STRengthening **A**nalytical **T**hinking for **O**bservational **S**tudies (**STRATOS**) initiative - very brief introduction

Willi Sauerbrei

for the Steering Group and the organizers of the workshop

http://stratos-initiative.org/

Institute for Medical Biometry and Statistics, Medical Center University of Freiburg, Germany

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
- Riccardo de Bin, Oslo, NOR

- Mitchell Gail, Bethesda, US
- Els Goetghebeur, Ghent, BE
- Georg Heinze, Vienne, AU
- Victor Kipnis, Bethesda, US
- Saskia le Cessie, Leiden, NL
- 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
- Joerg Rahnenfuehrer
- Andrew Vickers

STRATOS initiative: STRengthening Analytical Thinking for Observational Studies

Main goal

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 though **guidance for researchers** with different levels of statistical expertise.

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.

UNIVERSITÄTS KLINIKUM FREIBURG

STRATOS – History and Milestones

- **2011: Epi Subcom** at **I**nt **S**oc **C**lin **B**iostatistics (ISCB) in Ottawa

- **2013: First mini-symposium** at **ISCB** in Munich, **initiative launched**

- **2014: 1st STRATOS paper,** *Statistics in Medicine* 2014

- **2016 & 2019: 2 General meetings**, Banff Int Res Station (BIRS), Canada

- **Series in Biometric Bulletin (since 3/2017):** 27 articles published

- **2019:** Partner in the **Setting International Standards in Analysing Patient-Reported Outcomes and Quality of Life Endpoints (SISAQOL) project**, lead by EORTC (41 stakeholders, including pharma, regulators, patients)

- **2021: Memorandum of Understanding** with **ISCB**

- **2024: 3rd General meeting**, Lorentz Center, Leiden (NL)

As of 2024: >100 members (from 21 countries on 5 continents)

UNIVERSITÄTS
KLINIKUM FREIBURG

STRATOS Topic Groups (TGs)

	Topic Group	Chairs
1	Missing data	James Carpenter (UK), Kate Lee (AUS)
2	Selection of variables and functional forms in multivariable analysis	Georg Heinze (AUT), Aris Perperoglou (UK), Willi Sauerbrei (GER)
3	Initial data analysis	Marianne Huebner (US), Lara Lusa (SL), Carsten Oliver Schmidt (GER)
4	Measurement error and misclassification	Victor Kipnis (US), Pam Shaw (US)
5	Study design	Suzanne Cadarette (CAN), Mitchell Gail (US)
6	Evaluating diagnostic tests and prediction models	Ewout Steyerberg (NL), Ben van Calster (BEL)
7	Causal inference	Els Goetghebeur (BEL), Ingeborg Waernbaum (SWE)
8	Survival analysis	Michal Abrahamowicz (CAN), Malka Gorfine (IS), Terry Therneau (US)
9	High-dimensional data	Lisa McShane (US), Joerg Rahnenfuehrer (GER), Riccardo de Bin (NOR)

Chairs from 12 countries and 3 continents

UNIVERSITÄTS KLINIKUM FREIBURG

STRATOS Cross-cutting Panels

Panel		Chairs and Co-Chairs	
MP	Membership	Chairs:	James Carpenter (UK), Willi Sauerbrei (GER)
PP	Publications	Chairs:	Bianca De Stavola (UK), Pam Shaw (US), Mitchell Gail (US), Mark Baillie (CH)
GP	Glossary	Chairs:	Martin Boeker (GER), Marianne Huebner (US)
WP	Website	Chairs:	Joerg Rahnenfuehrer (GER), Willi Sauerbrei (GER)
RP	Literature Review	Chairs:	Gary Collins (UK), Carl Moons (NL)
BP	Bibliography	Chairs:	to be determined
SP	Simulation Studies	Chairs:	Michal Abrahamowicz (CAN), Anne-Laure Boulesteix (GER)
DP	Data Sets	Chairs:	Saskia Le Cessie (NL), Maarten van Smeden (NL)
TP	Knowledge Translation	Chair:	Maarten van Smeden (NL)
CP	Contact Organisations	Chair:	Willi Sauerbrei (GER)
VP	Visualisation	Chairs:	Mark Baillie (CH)
OS	Open Science	Chair:	Sabine Hoffmann (GER)

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???

