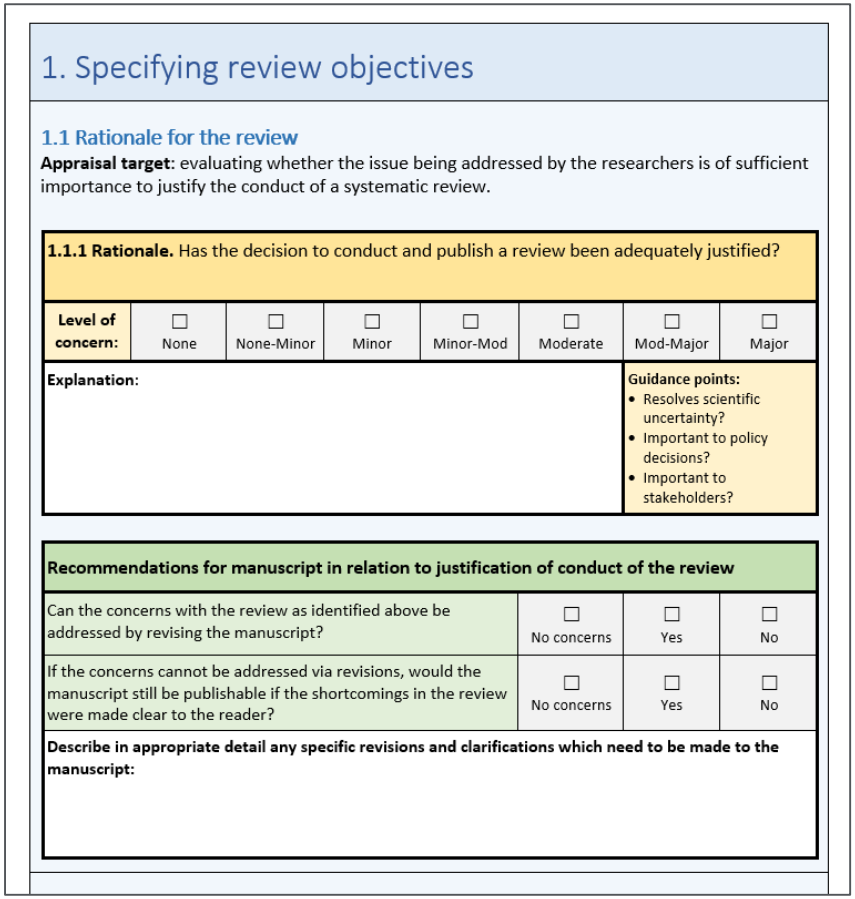
Specific issues raised regarding the search strategy:
[rep] There are issues with the reporting of the search strategy (e.g. it might not be reproducible), [miss] The search strategy will miss relevant evidence (e.g. issues with search strings, number of databases, etc.)

Comments:
The search strategy could be more clearly reported (e.g. in tables in supplemental information) than it is, as a narrative sequence in a paragraph in the main text. There is no obvious use of exploded search terms, while some seem either restrictive or redundant (e.g. searching "water" AND "bottled water"), which are a bit strange in terms of Boolean operator (why AND?) and redundancy ("water" should capture

Improved peer-review

- Target of 4 reviewers per submission
  - 2 topic experts
  - 2 methods experts
- Peer-review facilitation tool
  - Testing a Google Forms tool similar to Triage tool
  - Building CREST-SR for full-blooded implementation

Whaley et al. “A Tool for Critical Appraisal of Evidence Syntheses in Toxicology: Systematic Reviews (CREST-SR)” Under development



1. Specifying review objectives

1.1 Rationale for the review

Appraisal target: evaluating whether the issue being addressed by the researchers is of sufficient importance to justify the conduct of a systematic review.

1.1.1 Rationale. Has the decision to conduct and publish a review been adequately justified?

Level of concern:	☐ None	☐ None-Minor	☐ Minor	☐ Minor-Mod	☐ Moderate	☐ Mod-Major	☐ Major

