

# Experience and progress with developing guidance for the analysis of key topics in observational research

Sauerbrei W<sup>1</sup>, Abrahamowicz M<sup>2</sup>, Le Cessie S<sup>3</sup>, Huebner M<sup>4</sup>, Keogh R<sup>5</sup>, Carpenter J<sup>5</sup>  
for the STRATOS initiative

<sup>1</sup>Institute for Medical Biometry and Statistics, Faculty of Medicine and Medical Center University of Freiburg, Germany

<sup>2</sup>Department of Epidemiology and Biostatistics, McGill University, Montreal, Canada

<sup>3</sup>Department of Biomedical Data Sciences, Leiden University Medical Center, Leiden, The Netherlands

<sup>4</sup>Department of Statistics and Probability, Michigan State University, East Lansing, MI 48824, USA

<sup>5</sup>Department of Medical Statistics, London School of Hygiene & Tropical Medicine, London, UK

<http://stratos-initiative.org/>

# Steering Group of the STRATOS initiative

- Willi Sauerbrei, Freiburg, GER (Chair)
- Michal Abrahamowicz, Montreal, CA (Co-Chair)
- Marianne Huebner, East Lansing, US (Co-Chair)
- Ruth Keogh, London, UK (Co-Chair)
- Mark Baillie, Basel, CH
- Anne-Laure Boulesteix, Munich, GER
- James Carpenter, London, UK
- Gary Collins, Oxford, UK
- Mitchell Gail, Bethesda, US
- Els Goetghebeur, Ghent, BE
- Georg Heinze, Vienne, AU
- Victor Kipnis, Bethesda, US
- Saskia le Cessie, Leiden, NL
- Joerg Rahnenfuehrer, Dortmund, GER
- Pamela Shaw, Philadelphia, US
- Maarten van Smeden, Utrecht, NL
- Ewout Steyerberg, Leiden, NL

## Earlier members for at least two years

- Doug Altman
- Ray Carroll
- Harbajan Chadha-Boreham
- Stephen Evans
- Petra Macaskill
- Andrew Vickers

# Launch of the STRATOS initiative 2013

The STRATOS initiative was launched at a half-day [Mini-Symposium](#) on the last day of the [ISCB2013 meeting in Munich](#), in August 2013.

There were three general talks by Willi Sauerbrei, Doug Altman and James Carpenter. James also introduced a route map by using TG1 (missing data) as an example.

In addition, there were specific talks discussing key issues of TG2 through TG7:

- TG2 Selection of variables and functional form, Michal Abrahamowicz
- TG3 Descriptive and initial data analysis, Saskia Le Cessie
- TG4 Measurement error, Helmut Küchenhoff
- TG5 Design, Doug Altman
- TG6 Evaluating diagnostic tests and prediction models, Petra Macaskill
- TG7 Causal Inference, Els Goetghebeur

TG8 Survival analysis and TG9 High-dimensional data started in 2015

# Situation in 2013

- Funding for a guidance initiative was (and still is) a problem
- Funders and journals prioritize novelty over incremental replication research
- Reporting guidelines were still in the ‘early days’  
EQUATOR started in 2006, TRIPOD published in 2015
- Lancet series ‘Reduce waste, increase value’ not published (in 2014)  
Medical research needs to change – HOW?
- Open Access papers – rare and often rather negative reputation
- Center for Open Science started in 2013
- FAIR (**F**indability, **A**ccessibility, **I**nteroperability, and **R**euse) principles in 2016
- Meta Research still in its early days
- ..... more to mention???

---

# Why we need guidance documents for the design and analysis of observational studies



**Doug Altman**

***EQUATOR Network  
Centre for Statistics in Medicine  
University of Oxford, UK***



# Obligation

**“In return for the altruism and trust that make clinical research possible, the research enterprise has an obligation to **conduct** research ethically and to **report it honestly.**”** [i.e. transparently and completely]

[International Committee of Medical Journal Editors, 2004]

- **Medical research is very important – it affects people’s lives**
- **Researchers have an obligation to do high quality research**
  - Scientific, ethical, financial considerations
- **These issues are most obvious for RCTs but apply also to observational studies**

# Observational studies

- Most published research articles (towards 90%) report observational studies
- Evidence from numerous reviews of publications shows that design, analysis and reporting are often substandard



- RCTs
- Observational studies
- ... indeed everything!