An overview of activities is summarized in the Biometrical Bulletin, since 2017

(STRATOS): Six foci for the next three years

James Carpenter (1), Michal Abrahamowicz (2), Nan van Geloven (3), Paul Gustafson (4), Marianne Huebner (5), Ruth Keogh (1), Willi Sauerbrei (6), Pamela Shaw (7), Els Goetghebeur (8)

(STRATOS): Introducing the Open Science Panel

Sabine Hoffmann1, Kim Luijken2, Willi Sauerbrei3, Pamela Shaw4, Anne-Laure Boulesteix5

(STRATOS): On the importance of Data Quality Assessments and Initial Data Analysis

C. O. Schmidt[1], G Heinze[2], L Lusa[3], M Huebner[4]

(STRATOS): Overview of methodological issues when analyzing high-dimensional biomedical data

De Bin R., McShane L., Rahnenführer J. on behalf of STRATOS TG9 (2023)

(STRATOS): Guidance for analysts with limited statistical knowledge

Georg Heinze1, Anne-Laure Boulesteix2, Daniela Dunkler1, Mitchell H. Gail3, Katherine J. Lee4, Ben van Calster5, Michael Wallace6, Willi Sauerbrei7

a very brief update on the achievements of the STRATOS initiative in the last 5 years

Willi Sauerbrei[1], Michal Abrahamowicz[2], Mark Baillie[3], Bianca De Stavola[4], Mitchell Gail[5], Marianne Huebner[6], Ruth Keogh[7] and Pamela A. Shaw[8] for the STRATOS initiative

(STRATOS): Progress in the Topic Group on Evaluating Diagnostic Tests and Prediction Models (TG6)

Ewout W Steyerberg[1], Ben Van Calster[1,2], on behalf of STRATOS TG6

STRATOS foci for the next three years

1. Simulation studies
2. Open science
3. Initial Data Analysis (IDA)
4. Machine learning (ML) enhanced statistical methods
5. Estimands in observational data analysis
6. More guidance for researchers with limited statistical knowledge and experience

UNIVERSITÄTS
KLINIKUM FREIBURG

What about statistical guidance from journals?

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

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

Statistical guidance at 15 journals (top ranked by IF) in each of 22 scientific disciplines
n = 330 journals

