The STRengthening Analytical Thinking for Observational Studies (STRATOS) initiative – introduction and brief overview

Willi Sauerbrei[1]
Michal Abrahamowicz[2], Georg Heinze[3],
Marianne Hübner[4], Ruth Keogh[5]

Medical Center – University of Freiburg, Germany
http://stratos-initiative.org/

[1] Medical Center – University of Freiburg ,Germany [2] McGill University Montreal , Canada [3] Medical University of Vienna, CeMSIIS, Section for Clinical Biometrics, Austria, [4] Michigan State University, USA [5] London School of Hygiene and Tropical Medicine, UK

Aims of the STRATOS virtual conference

- 3 days: 2 Nov, 9 Nov, 14 Dec

- Day 1: To provide an overview of current and future TG activities to
  - Long-term members
  - Newbies
  - Interested guests
- Day 2: To report on and stimulate new cross-TG activities
  - Panels
  - Existing STRATOS projects
  - New cross-TG activities
- Day 3: To follow-up on new STRATOS activities

NEED for GUIDANCE

- Profusion of new, complex statistical techniques and algorithms
- Unclear which methods are useful in practice, and under what conditions?
- Insufficient awareness and understanding, among practitioners, of both well-established and, especially, new approaches and methods
- For some complex analytical challenges, there is *no* consensus, even among experts, as to the best approach
- Very **limited guidance** on key issues that are **vital in practice** discourages analysts from utilizing possibly more appropriate methods in their real-life applications, thus, reducing the scientific yield of empirical research

STRATOS Initiative: STRengthening Analytical Thinking for Observational Studies

- **The overarching long-term goal:**

  To improve design and statistical analyses of observational studies in practice

  by '*closing the gap*' between (i) recent relevant developments in statistical methodology *versus* (ii) methods applied in real-life observational studies

- Specific aims:
  - Develop **evidence-supported guidance** for statistical issues of practical importance (*through discussions among experts with different views, and simulations to systematically assess and compare alternative methods*)
  - Provide **guidance at several levels** of statistical knowledge
  - Start with **state-of-the-art** guidance for issues where there is consensus and necessary evidence
  - **Identify and explore complex analytical challenges requiring more primary research** and/or **combining expertise** in different areas of statistical research

Different levels of statistical knowledge

Level 1: Low statistical knowledge

- Most analyses are done by analysts at this level

Level 2: Experienced statistician

- Methodology perhaps slightly below state of the art, but doable by every experienced analyst

Level 3: Expert in a specific area

- To improve statistical models and to adapt them to complex real problems, researchers develop new and more complex approaches. Advantages and usefulness in practice need to be assessed

Guidance for analysis is needed for many stakeholders (analysts with different levels of knowledge, reviewers, teachers, journalists, ……)

Researchers

First in a Series of Papers for the Biometric Bulletin

STRATOS initiative – Guidance for designing and analyzing observational studies

STRATOS INITIATIVE

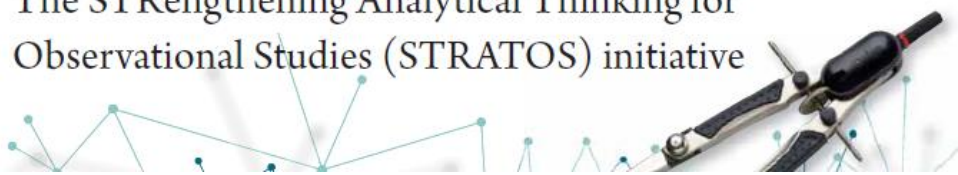
Willi Sauerbrei[1], Marianne Huebner[2] , Gary S. Collins[3], Katherine Lee[4], Laurence Freedman[5], Mitchell Gail[6], Els Goetghebeur[7], Joerg Rahnenfuehrer[8] and Michal Abrahamowicz[9] on behalf of the STRATOS initiative.

➡ Short papers from all TGs and some panels

Consumers

Guidance for designing and analysing observational studies:

The STRengthening Analytical Thinking for Observational Studies (STRATOS) initiative



Willi Sauerbrei[1], Gary S. Collins[2], Marianne Huebner[3], Stephen D. Walter[4], Suzanne M. Cadarette[5], and Michal Abrahamowicz[6] on behalf of the STRATOS initiative

Volume 26 Number 3 | **Medical Writing** September 2017 | 17

Journal of the European Medical Writers Association (EMWA)

STRATOS – History and Milestones

- **2011 Epi Subcom** at 42th Int Soc Clin Biostatistics (ISCB) in Ottawa
- **2013: Initiative launched** at 44th ISCB in Munich
- **2014: 1st STRATOS paper [1]:** *Statistics in Medicine* 2014; 33(30):5413-5432.