# Guidelines for observational studies

- **Several guidelines have outlined the essential elements of reporting observational studies of different designs (see [equator-network.org](http://equator-network.org))**
  - STROBE (epidemiological cohort, case-control, cross-sect) 2007
    - Extensions STREGA, STROBE-ME,...
  - REMARK (tumour marker prognostic studies) 2005
  - TRIPOD (multivariable prediction models ) 2015
  - GRIPS (genetic risk prediction studies) 2011
- **There is a clear need for companion guidelines for research conduct**
  - Would be of particular benefit to those without formal training or limited experience



# What should be our scope?

- **Observational studies cover a wide variety of research questions**
  - Effects of interventions
  - Incidence, Aetiology, Prognosis, Diagnosis, ...
- **... and study designs**
  - Cohort, Case-control, Cross-sectional
  - Interrupted time series, Ecological, ...
- **... and data sources**
  - Prospective planned studies
  - Routinely collected data
- **Guidance can be generic (e.g. missing data) or specific (e.g. design of case-control studies)**

**Guidance Initiative:—  
a route map, and the journeying of the missing data topic group (TG1)**

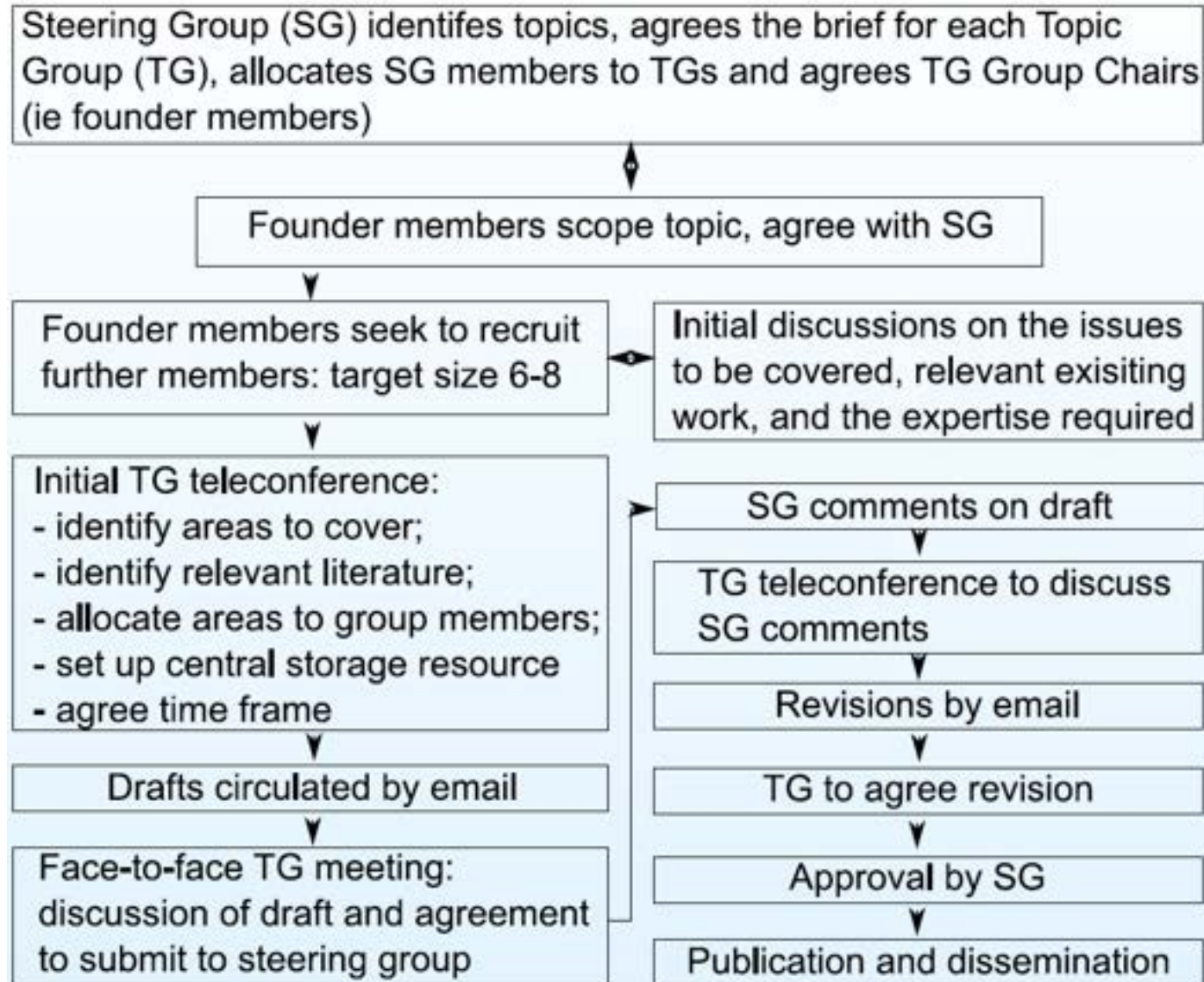
**James Carpenter**

London School of Hygiene & Tropical Medicine *and* MRC Clinical Trials Unit

Email: [james.carpenter@lshtm.ac.uk](mailto:james.carpenter@lshtm.ac.uk)

August 28, 2013

# Initial route map



# Topics

The steering group has identified the following topics on the basis of their importance for reliable inference from medical data:

- TG1 Missing data
- TG2 Selection of variable and functional form
- TG3 Descriptive and initial data analysis
- TG4 Measurement error
- TG5 Design
- TG6 Evaluating diagnostic tests and prediction models
- TG7 Causal inference
- TG8 ...
- TG9 ...
- TG10 ...

The numbering of the topic groups does not represent either the relative importance of the topics or the order in which they should be considered!

# First port of call

- Level 2: Experienced statistician  
(typically Masters level plus work experience  
= ‘somewhat below state-of-the-art’)
  - outline issues
  - outline methodological approaches, and their pros and cons
  - guidance on software

## Output:

1. SG publication about initiative aims *etc*
2. For each TG:
  - a longer report (around 10,000 words) on the Guidance Initiative’s website—open for external comments, and
  - publications deriving for this.

# Subsequently:

Guidance for:

1. Level 1: Less experienced analysts, eg
  - those without formal statistical training
