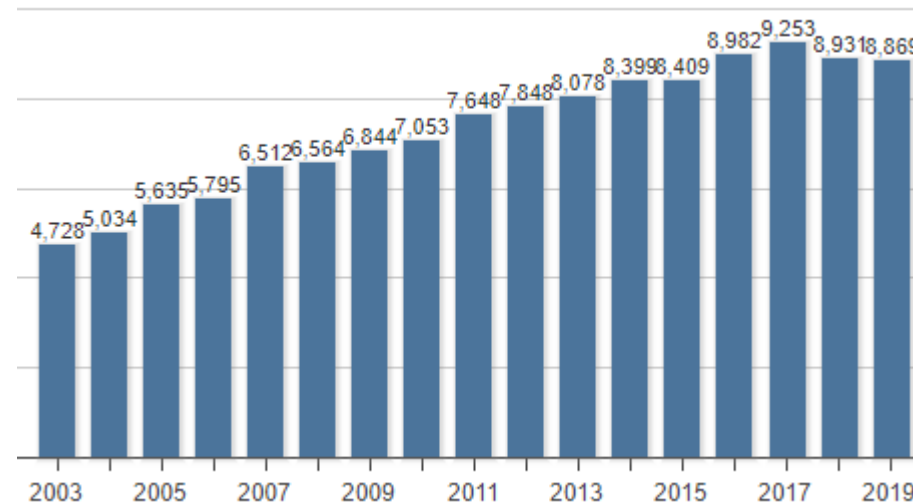# STRATOS initiative – more on guidance for analysis of observational studies

Georg Heinze & Willi Sauerbrei

Statistical research...

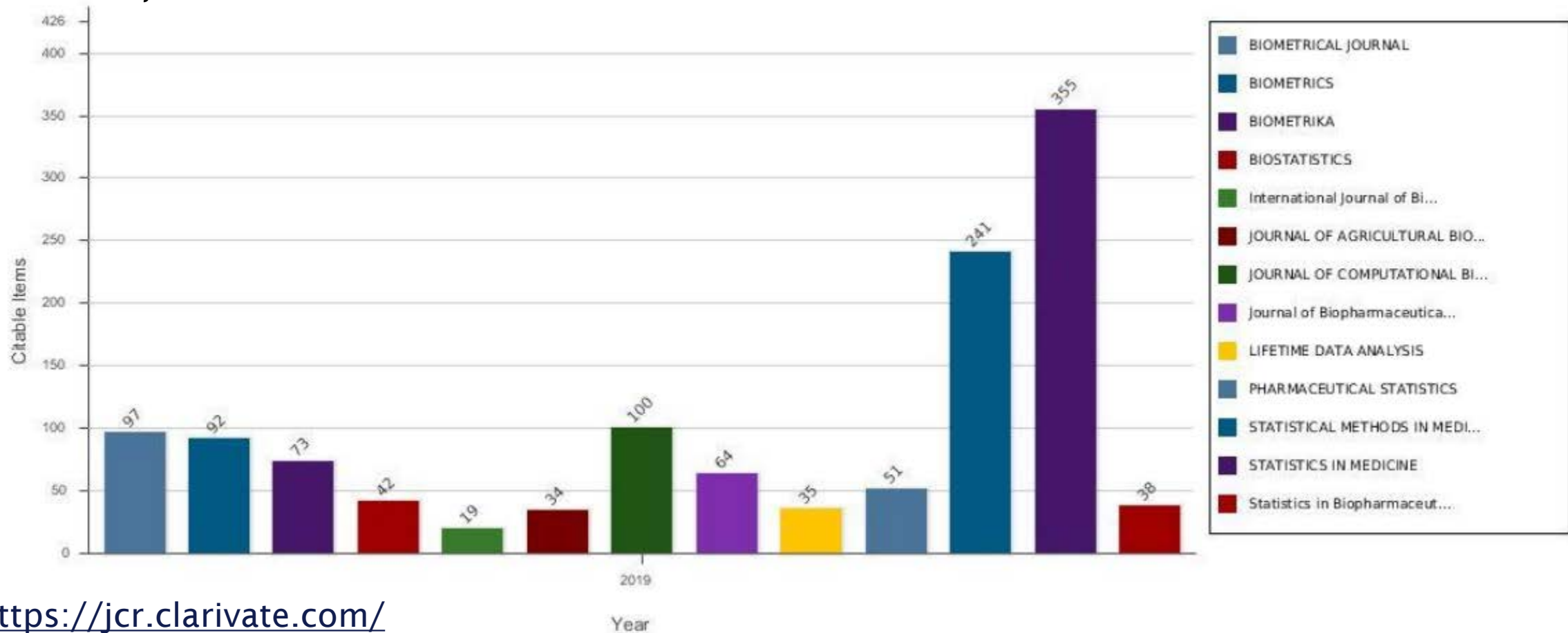- Number of articles per year in 124 ‚Statistics & Probability' journals

https://jcr.clarivate.com/

- Now, every year more than 8,000 new articles
  full of novel methodology are published…

... **bio**statistical research ...