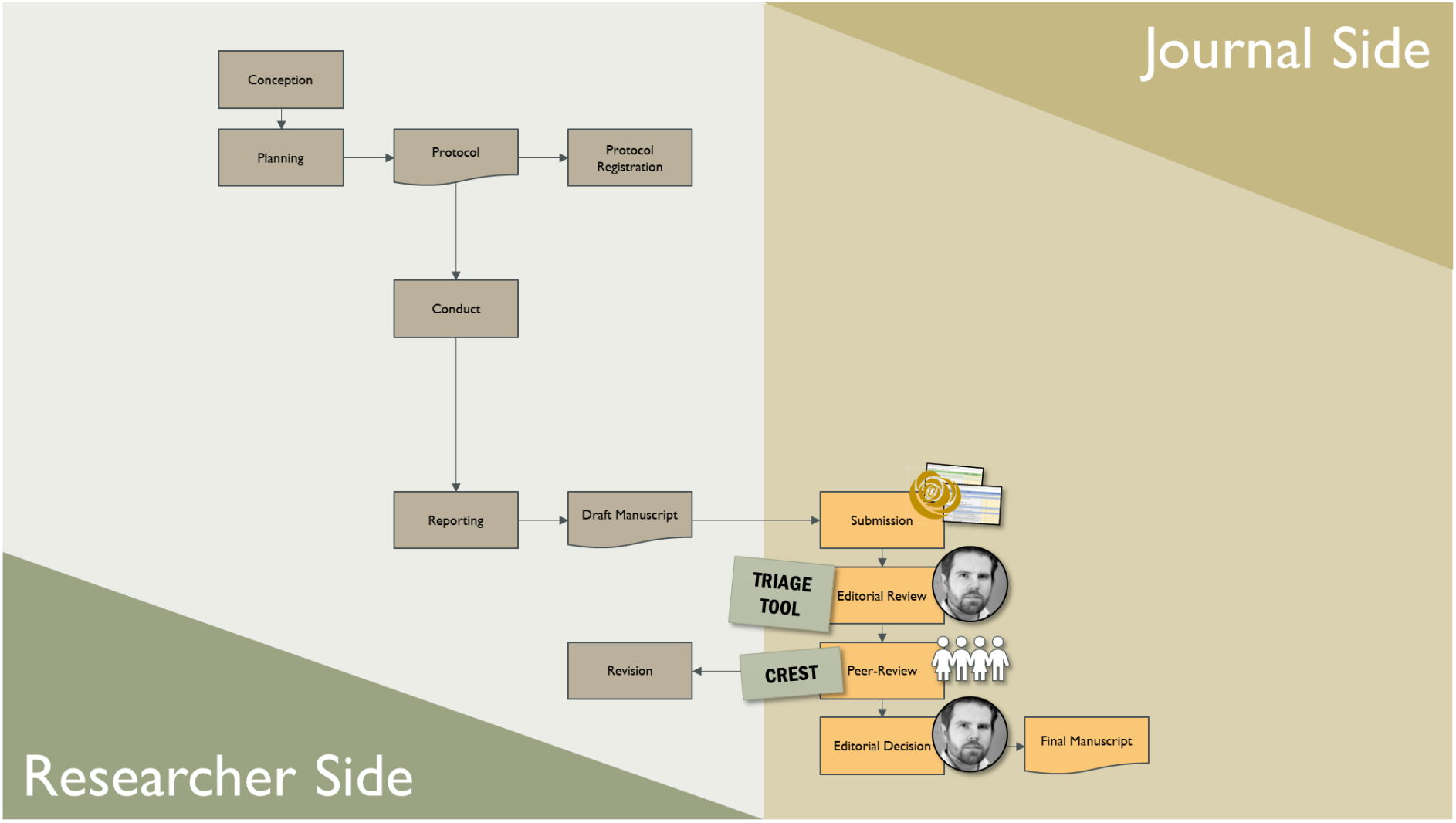
Explanation:

Guidance points:
- Resolves scientific uncertainty?
- Important to policy decisions?
- Important to stakeholders?

Recommendations for manuscript in relation to justification of conduct of the review

Can the concerns with the review as identified above be addressed by revising the manuscript?	☐ No concerns	☐ Yes	☐ No
If the concerns cannot be addressed via revisions, would the manuscript still be publishable if the shortcomings in the review were made clear to the reader?	☐ No concerns	☐ Yes	☐ No

Describe in appropriate detail any specific revisions and clarifications which need to be made to the manuscript:


Journal Side
Researcher Side
Conception
Planning
Protocol
Protocol Registration
Conduct
Reporting
Draft Manuscript
Submission
TRIAGE TOOL
Editorial Review
CREST
Peer-Review
Revision
Editorial Decision
Final Manuscript

Progress so far?