  - recently qualified Masters students
2. Level 3: Experienced statisticians, eg
  - active researchers
  - research group leaders

Level 1: probably the same TG

Level 3: some additional experts maybe added

# Missing data TG

Scope (principally clinical, but also other application areas):

- Likely effects of missing data:
  - marginal statistics;
  - regression;
  - implications for more complex settings....
- Methodology:
  - Complete Records
  - Multiple imputation
  - EM algorithm
  - Inverse probability weighting
  - Combinations of the above
- Reporting

# ‘We’ve been here before...’

Do existing review papers in a particular area mean there is no need for further work—just dissemination?

E.g. missing data in clinical trials, where in 2010 the National Academies Press published a report ‘The prevention and Treatment of Missing Data in Clinical Trials’ prepared by a panel of experts chaired by Rod Little [2].

This

- convened an expert group;
- took evidence;
- wrote a report with recommendations for action, and
- published commentaries published in various journals

Further, the European Medicines Association published a ‘Guideline on Missing Data in Confirmatory Clinical Trials’ (2010)[1].

What do we add to these?



## ...or have we?

- the US National Academies of Science report did not aim to be a guideline; rather it was commissioned by the FDA to inform their guidelines.
- The EMEA guidelines are closer to our remit, but—partly because of process – are less statistically credible.
- Neither
  - has focus on observational data;
  - sets out to provide a systematic guide for analysts at an appropriate level;
  - is part of a unified initiative;
  - has a web-based home providing an up-to-date links to the literature, and documenting comments

The aim is an accessible document raising awareness of the issues to be considered and how various methods may be applied, not a cookbook.

# Why do we need such an initiative?

- Weaknesses of many analyses
- Many analyses are conducted by people with limited statistical knowledge
- Analysis of observational studies – issues are very similar in all areas of science
- Rapid developments of statistical methodology requires guidance and education

# Guidance and education required

- Statistical models are always a simplification of real life processes. To improve these models, researchers develop new and more complicated approaches
- Different and partly conflicting approaches are proposed
- Expert knowledge is required to use methods
- Statistical software has to be available