UNIVERSITÄTS
KLINIKUM FREIBURG
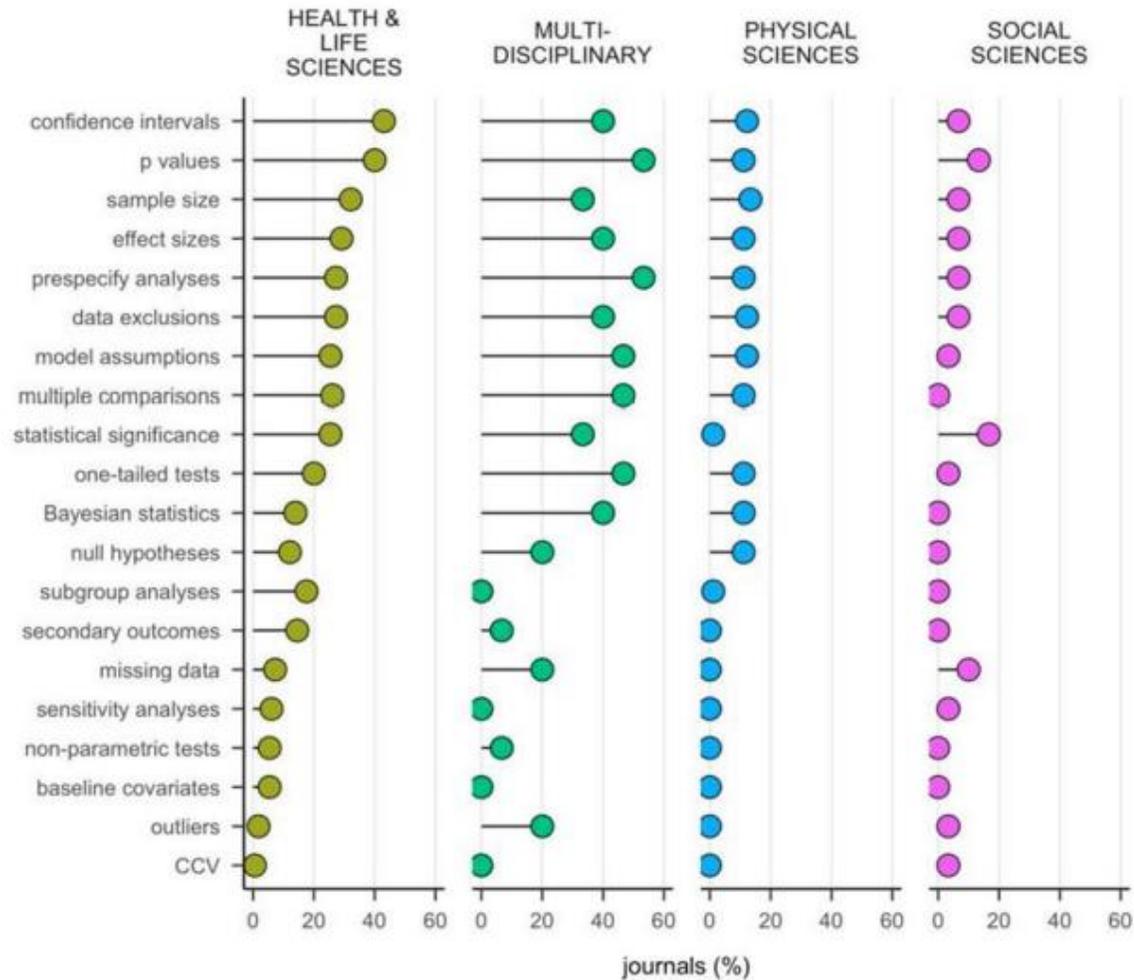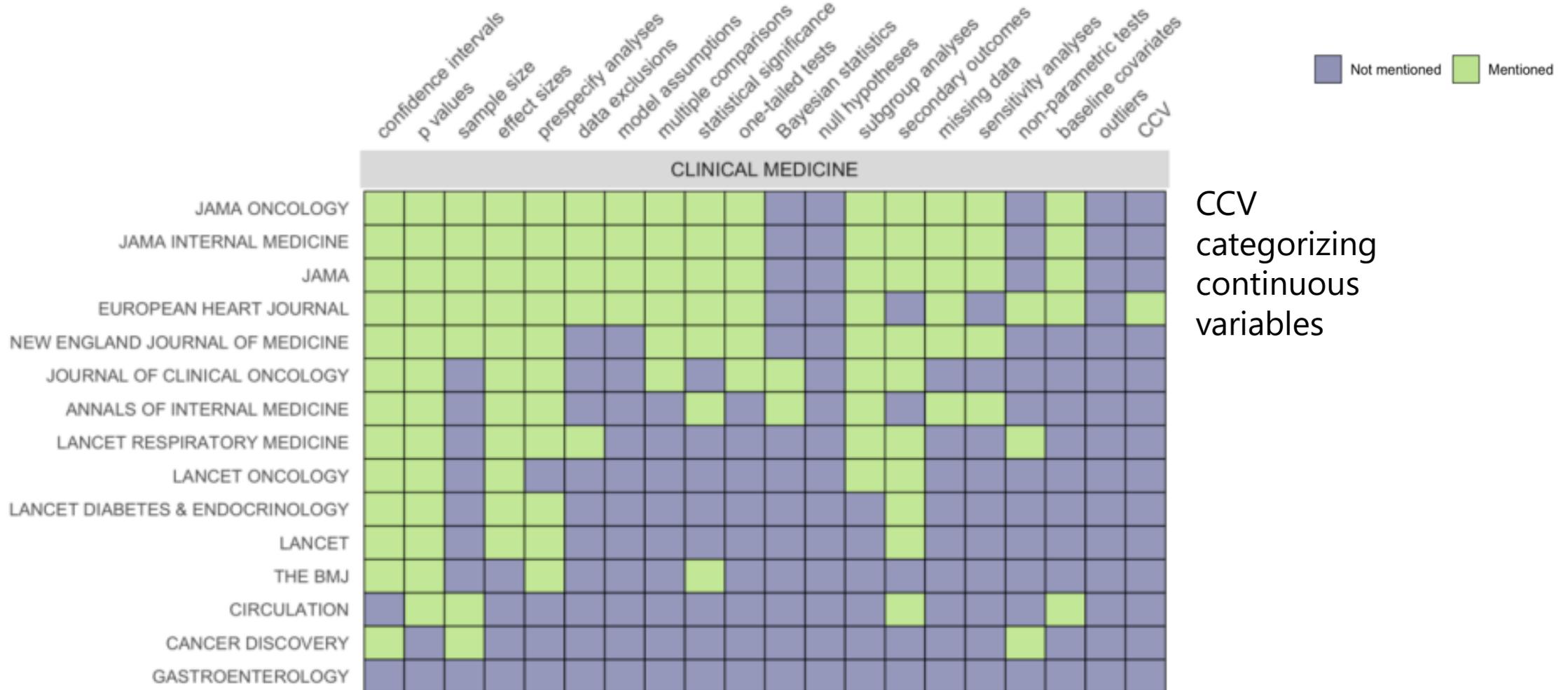
Mentioning of twenty prespecified statistical topics