- Restricting to ‚bio*‘, ‚medic*‘, ‚pharma*‘, ‚life*‘ journals:
  1,241 articles in 2019

https://jcr.clarivate.com/

...and how it is uptaken by medical research

Prediction models for diagnosis and prognosis of covid-19: systematic review and critical appraisal

Laure Wynants,[1,2] Ben Van Calster,[2,3] Gary S Collins,[4,5] Richard D Riley,[6] Georg Heinze,[7] Ewoud Schuit,[8,9] Marc M J Bonten,[8,10] Darren L Dahly,[11,12] Johanna A A Damen,[8,9] Thomas P A Debray,[8,9] Valentijn M T de Jong,[8,9] Maarten De Vos,[2,13] Paula Dhiman,[4,5] Maria C Haller,[7,14] Michael O Harhay,[15,16] Liesbet Henckaerts,[17,18] Pauline Heus,[8,9] Nina Kreuzberger,[19] Anna Lohmann,[20] Kim Luijken,[20] Jie Ma,[5] Glen P Martin,[21] Constanza L Andaur Navarro,[8,9] Johannes B Reitsma,[8,9] Jamie C Sergeant,[22,23] Chunhu Shi,[24] Nicole Skoetz,[19] Luc J M Smits,[1] Kym I E Snell,[6] Matthew Sperrin,[25] René Spijker,[8,9,26] Ewout W Steyerberg,[3] Toshihiko Takada,[8] Ioanna Tzoulaki,[27,28] Sander M J van Kuijk,[29] Florien S van Royen,[8] Jan Y Verbakel,[30,31] Christine Wallisch,[7,32,33] Jack Wilkinson,[22] Robert Wolff,[34] Lotty Hooft,[8,9] Karel G M Moons,[8,9] Maarten van Smeden[8]

RESULTS

14 217 titles were screened, and 107 studies describing 145 prediction models were included. The review identified four models for identifying people at risk in the general population; 91 diagnostic models for detecting covid-19 (60 were based on medical imaging, nine to diagnose disease severity); and 50 prognostic models for predicting mortality risk, progression to severe disease, intensive care unit admission, ventilation, intubation, or length of hospital stay. The most frequently reported predictors of diagnosis and prognosis of covid-19 are age, body temperature, lymphocyte count, and lung imaging features. Flu-like symptoms and neutrophil count are frequently predictive in diagnostic models, while comorbidities, sex, C reactive protein, and creatinine are frequent prognostic factors. C index estimates ranged from 0.73 to 0.81 in prediction models for the general population, from 0.65 to more than 0.99 in diagnostic models, and from 0.68 to 0.99 in prognostic models. All models were rated at high risk of bias, mostly because of non-representative selection of control patients, exclusion of patients who had not experienced the event of interest by the end of the study, high risk of model overfitting, and vague reporting. Most reports did not include any description of the study population or intended use of the models, and calibration of the model predictions was rarely assessed.

STRATOS INITIATIVE

STRengthening Analytical Thinking for Observational Studies
Accessible and Accurate Guidance in the Design and Analysis of Observational Studies

Context

- Observational studies pose many design and statistical challenges
- Valid observational research depends on careful study design, high data quality, appropriate statistical methods and accurate interpretation of results

The Problem

- Statistical methods has seen exponential advancements
  - diffusion of methodological innovation is slow
  - many developments are not applied in practice
- Even worse, 'standard' analyses reported in the medical literature are often based on unrealistic assumptions or use inappropriate methods, casting doubt on their results and conclusions
- Analysts, reviewers, editors, readers and many more stakeholders and consumers need guidance for key issues in the design and analysis of observational studies