  Sauerbrei W, Abrahamowicz M, Altman D, le Saskia, Carpenter J. *STRengthening Analytical Thinking for Observational Studies: The STRATOS initiative.*

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

….. STRATOS – History and Milestones

- **Invited STRATOS Sessions and Mini-Symposia:**
  - Int Soc Clin Biost (ISCB): 2014, 2015, 2016, 2018, 2019, 2020, 2021,2022
  - Int Biometric Conf (IBC): 2016, 2020, 2022 + Regional IBS meetings: 2017, 2018, 2021, 2022
  - Royal Statistical Soc (RSS): 2018, 2020, 2021
  - Other international conferences: HEC 2016, CEN 2018, GMDS 2017, Soc Epi Res (SER) 2021, DAGStat 2022
- **Series in Biometric Bulletin (since 3/2017), 16 articles published, to proceed until 4/2024**
- **2021 Memorandum of Understanding with ISCB**
- **Partner in the SISAQOL project lead by EORTC (>40 stakeholders, including pharma and regulators)**

- **By 2021: >100 members (from 19 countries on 5 continents)**

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), Saskia le Cessie(NL), Carsten Oliver Schmidt (GER)
4	**Measurement error and misclassification**	Laurence Freedman (ISR), Victor Kipnis (US)
5	**Study design**	Mitchell Gail (US), Suzanne Cadarette (CAN)
6	**Evaluating diagnostic tests and prediction models**	Ewout Steyerberg (NL), Ben van Calster (NL)
7	**Causal inference**	Els Goetghebeur (BEL), Ingeborg Waernbaum (SWE)
8	**Survival analysis**	Michal Abrahamowicz (CAN), Per Kragh Andersen (DEN), Terry Therneau (US)
9	**High-dimensional data**	Lisa McShane (US), Joerg Rahnenfuehrer (GER), Riccardo de Bin (NOR)

Chairs from 12 countries and 3 continents

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)
		Co-Chairs:	Mitchell Gail (US), Petra Macaskill (AUS)
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:	Rolf Groenwold (NL), Maarten van Smeden (NL)
CP	**Contact Organisations**	Chairs:	Willi Sauerbrei (GER)
VP	**Visualisation**	Chairs:	Mark Baillie (SWITZ/CH)

TG 2: Selection of variables and functional forms in multivariable analysis

Sauerbrei *et al. Diagnostic and Prognostic Research* (2020) 4:3
https://doi.org/10.1186/s41512-020-00074-3

State of the art in selection of variables and functional forms in multivariable analysis—outstanding issues

Willi Sauerbrei[1*], Aris Perperoglou[2], Matthias Schmid[3], Michal Abrahamowicz[4], Heiko Becher[5], Harald Binder[1], Daniela Dunkler[6], Frank E. Harrell Jr[7], Patrick Royston[8], Georg Heinze[6] and for TG2 of the STRATOS initiative

Table 1 Relevant issues in deriving evidence-supported state of the art guidance for multivariable modelling

No.	Item
1	Investigation and comparison of the properties of variable selection strategies
2	Comparison of spline procedures in both univariable and multivariable contexts
3	How to model one or more variables with a 'spike-at-zero'?
4	Comparison of multivariable procedures for model and function selection
5	Role of shrinkage to correct for bias introduced by data-dependent modelling
6	Evaluation of new approaches for post-selection inference
7	Adaption of procedures for very large sample sizes needed?