CCV categorizing continuous variables

Statistical guidance – mainly from reporting guidelines!

Supplementary Material J. External sources of statistical guidance.

Supplementary Table J1. Number of journals referring to specific reporting guidelines.

Reporting guideline	Journals (n)
Consolidated Standards of Reporting Trials (CONSORT)	95
Animal Research: Reporting of In Vivo Experiments (ARRIVE)	80
Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA)	47
REporting recommendations for tumour MARKer prognostic studies (REMARK)	40
The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE)	39
Standards for Reporting Diagnostic accuracy studies (STARD)	36
International Committee of Medical Journal Editors (ICMJE) Recommendations	21
Consolidated Health Economic Evaluation Reporting Standards (CHEERS)	18
STrengthening the REporting of Genetic Association Studies (STREGA)	16
National Institutes of Health (NIH) Recommendations	14
Meta-analysis Of Observational Studies in Epidemiology (MOOSE)	12
CAse REport (CARE)	11
Minimum Information About a Microarray Experiment (MIAME)	11
Consolidated criteria for reporting qualitative research (COREQ)	9
Standard Protocol Items: Recommendations for Interventional Trials (SPIRIT)	9
Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD)	9

Examples of guidance – sufficient details?

Bayesian statistics

"For Bayesian analysis, [report] information on the choice of priors and Markov chain Monte Carlo settings."

(Scientific Data)

Categorisation of continuous data

"Categorizing of continuous data (e.g. into quartiles, quintiles) is discouraged. It leads to a loss of information, usually needs more complicated methods than for continuous data and introduces demarcations which are valid only for this particular study."

(European Heart Journal)

Handling outliers

"How were outliers defined and handled? Were they defined before the beginning of the study? Have you reported outliers that were excluded?…Data pre-processing steps such as transformations, re-coding, re-scaling, normalization, truncation, and handling of below detectable level readings and outliers should be fully described; any removal or modification of data values must be fully acknowledged and justified."

(Science Translational Medicine)

Handling missing data

"Report losses to observation, such as dropouts from a clinical trial or those lost to follow-up or unavailable in an observational study. Consider multiple imputation methods to impute missing data and include an assessment of whether data were missing at random. Approaches based on "last observation carried forward" should not be used."

(JAMA Internal Medicine)