Filling the Gap: The STRATOS initiative

- The STRengthening Analytical Thinking for Observational Studies (STRATOS) initiative was launched in August 2013
  - In November 2019
    - ~100 researchers from 18 countries worldwide with background in biostatistical and epidemiological methods
    - Connected with leading international organisations, e.g.,
      - International Society of Clinical Biostatistics (ISCB), and
      - International Biometric Society (IBS)
- Website includes member details, publications & resources http://www.stratos-initiative.org/

STRATOS Objectives

- Provide accessible and evidence-based guidance for key topics in the design and analysis of observational studies
- Guidance is intended for applied statisticians and other data analysts with varying levels of statistical education, experience and interests

Organisational Structure

STRATOS is led by an **Executive Committee** and a **Steering Group**. It has the following topic groups and cross-cutting panels:

Topic Groups (TGs)		Panels	
1	Missing data	MP	Membership
2	Selection of variables and functional forms in multivariable analysis	PP	Publications
3	Initial data analysis	GP	Glossary
4	Measurement error and misclassification	WP	Website
5	Study design	RP	Literature Review
6	Evaluating diagnostic tests and prediction models	BP	Bibliography
7	Causal Inference	SP	Simulation Studies
8	Survival analysis	DP	Data Sets
9	High-dimensional data	TP	Knowledge Translation
		CP	Contact Organizations
		VP	Visualisation

Membership:

Regular: Experienced researchers who contribute regularly to STRATOS TG +/- panel activities
Experienced Adjunct: Experienced researchers who contribute occasionally to STRATOS activities
Early Career Adjunct: Researchers at early career stages who participate in TG +/- panel activities
Clinical Affiliates: Clinicians engaged in research with interest and practical experience with statistics

Framework for Guidance Development

Guidance is aimed at users with three levels of statistical knowledge:
Level 1: Low statistical knowledge
  - Propose acceptable methods that are easily implemented
  - Highlight weaknesses of common approaches
Level 2: Experienced statistician
  - Refer to advantages and disadvantages of competing approaches
  - Propose advanced methodology feasible by experienced analysts
Level 3: Expert in a specific area
  - Consider recent developments with statements about possible advantages and disadvantages
  - Identify areas needing more methodological research or guidance

Stages in Guidance Development
Phase I: Experts need to work on state-of-the-art methods and develop guidance for knowledge level 2. Comparisons (simulations) required!
Phase II: Extend guidance to level 1 (simpler methods but still acceptable) and experts work on improved methodology aiming to improve level 2 guidance

Activities & Selected Publications

Important Meetings/Mini-Symposia
- ISCB conferences (2013, 2014, 2015, 2016, 2018, 2019, 2020)
- IBS - Invited sessions at IBC (2016, 2020)
    - Meetings of regional groups of IBS (2017, 2018, 2020)
- Banff International Research Station, Canada (general meetings in 2016, 2019)

Selected Publications
- Sauerbrei et al for the STRATOS initiative. STRengthening Analytical Thinking for Observational Studies: the STRATOS initiative. Stat Med 2014; 33:5413-5432
- Huebner et al for TG3. A contemporary conceptual framework for initial data analysis. Obs Stud 2018; 4: 171-192
- Shaw et al for TG4. Epidemiologic analyses with error-prone exposures: review of current practice and recommendations. Ann Epidemiol 2018; 28: 821–828
- Boulesteix et al for the Simulation Studies panel. On the necessity and design of studies comparing statistical methods. Biom J 2018; 60: 216-218
- **Series of short papers** from the TGs in the Biometric Bulletin of IBS (available on the STRATOS website) provide an overview of TG activities.

How could STRATOS help?

14 217 titles were screened, and 107 studies describing 145 prediction models were included. review identified four models for identifying peopl risk in the general population; 91 diagnostic models for detecting covid-19 (60 were based on medical imaging, nine to diagnose disease severity); and 50 prognostic models for predicting mortality risk, progression to severe disease, intensive care unit admission, ventilation, intubation, or length of hospital stay. The most frequently reported predictors of diagnosis and prognosis of covid-19 are age, body temperature, lymphocyte count, and lung imaging features. Flu-like symptoms and neutrophil count are frequently predictive in diagnostic models, while comorbidities, sex, C reactive protein, and creatinine are frequent prognostic factors. C index estimates ranged from 0.73 to 0.81 in prediction models for the general population, from 0.65 to more than 0.99 in diagnostic models, and from 0.68 to 0.99 in prognostic models. All models were rated at high risk of bias, mostly because of non-representative selection of control patients, exclusion of patients who had not experienced the event of interest by the end of the study, high risk of model overfitting, and vague reporting. Most reports did not include any description of the study population or intended use of the models, and calibration of the model predictions was rarely assessed.