Perperoglou *et al. BMC Medical Research Methodology* (2019) 19:46
https://doi.org/10.1186/s12874-019-0666-3

A review of spline function procedures in R

Aris Perperoglou[1*], Willi Sauerbrei[2], Michal Abrahamowicz[3], Matthias Schmid[4] on behalf of TG2 of the STRATOS initiative

See also „Educational work“

TG 4: Measurement error and misclassification

Annals of Epidemiology
Volume 28, Issue 11, November 2018, Pages 821-828

Epidemiologic analyses with error-prone exposures: review of current practice and recommendations

Pamela A. Shaw, PhD [a,*], Veronika Deffner, PhD [b], Ruth H. Keogh, DPhil [c], Janet A. Tooze, PhD [d], Kevin W. Dodd, PhD [e], Helmut Küchenhoff, PhD [b], Victor Kipnis, PhD [e], Laurence S. Freedman, PhD [f,g], on behalf of Measurement Error and Misclassification Topic Group (TG4) of the STRATOS Initiative

TUTORIAL IN BIOSTATISTICS

STRATOS guidance document on measurement error and misclassification of variables in observational epidemiology: Part 1—Basic theory and simple methods of adjustment

Ruth H. Keogh[1] | Pamela A. Shaw[2] | Paul Gustafson[3] | Raymond J. Carroll[4,5] | Veronika Deffner[6] | Kevin W. Dodd[7] | Helmut Küchenhoff[8] | Janet A. Tooze[9] | Michael P. Wallace[10] Victor Kipnis[11] | Laurence S. Freedman[12,13]

Statistics in Medicine (2020).

TUTORIAL IN BIOSTATISTICS

STRATOS guidance document on measurement error and misclassification of variables in observational epidemiology: Part 2—More complex methods of adjustment and advanced topics

Pamela A. Shaw[1] | Paul Gustafson[2] | Raymond J. Carroll[3,4] | Veronika Deffner[5] | Kevin W. Dodd[6] | Ruth H. Keogh[7] | Victor Kipnis[6] | Janet A. Tooze[8] | Michael P. Wallace[9] | Helmut Küchenhoff[5] | Laurence S. Freedman[10,11]

Statistics in Medicine (2020).

See also „Educational work“

TG 1: Missing data

Journal of Clinical Epidemiology
Available online 2 February 2021
In Press, Journal Pre-proof

Framework for the Treatment And Reporting of Missing data in Observational Studies: The TARMOS framework

Lee KJ [1, 2], Tilling K [3], Cornish RP [3], Little RJA [4], Bell ML [5], Goetghebeur E [6], Hogan JW [7], Carpenter JR [8], the STRATOS initiative

TG 3: Initial data analysis

Huebner *et al. BMC Medical Research Methodology* (2020) 20:61
https://doi.org/10.1186/s12874-020-00942-y

Hidden analyses: a review of reporting practice and recommendations for more transparent reporting of initial data analyses

Marianne Huebner[1,2*], Werner Vach[3], Saskia le Cessie[4], Carsten Oliver Schmidt[5], Lara Lusa[6,7] and on behalf of the Topic Group "Initial Data Analysis" of the STRATOS Initiative (STRengthening Analytical Thinking for Observational Studies, http://www.stratos-initiative.org)

Observational Studies 4 (2018) 171-192

A Contemporary Conceptual Framework for Initial Data Analysis

Marianne Huebner huebner@stt.msu.edu
Department of Statistics and Probability
Michigan State University
East Lansing, MI 48824, USA

Saskia le Cessie S.le Cessie@lumc.nl
Department of Clinical Epidemiology and Department of Medical Statistics and Bioinformatics
Leiden University Medical Center
Leiden, The Netherlands

Carsten O. Schmidt Carsten.schmidt@uni-greifswald.de
Institute for Community Medicine, SHIP-KEF
University Medicine of Greifswald
Greifswald, Germany

Werner Vach Werner.vach@usb.ch
Department of Orthopaedics and Traumatology
University Hospital Basel
Basel, Switzerland

on behalf of the Topic Group "Initial Data Analysis" of the STRATOS Initiative (STRengthening Analytical Thinking for Observational Studies, http://www.stratos-initiative.org). Membership of the Topic Group is provided in the Acknowledgments.

TG 5: Study design

Gail MH, *et al. BMJ Open* 2019;**9**:e031031. doi:10.1136/bmjopen-2019-031031

Design choices for observational studies of the effect of exposure on disease incidence

Mitchell H Gail ,[1] Douglas G Altman,[2] Suzanne M Cadarette,[3] Gary Collins,[4] Stephen JW Evans,[5] Peggy Sekula ,[6] Elizabeth Williamson,[7] Mark Woodward[8]

TG 6: Evaluating diagnostic tests and prediction models

Wynants *et al. BMC Medicine* (2019) 17:192
https://doi.org/10.1186/s12916-019-1425-3