# Current situation

- Statistical methodology has seen some substantial development
- Computer facilities can be viewed as the cornerstone
- Possible to assess properties and compare complex model building strategies using simulation studies
- Resampling and Bayesian methods allow investigations that were impossible two decades ago
- Machine learning approaches may be interesting alternatives to more traditional approaches
- Wealth of new statistical software packages allow a rapid implementation and verification of new statistical ideas

**Unfortunately, many sensible improvements are ignored in practical statistical analyses**

# Reasons that improved strategies are ignored

- Overwhelming concern with **theoretical aspects**
- Very **limited guidance** on key issues that are **vital in practice**, discourages analysts from utilizing more sophisticated and possibly more appropriate methods in their analyses

# Improvement

At least **two tasks** are essential

1. **Experts** in specific methodological areas have to work towards **developing guidance documents**
2. An ever-increasing need for **continuing education** at all stages of the career

For busy applied researchers it is often **difficult to follow methodological progress** even in their principal application area

- Reasons are diverse
- Consequence is that analyses may be deficient
- **Knowledge** gained through research on statistical methodology needs to be **transferred** to the broader community
- Many **analysts** would be **grateful for** an overview on the current **state of the art** and for **practical expert guidance**

# Aims of the initiative

- **Provide guidance documents** for highly relevant issues
- As the statistical **knowledge** of the analyst **varies** substantially, guidance has to keep this background in mind. **Guidance** documents have to be provided **at several levels**
- For the **start** we will concentrate on **state-of-the-art** documents for experienced statisticians (**level 2**)
- Help to identify questions **requiring more primary research**

**The overarching long-term aim is to improve key parts of statistical analyses of observational studies in practice**

# Short term aims

- For a small number of **highly relevant topics** we will try to assess the current **state of practice** and identify current documents which try provide some guidance
- Try to find **'agreement' what to recommend** within each of the three levels mentioned
- Documents have to be **understood** and at least broadly **accepted**
- **Positive examples** from the current literature will help to reach the latter goal



# Long term aims

- **Improvement of statistical analyses**, acceptance of guidance documents by analysts could be a cornerstone
- The percentage of analysts who reflect only the **'level 1'** knowledge should **decrease substantially**
- Guidance **documents** have to be **regularly improved**. Based on evidence some approaches should be ready to be moved from level 3 into a level 2 recommendation
- **Software** is generally available and usable at a broader level
- Number of **topics** is large and there is often a **relation between them** or one is influenced by the other.
- For example, variable selection and missing data: First it is important to **derive guidance for both of them separately**. For practical reasons it is highly relevant to consider the **implications of** missing data **on** guidance for variable selection.

# Main goal of STRATOS

To **improve the current practice** in design and statistical analyses of observational studies in practice by **closing the gap** between available statistical methodology and methods applied in real-life through **guidance for researchers** with different levels of statistical expertise.

The focus is on health sciences research, but the content is also relevant for applications of statistics in other empirical sciences.

## Different levels of statistical knowledge

### Level 1: Low statistical knowledge

- Many studies are analysed by researchers with this level

### Level 2: Experienced statistician

- Uses methodology which is ok, however often not according to state of the art







### Level 3: Expert in a specific area

- Performs research to improve statistical models and adapts them to complex real problems.