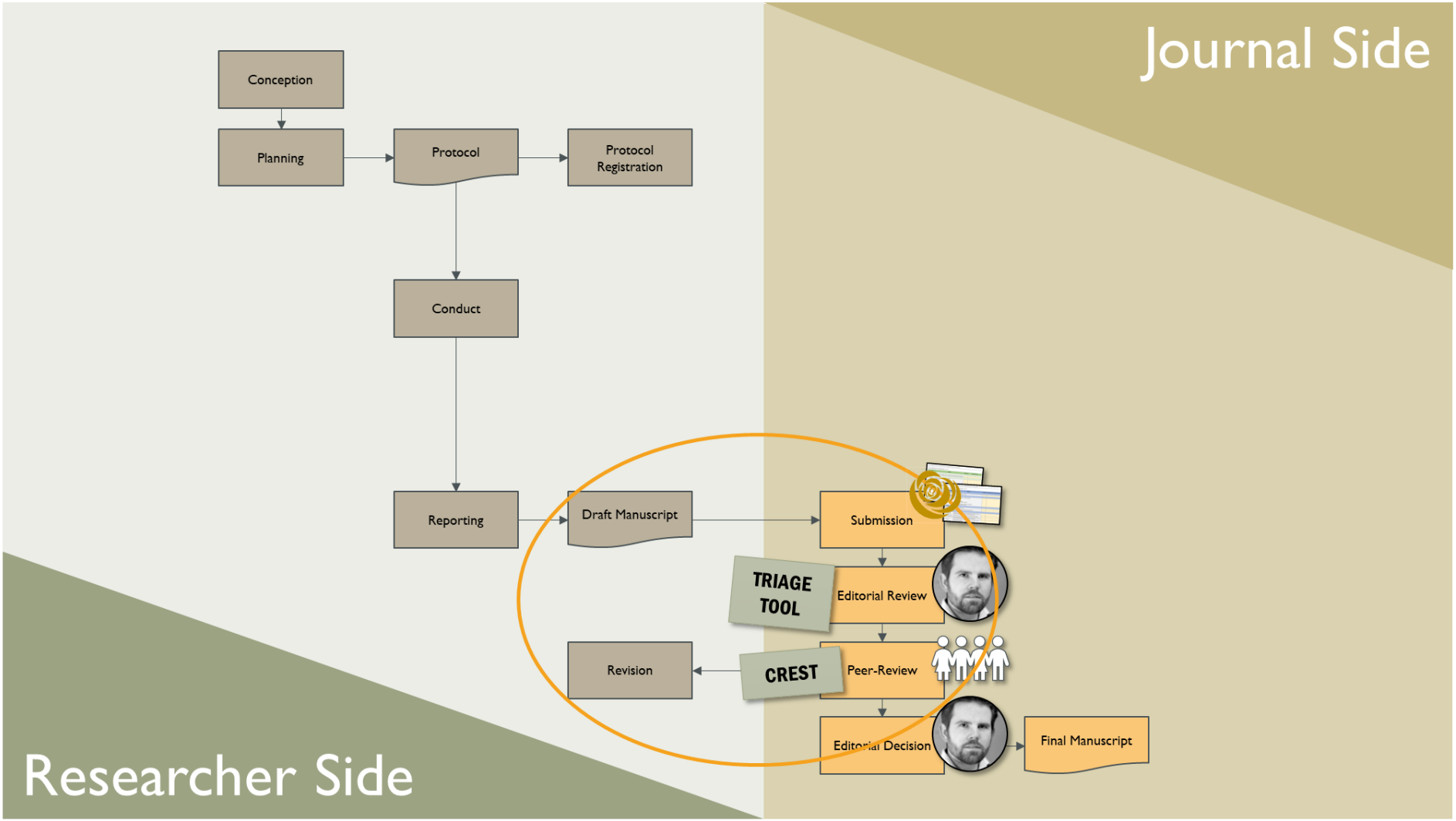
- 46 of 67 submissions rejected since using EVISE (~18 months)
  - 10 in process, 10 sent to production, one declined resubmission
  - 6 SRs, one SM, 2 commentaries, one correspondence
  - Only 3 SRs rejected post peer-review, 43 pre peer-review
- Hopefully that means we are at least filtering out the SRs which are not fit for purpose