Three myths about risk thresholds for prediction models

Laure Wynants[1,2*], Maarten van Smeden[3,4], David J. McLernon[5], Dirk Timmerman[1,6], Ewout W. Steyerberg[4], Ben Van Calster[1,4] and on behalf of the Topic Group 'Evaluating diagnostic tests and prediction models' of the STRATOS initiative

Van Calster *et al. BMC Medicine* (2019) 17:230
https://doi.org/10.1186/s12916-019-1466-7

Calibration: the Achilles heel of predictive analytics

Ben Van Calster[1,2,6*], David J. McLernon[3,6], Maarten van Smeden[2,4,6], Laure Wynants[1,5], Ewout W. Steyerberg[2,6]
On behalf of Topic Group 'Evaluating diagnostic tests and prediction models' of the STRATOS initiative[6]

Gynecologic Oncology Reports 18 (2016) 49–50

Correspondence

Flawed external validation study of the ADNEX model to diagnose ovarian cancer on behalf of TG6 of the STRATOS initiative

TG 7: Causal inference

TUTORIAL IN BIOSTATISTICS

Formulating causal questions and principled statistical answers

Els Goetghebeur[1,2] | Saskia le Cessie[3] | Bianca De Stavola[4] | Erica EM Moodie[5] | Ingeborg Waernbaum[6] | "on behalf of" the topic group Causal Inference (TG7) of the STRATOS initiative

Statistics in Medicine (2020).

TG 8: Survival analysis

TUTORIAL IN BIOSTATISTICS

Analysis of time-to-event for observational studies: Guidance to the use of intensity models

Per Kragh Andersen[1] | Maja Pohar Perme*[2] | Hans C. van Houwelingen[3] | Richard J. Cook[4] | Pierre Joly[5] | Torben Martinussen[1] | Jeremy M.G. Taylor[6] | Michal Abrahamowicz[7] | Terry M. Therneau[8] | for the STRATOS TG8 topic group

Statistics in Medicine (2020).

TG 9: High-dimensional data

TO CHANGE and EXTEND

Simulation panel

Comparison of statistical methods: How?

Simulation studies play a key role

LETTER TO THE EDITOR

Biometrical Journal

On the necessity and design of studies comparing statistical methods

Anne-Laure Boulesteix[1]
Harald Binder[2]
Michal Abrahamowicz[3]
Willi Sauerbrei[2]
for the Simulation Panel of the STRATOS Initiative

See also „Educational work“

TO CHANGE and EXTEND!!!

Educational work

TG 2

PLOS ONE | https://doi.org/10.1371/journal.pone.0241427 December 21, 2020

REGISTERED REPORT PROTOCOL

Systematic review of education and practical guidance on regression modeling for medical researchers who lack a strong statistical background: Study protocol

Paul Bach[1,2,3], Christine Wallisch[1,2,4], Nadja Klein[3], Lorena Hafermann[1,2], Willi Sauerbrei[5], Ewout W. Steyerberg[6], Georg Heinze[4], Geraldine Rauch[1,2]*, for topic group 2 of the STRATOS initiative[¶]

Introduction to statistical simulations in health research

Anne-Laure Boulesteix ,[1] Rolf HH Groenwold,[2,3] Michal Abrahamowicz,[4] Harald Binder,[5] Matthias Briel,[6,7] Roman Hornung,[1] Tim P Morris ,[8] Jörg Rahnenführer,[9] Willi Sauerbrei,[5] for the STRATOS Simulation Panel

SIGNIFICANCE | February 2020

Analysis in an imperfect world

TG 4

When we observe the world, we sometimes make mistakes. **Michael Wallace**, on behalf of the measurement error topic group of the STRATOS Initiative, explains the potentially severe consequences of this often overlooked issue, and how statistics can help bring us back – or at least a little closer – to the truth

Videos, Shiny apps

Summary

- Data and data science becomes more and more important
- Answering questions empirically through data analyses often requires the use of complex methodology. It is important to develop suitable approaches; needs to be done by experts (Level 3)
- Experienced statisticians (Level 2) need to be supported by suitable guidance. There are (too) many approaches (some are useless) available and suitable comparisons are missing
- Better simulation studies are required to assess properties, compare approaches and derive evidence based guidance for practice.
- Suitable educational material is the key to improve analyses at a broad level
- For practically relevant topics we need greater emphasis on development of Level 1 and 2 guidance

SOMETHING about the PROGRAM and the AIMS