Topic group 9: High-dimensional data

Topic Group 4: Measurement error

Topic Group 5: Study design

Topic Group 1: Missing data

Topic Group 8: Survival analysis

Topic Group 3: Initial data analysis

Topic Group 2: multivariable analysis

Topic Group 6: Evaluating diagnostic tests and prediction models

Topic Group 7: Causal inference

Line-up for the minisymposium

Session: **MS2: STRATOS initiative - more on guidance for analysis of observational studies**

Organisers / Chairs: **Georg Heinze**, *Medical University of Vienna, Austria*

Willi Sauerbrei, *University of Freiburg, Germany*

14.00 - 17.15

	14.00 - 14.10	Georg Heinze	Introduction
MS2.1	14.10 - 14.35	Willi Sauerbrei	Outstanding issues in selection of variables and functional forms in multivariable analysis
MS2.2	14.35 - 15.00	Ben van Calster	Calibration of risk prediction models: making decisions with the lights on or off?
MS2.3	15.00 - 15.25	Helmut Kuechenhoff and Veronika Deffner	Measurement error and misclassification of variables in observational epidemiology - an overview
MS2.4	15.45 - 16.10	Maja Pohar Perme	Analysis of time-to-event for observational studies: Guidance to the use of intensity models
MS2.5	16.10 - 16.35	James Carpenter and Katherine Lee	Framework for the Treatment And Reporting of Missing data in Observational Studies: The TARMOS framework
MS2.6	16.35 - 17.00	Anne-Laure Boulesteix	A replication crisis in methodological research? On the design of comparison studies
	17.00 - 17.15		General discussion

or the minisymposium

Sessio ... OS initiative - more on guidance for analysis of observational studies

irs: **Georg Heinze**, *Medical University of Vienna, Austria*
Willi Sauerbrei, *University of Freiburg, Germany*

14.00

		Georg Heinze	Introduction
MS2.1	14.10 - 14.35	Willi Sauerbrei	Outstanding issues in selection of variables and functional forms in multivariable analysis
MS2.2	14.35 - 15.00	Ben van Calster	Calibration of risk prediction models: making decisions with the lights on or off?
MS2.3	15.00 - 15.25	Helmut Kuechenhoff and Veronika Deffner	Measurement error and misclassification of variables in observational epidemiology - an overview
MS2.4	15.45 - 16.10	Maja Pohar Perme	Analysis of time-to-event for observational studies: Guidance to the use of intensity models
MS2.5	16.10 - 16.35	James Carpenter and Katherine Lee	Framework for the Treatment And Reporting of Missing data in Observational Studies: The TARMOS framework
MS2.6	16.35 - 17.00	Anne-Laure Boulesteix	A replication crisis in methodological research? On the design of comparison studies
	17.00 - 17.15		General discussion

TG 2

Line-up for the minisymposium

Ses... ...OS initiative - more on guidance for analysis of observational studies

...irs: **Georg Heinze**, *Medical University of Vienna, Austria*
Willi Sauerbrei, *University of Freiburg, Germany*

14...

TG 6

		Georg Heinze	Introduction
		Willi Sauerbrei	Outstanding issues in selection of variables and functional forms in multivariable analysis
MS2.2	14.35 - 15.00	Ben van Calster	Calibration of risk prediction models: making decisions with the lights on or off?
MS2.3	15.00 - 15.25	Helmut Kuechenhoff and Veronika Deffner	Measurement error and misclassification of variables in observational epidemiology - an overview
MS2.4	15.45 - 16.10	Maja Pohar Perme	Analysis of time-to-event for observational studies: Guidance to the use of intensity models
MS2.5	16.10 - 16.35	James Carpenter and Katherine Lee	Framework for the Treatment And Reporting of Missing data in Observational Studies: The TARMOS framework
MS2.6	16.35 - 17.00	Anne-Laure Boulesteix	A replication crisis in methodological research? On the design of comparison studies
	17.00 - 17.15		General discussion

Line-up for the minisymposium

Session: **MS2: STRATOS initiative - more on guidance for analysis of observational studies**
Organisers / Chairs: **Georg Heinze**, *Medical University of Vienna, Austria*
Willi Sauerbrei, *University of Freiburg, Germany*