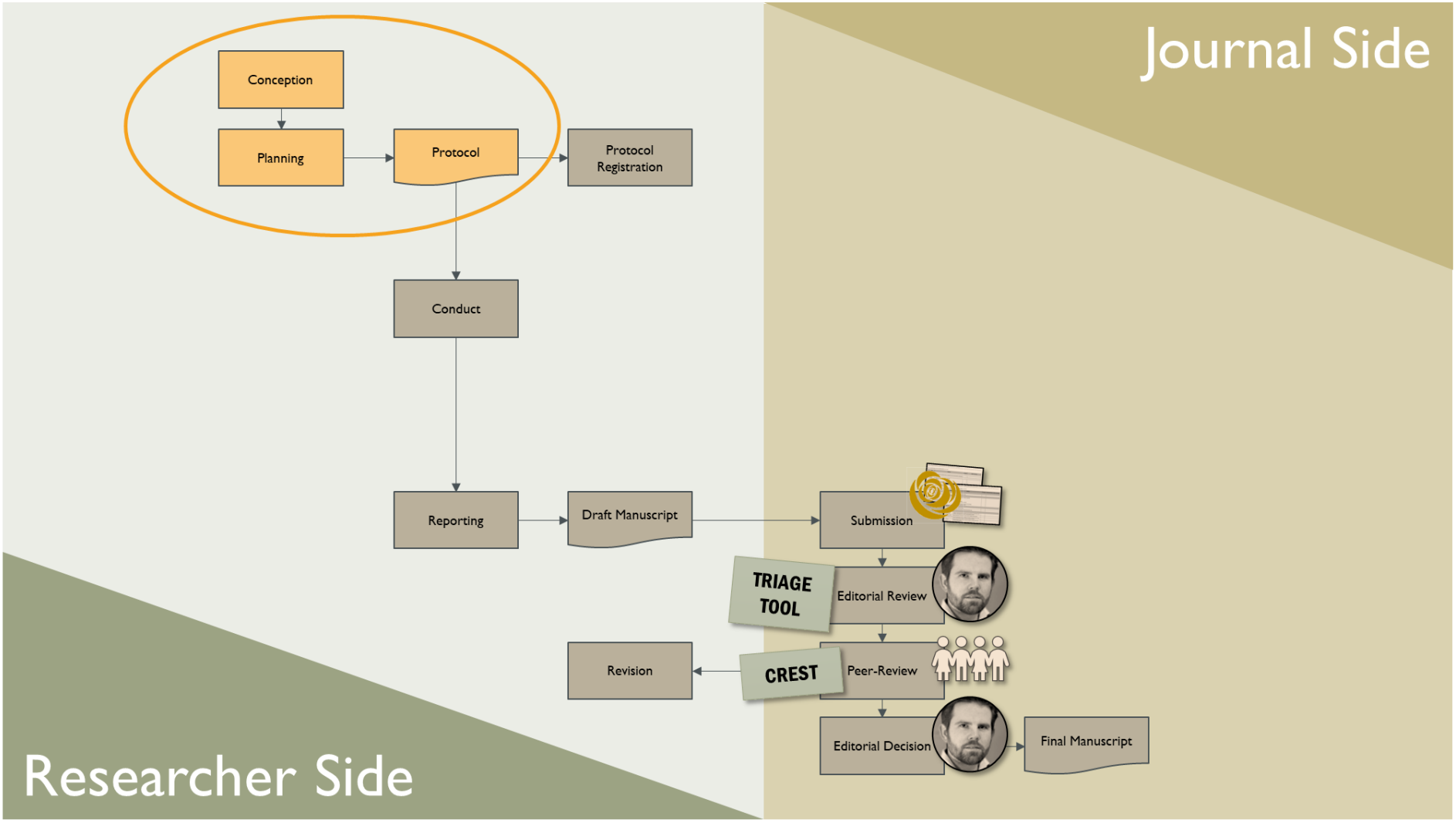
Is it really progress?

- We are mainly getting low-quality systematic reviews long after it's too late for the authors to address major issues (43 of 46 rejections are at desk; 2 years of work rejected in 2 minutes)
  - Objectives lacking research value and/or focus
  - Insensitive search strategies
  - Inappropriate inclusion criteria