# What about statistical guidance from journals?

---

## Statistical Guidance to Authors at Top-Ranked Journals across Scientific Disciplines

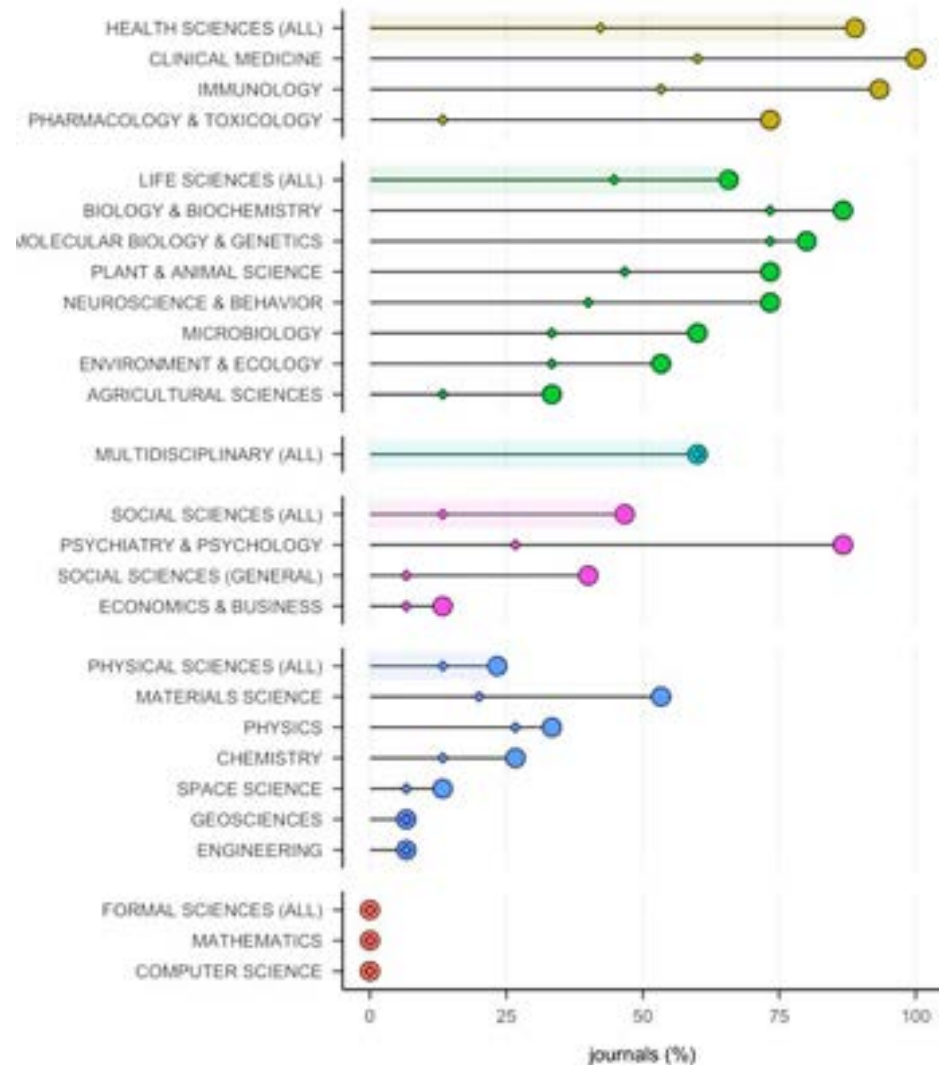
Tom E. Hardwicke<sup>a,b</sup> , Maia Salholz-Hillel<sup>c</sup> , Mario Malički<sup>d</sup> , Dénes Szűcs<sup>e</sup> , Theiss Bendixen<sup>f</sup> ,  
and John P. A. Ioannidis<sup>d,g,h</sup> 

THE AMERICAN STATISTICIAN  
2022, VOL. 00, NO. 0, 1–9: General  
<https://doi.org/10.1080/00031305.2022.2143897>