  - Inadequate or non-existent risk of bias assessment methods
  - Unstructured, unsystematic interpretation of strength of evidence
- We are making sure readers aren't receiving misleading research (at least through our own journal) but could do much more to help submitting authors develop high-quality manuscripts

EDITOR AS MIDWIFE

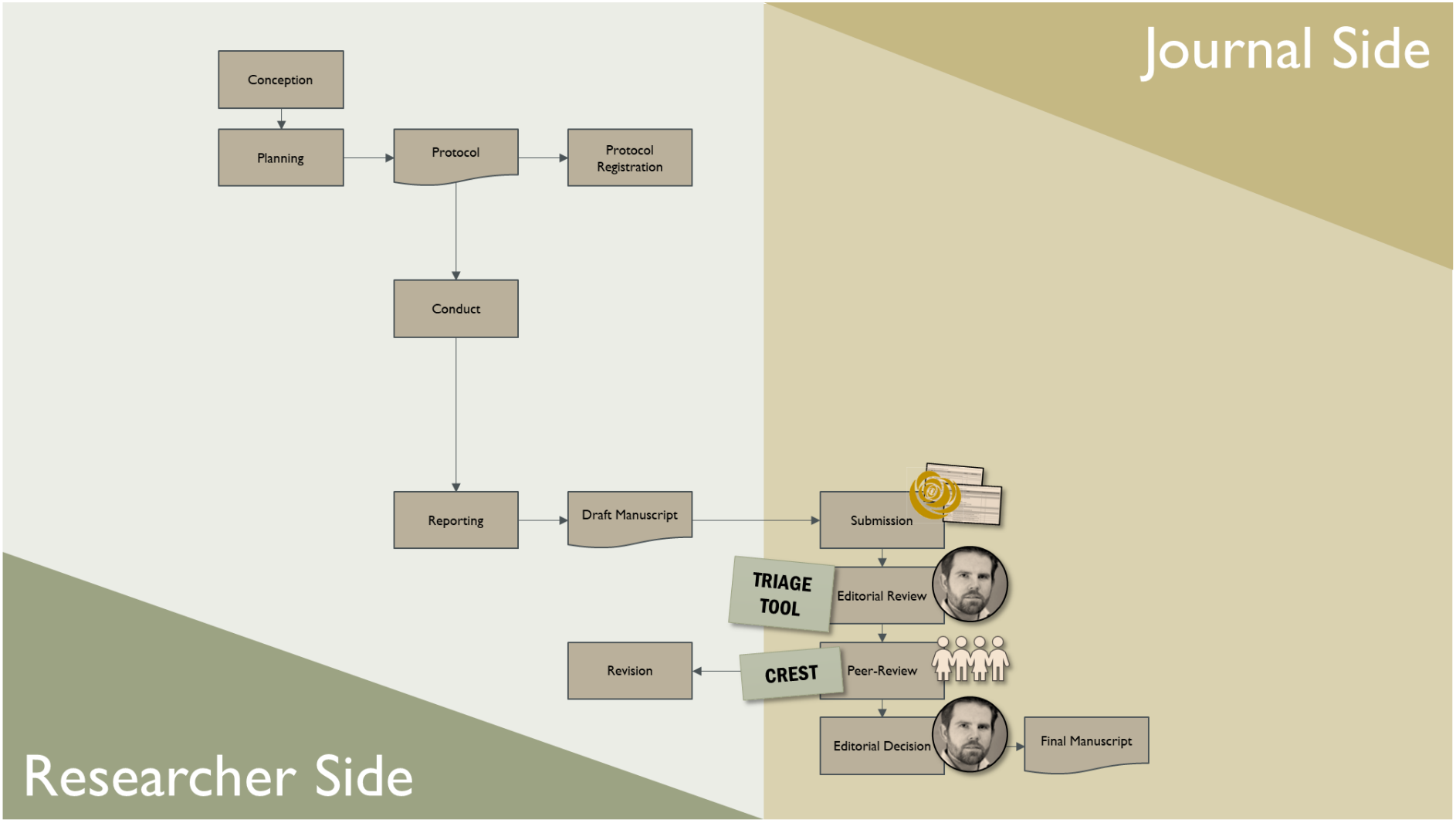
Rethinking the SR workflow and submission process

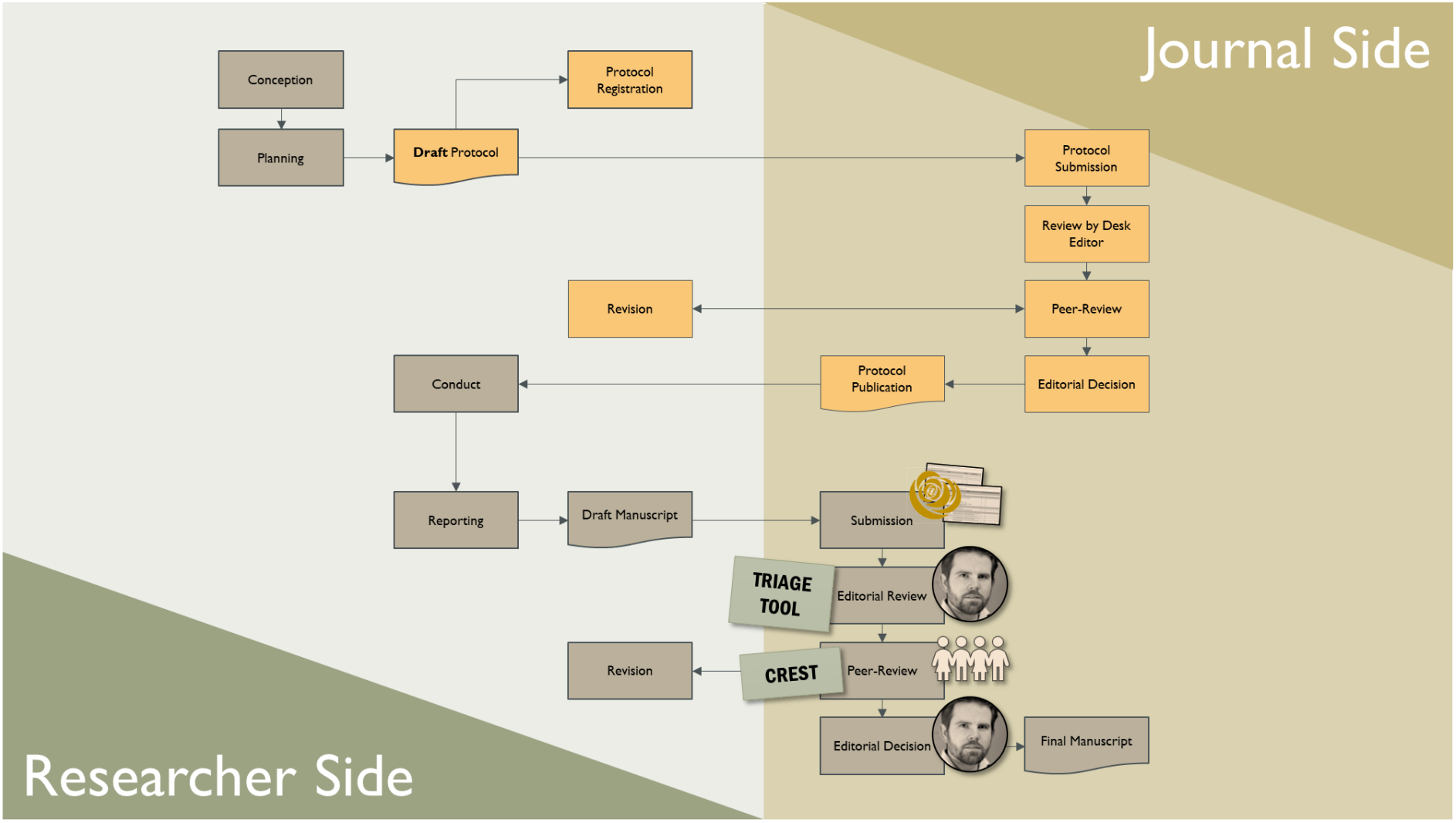

Journal Side
Conception
Planning
Protocol
Protocol Registration
Conduct
Reporting
Draft Manuscript
Submission
TRIAGE TOOL
Editorial Review
Revision
CREST
Peer-Review
Editorial Decision
Final Manuscript
Researcher Side

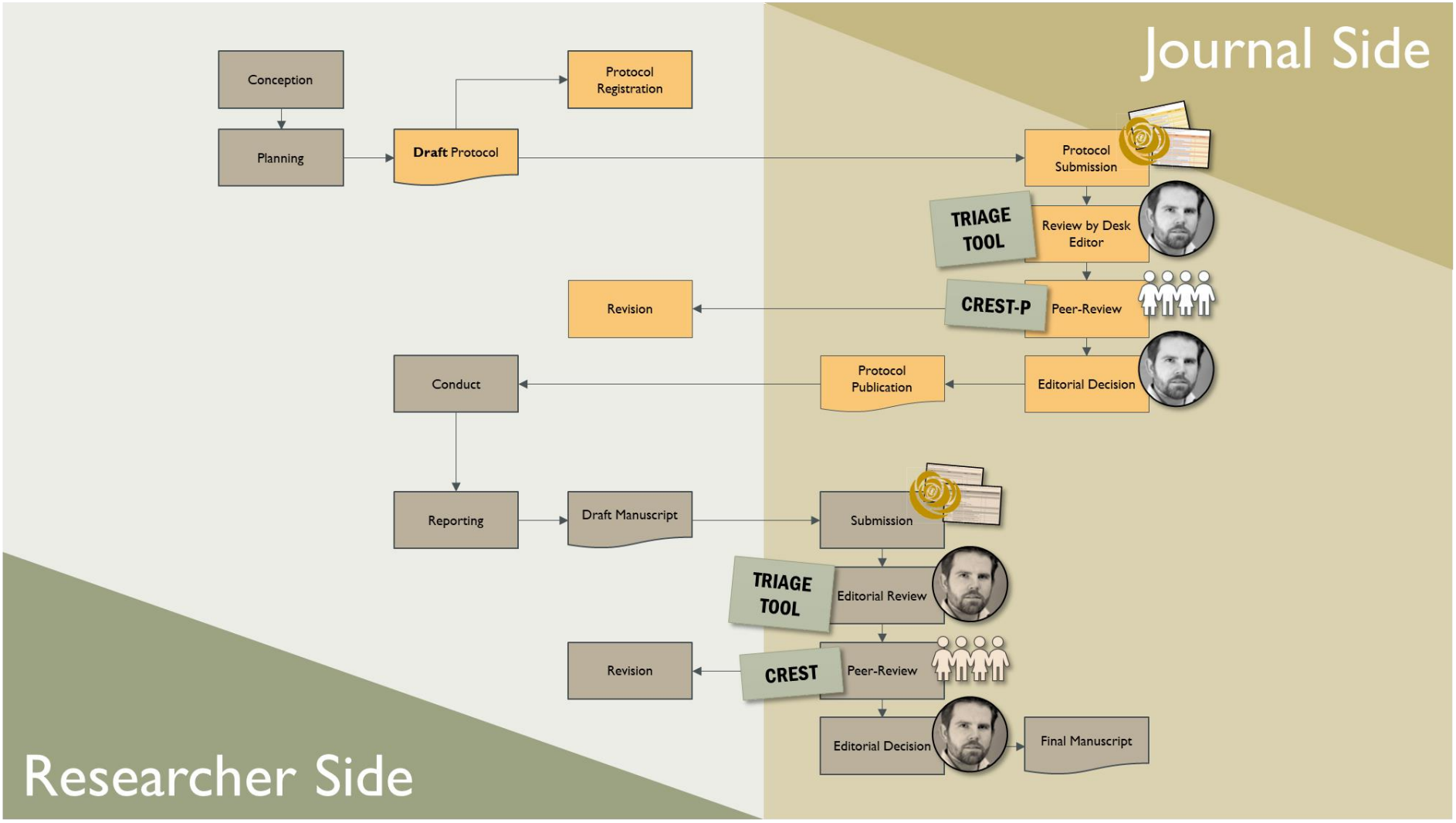

Journal Side
Researcher Side
Conception
Planning
Protocol
Protocol Registration
Conduct
Reporting
Draft Manuscript
Submission
TRIAGE TOOL
Editorial Review
CREST
Peer-Review
Revision
Editorial Decision
Final Manuscript