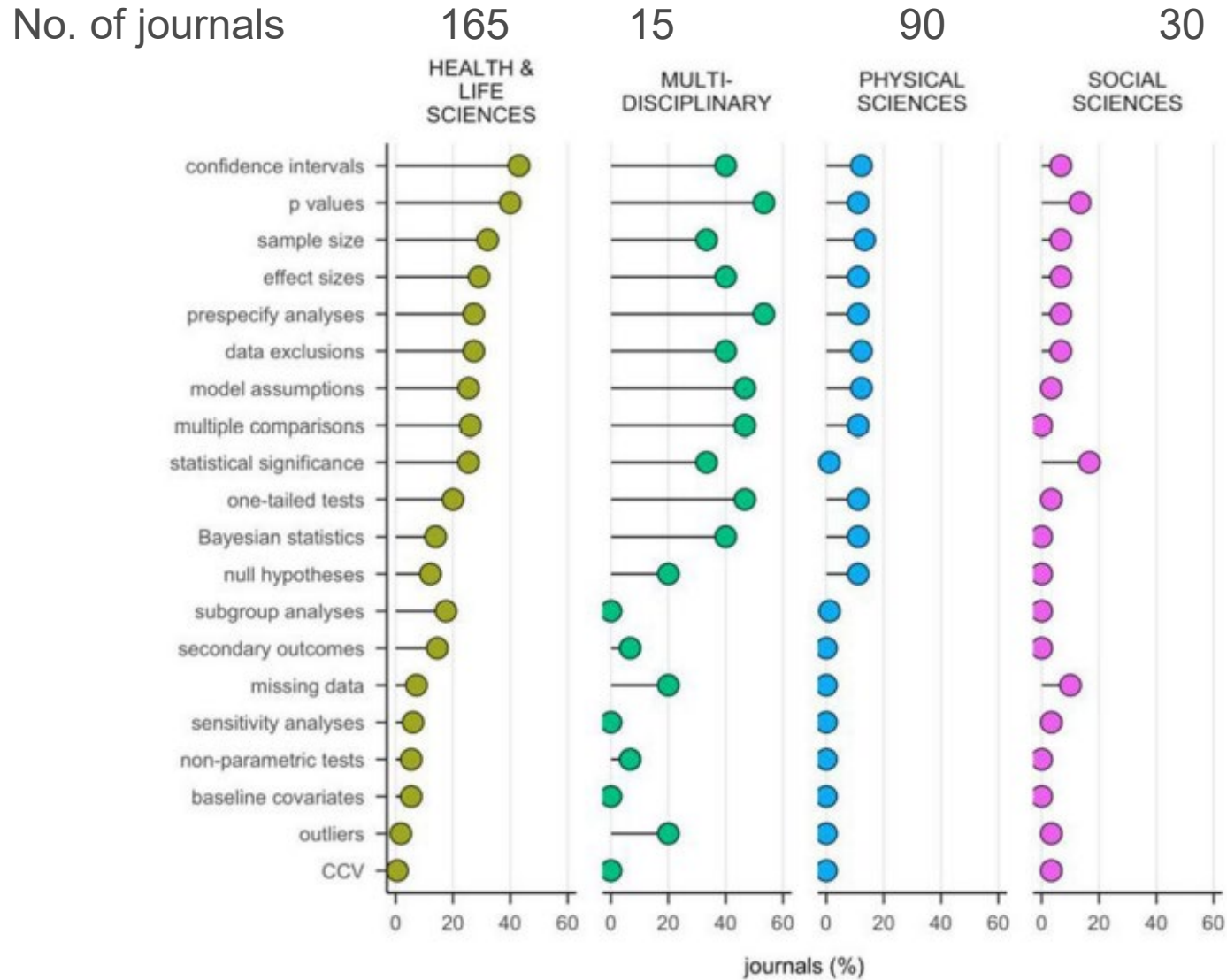
Statistical guidance at 15 journals (top ranked by IF) in each of 22 scientific disciplines

n = 330 journals

# Frequency of journals offering some statistical guidance (circles) or having a dedicated statistical guidance section(diamonds)



# Frequency of journals offering guidance on 20 prespecified statistical topics



# Mentioning of twenty prespecified statistical topics

Not mentioned (blue square)    Mentioned (green square)

confidence intervals  
 p values  
 sample size  
 effect sizes  
 prespecify analyses  
 data exclusions  
 model assumptions  
 multiple comparisons  
 statistical significance  
 one-tailed tests  
 Bayesian statistics  
 null hypotheses  
 subgroup analyses  
 secondary outcomes  
 missing data  
 sensitivity analyses  
 non-parametric tests  
 baseline covariates  
 outliers  
 CCV