The solution: accept protocol submissions

- *Environment International* counts protocols as full publications
- First environmental health journal to do this
- Opens up multiple opportunities for editorial interventions

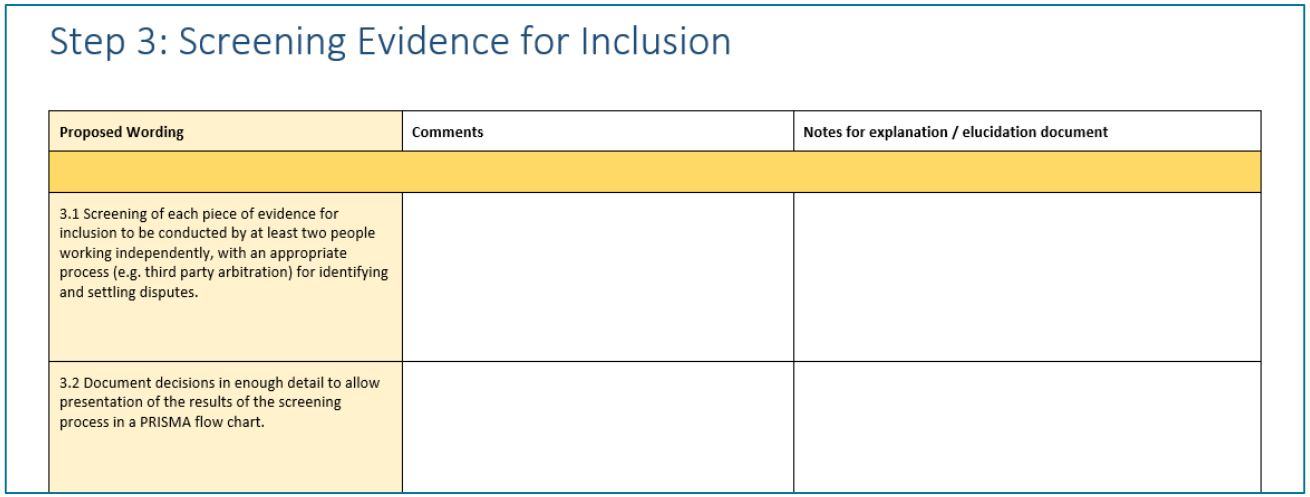

Journal Side
Researcher Side
Conception
Planning
Protocol
Protocol Registration
Conduct
Reporting
Draft Manuscript
Submission
TRIAGE TOOL
Editorial Review
CREST
Peer-Review
Revision
Editorial Decision
Final Manuscript


Journal Side
Researcher Side
Conception
Planning
Draft Protocol
Protocol Registration
Protocol Submission
Review by Desk Editor
Peer-Review
Editorial Decision
Protocol Publication
Revision
Conduct
Reporting
Draft Manuscript
Submission
TRIAGE TOOL
Editorial Review
CREST
Peer-Review
Revision
Editorial Decision
Final Manuscript


Journal Side
Researcher Side
Conception
Planning
Draft Protocol
Protocol Registration
Protocol Submission
TRIAGE TOOL
Review by Desk Editor
CREST-P
Peer-Review
Editorial Decision
Protocol Publication
Revision
Conduct
Reporting
Draft Manuscript
Submission
TRIAGE TOOL
Editorial Review
CREST
Peer-Review
Revision
Editorial Decision
Final Manuscript

Final piece of the puzzle

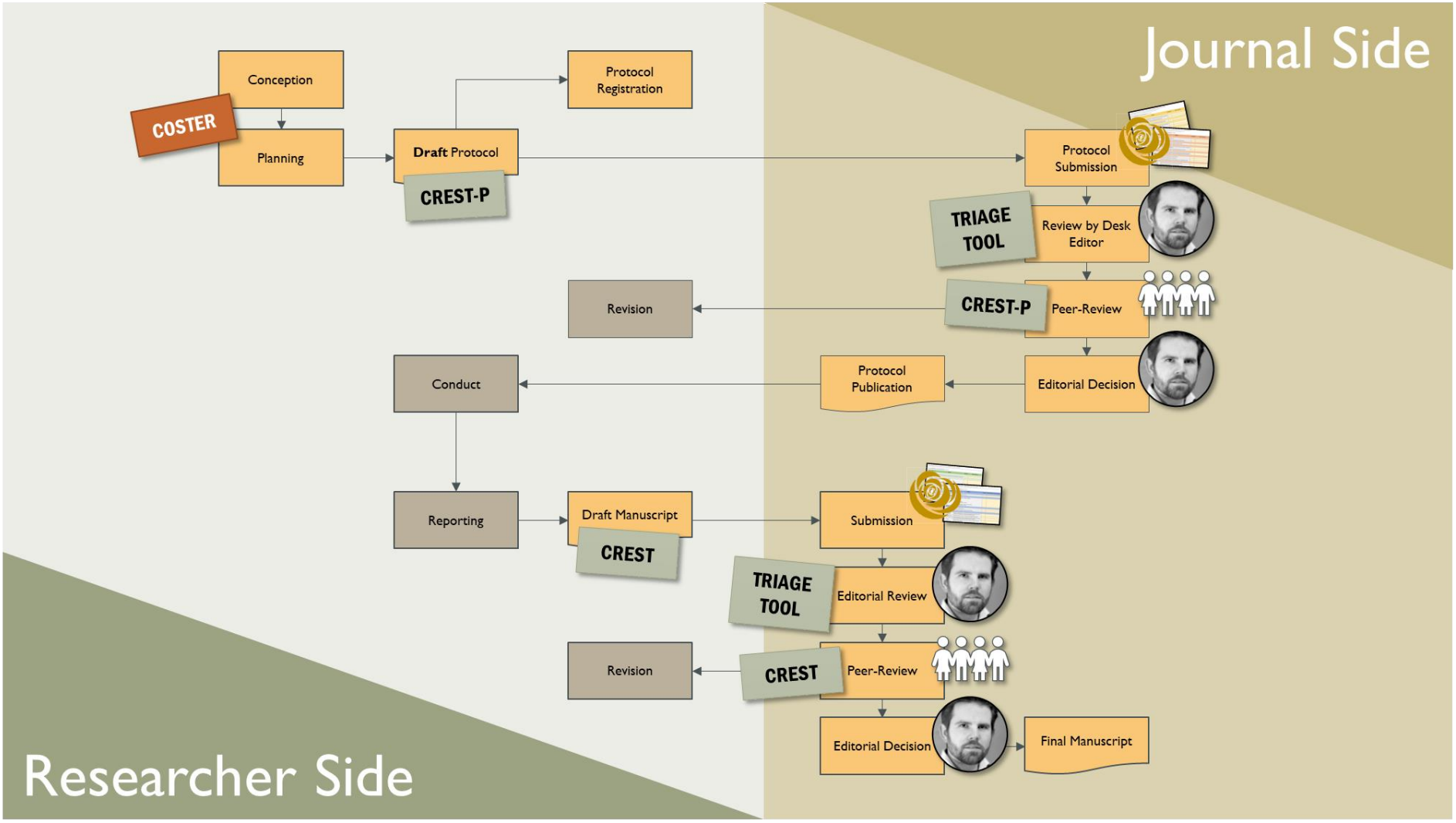
- "Recipe-book" for what researchers ought to do, to maximise chance of producing a fit-for-purpose systematic review
- Developing a tool called COSTER – 70 provisions across 8 stages of conducting a systematic review
- Makes explicit the required processes for fulfilling the criteria of e.g. PRISMA or ROSES, and for critical appraisal tools such as CREST



Step 3: Screening Evidence for Inclusion

Proposed Wording	Comments	Notes for explanation / elucidation document
3.1 Screening of each piece of evidence for inclusion to be conducted by at least two people working independently, with an appropriate process (e.g. third party arbitration) for identifying and settling disputes.		
3.2 Document decisions in enough detail to allow presentation of the results of the screening process in a PRISMA flow chart.		

Whaley et al. "A Code of Practice for Conduct of Systematic Reviews in Toxicology and Environmental Health Research (COSTER)" Under development


Journal Side
Researcher Side
Conception
COSTER
Planning
Draft Protocol
CREST-P
Protocol Registration
Protocol Submission
TRIAGE TOOL
Review by Desk Editor
CREST-P
Peer-Review
Editorial Decision
Protocol Publication
Revision
Conduct
Reporting
Draft Manuscript
CREST
Submission
TRIAGE TOOL
Editorial Review
CREST
Peer-Review
Editorial Decision
Final Manuscript
Revision

Implications for submitting authors

- Take advantage of our offer to review and publish protocols
- Follow best-practice standards for conduct of systematic reviews
- Think about the conduct implied by reporting standards
- For internal QC, use the same triage and peer-review tools we do
- Don’t assume that any stage of a systematic review is optional
- It’s good to be boring (results are irrelevant if methods are good)
- Find out more? **Subscribe to our newsletter**: http://bit.ly/overcite

Thank you.

Questions?

Image rights (in order of use)

Smokestacks / Guy Gorek / Flickr / CC BY-NC-ND

Plastic bottles / zhrefch / Flickr / public domain

Pesticide application / Oregon State University / Flickr / CC BY-SA

Food cans / King of Hearts / Wikimedia Commons

Till receipt / Till Dettmering / Wikimedia Commons

Drinks bottles / Amraepowell / Wikimedia Commons

Injected egg / Ekem / Wikimedia Commons