	CLINICAL MEDICINE																			
JAMA ONCOLOGY	Mentioned	Mentioned	Mentioned	Mentioned	Mentioned	Mentioned	Mentioned	Mentioned	Mentioned	Mentioned	Not mentioned	Not mentioned	Mentioned	Mentioned	Mentioned	Mentioned	Not mentioned	Mentioned	Not mentioned	Not mentioned
JAMA INTERNAL MEDICINE	Mentioned	Mentioned	Mentioned	Mentioned	Mentioned	Mentioned	Mentioned	Mentioned	Mentioned	Mentioned	Not mentioned	Not mentioned	Mentioned	Mentioned	Mentioned	Mentioned	Not mentioned	Mentioned	Not mentioned	Not mentioned
JAMA	Mentioned	Mentioned	Mentioned	Mentioned	Mentioned	Mentioned	Mentioned	Mentioned	Mentioned	Mentioned	Not mentioned	Not mentioned	Mentioned	Mentioned	Mentioned	Mentioned	Not mentioned	Mentioned	Not mentioned	Not mentioned
EUROPEAN HEART JOURNAL	Mentioned	Mentioned	Mentioned	Mentioned	Mentioned	Mentioned	Mentioned	Mentioned	Mentioned	Mentioned	Not mentioned	Not mentioned	Mentioned	Not mentioned	Mentioned	Mentioned	Not mentioned	Mentioned	Not mentioned	Mentioned
NEW ENGLAND JOURNAL OF MEDICINE	Mentioned	Mentioned	Mentioned	Mentioned	Not mentioned	Not mentioned	Mentioned	Mentioned	Mentioned	Mentioned	Not mentioned	Not mentioned	Mentioned	Mentioned	Mentioned	Mentioned	Not mentioned	Not mentioned	Not mentioned	Not mentioned
JOURNAL OF CLINICAL ONCOLOGY	Mentioned	Not mentioned	Mentioned	Mentioned	Not mentioned	Not mentioned	Mentioned	Not mentioned	Mentioned	Mentioned	Not mentioned	Not mentioned	Mentioned	Mentioned	Mentioned	Mentioned	Not mentioned	Not mentioned	Not mentioned	Not mentioned
ANNALS OF INTERNAL MEDICINE	Mentioned	Not mentioned	Mentioned	Mentioned	Not mentioned	Not mentioned	Mentioned	Mentioned	Mentioned	Mentioned	Not mentioned	Not mentioned	Mentioned	Mentioned	Mentioned	Mentioned	Not mentioned	Not mentioned	Not mentioned	Not mentioned
LANCET RESPIRATORY MEDICINE	Mentioned	Not mentioned	Mentioned	Mentioned	Not mentioned	Not mentioned	Mentioned	Mentioned	Mentioned	Mentioned	Not mentioned	Not mentioned	Mentioned	Mentioned	Mentioned	Mentioned	Not mentioned	Not mentioned	Not mentioned	Not mentioned
LANCET ONCOLOGY	Mentioned	Not mentioned	Mentioned	Not mentioned	Not mentioned	Not mentioned	Mentioned	Mentioned	Mentioned	Mentioned	Not mentioned	Not mentioned	Mentioned	Mentioned	Mentioned	Mentioned	Not mentioned	Not mentioned	Not mentioned	Not mentioned
LANCET DIABETES & ENDOCRINOLOGY	Mentioned	Not mentioned	Mentioned	Mentioned	Not mentioned	Not mentioned	Mentioned	Mentioned	Mentioned	Mentioned	Not mentioned	Not mentioned	Mentioned	Mentioned	Mentioned	Mentioned	Not mentioned	Not mentioned	Not mentioned	Not mentioned
LANCET	Mentioned	Not mentioned	Mentioned	Mentioned	Not mentioned	Not mentioned	Mentioned	Mentioned	Mentioned	Mentioned	Not mentioned	Not mentioned	Mentioned	Mentioned	Mentioned	Mentioned	Not mentioned	Not mentioned	Not mentioned	Not mentioned
THE BMJ	Mentioned	Not mentioned	Mentioned	Mentioned	Not mentioned	Not mentioned	Mentioned	Mentioned	Mentioned	Mentioned	Not mentioned	Not mentioned	Mentioned	Mentioned	Mentioned	Mentioned	Not mentioned	Not mentioned	Not mentioned	Not mentioned
CIRCULATION	Not mentioned	Mentioned	Mentioned	Not mentioned	Not mentioned	Not mentioned	Mentioned	Mentioned	Mentioned	Mentioned	Not mentioned	Not mentioned	Mentioned	Mentioned	Mentioned	Mentioned	Not mentioned	Not mentioned	Mentioned	Not mentioned
CANCER DISCOVERY	Mentioned	Not mentioned	Mentioned	Not mentioned	Not mentioned	Not mentioned	Mentioned	Mentioned	Mentioned	Mentioned	Not mentioned	Not mentioned	Mentioned	Mentioned	Mentioned	Mentioned	Not mentioned	Not mentioned	Not mentioned	Not mentioned
GASTROENTEROLOGY	Not mentioned	Not mentioned	Not mentioned	Not mentioned	Not mentioned	Not mentioned	Not mentioned	Not mentioned	Not mentioned	Not mentioned	Not mentioned	Not mentioned	Not mentioned	Not mentioned	Not mentioned	Not mentioned	Not mentioned	Not mentioned	Not mentioned	Not mentioned

CCV  
 categorizing  
 continuous  
 variables