14.00 - 17.1

		Georg Heinze	Introduction
MS2.1		Willi Sauerbrei	Outstanding issues in selection of variables and functional forms in multivariable analysis
MS2.2		Ben van Calster	Calibration of risk prediction models: making decisions with the lights on or off?
MS2.3	15.00 - 15.25	Helmut Kuechenhoff and Veronika Deffner	Measurement error and misclassification of variables in observational epidemiology - an overview
MS2.4		Maja Pohar Perme	Analysis of time-to-event for observational studies: Guidance to the use of intensity models
MS2.5		James Carpenter and Katherine Lee	Framework for the Treatment And Reporting of Missing data in Observational Studies: The TARMOS framework
MS2.6		Anne-Laure Boulesteix	A replication crisis in methodological research? On the design of comparison studies
	17.00 - 17.15		General discussion

TG 4

TG 4

Line-up for the minisymposium

Session: **MS2: STRATOS initiative - more on guidance for analysis of observational studies**

Organisers / Chairs: **Georg Heinze**, *Medical University of Vienna, Austria*
Willi Sauerbrei, *University of Freiburg, Germany*

14.00 - 17.15

	14.00 - 14.10	Georg Heinze	Introduction
MS		Willi Sauerbrei	Outstanding issues in selection of variables and functional forms in multivariable analysis
MS		Ben van Calster	Calibration of risk prediction models: making decisions with the lights on or off?
MS		Helmut Kuechenhoff and Veronika Deffner	Measurement error and misclassification of variables in observational epidemiology - an overview
MS2.4	15.45 - 16.10	Maja Pohar Perme	Analysis of time-to-event for observational studies: Guidance to the use of intensity models
MS2.5	16.10 - 16.35	James Carpenter and Katherine Lee	Framework for the Treatment And Reporting of Missing data in Observational Studies: The TARMOS framework
MS2.6	16.35 - 17.00	Anne-Laure Boulesteix	A replication crisis in methodological research? On the design of comparison studies
	17.00 - 17.15		General discussion

TG 8

Line-up for the minisymposium

Session: **MS2: STRATOS initiative - more on guidance for analysis of observational studies**
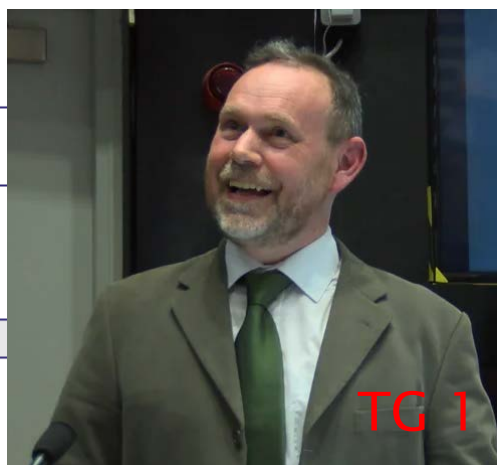
Organisers / Chairs: **Georg Heinze**, *Medical University of Vienna, Austria*

Willi Sauerbrei, *University of Freiburg, Germany*

14.00 - 17.15

	14.00 - 14.10	Georg Heinze	Introduction
			Outstanding issues in selection of variables and functional forms in multivariable analysis
			Calibration of risk prediction models: making decisions with the lights on or off?
			Measurement error and misclassification of variables in observational epidemiology - an overview
			Analysis of time-to-event for observational studies: Guidance to the use of intensity models
MS2.5	16.10 - 16.35	James Carpenter and Katherine Lee	Framework for the Treatment And Reporting of Missing data in Observational Studies: The TARMOS framework
MS2.6	16.35 - 17.00	Anne-Laure Boulesteix	A replication crisis in methodological research? On the design of comparison studies
	17.00 - 17.15		General discussion

TG 1   TG 1

Line-up for the minisymposium

Session: **MS2: STRATOS initiative - more on guidance for analysis of observational studies**

Organisers / Chairs: **Georg Heinze**, *Medical University of Vienna, Austria*
Willi Sauerbrei, *University of Freiburg, Germany*

14.00 - 17.15

14.00 - 14.10	Georg Heinze	Introduction
	Willi Sauerbrei	Outstanding issues in selection of variables and functional forms in multivariable analysis
	Ben van Calster	Calibration of risk prediction models: making decisions with the lights on or off?
	Helmut Kuechenhoff and Veronika Deffner	Measurement error and misclassification of variables in observational epidemiology - an overview
	Maja Pohar Perme	Analysis of time-to-event for observational studies: Guidance to the use of intensity models
	James Carpenter and Katherine Lee	Framework for the Treatment And Reporting of Missing data in Observational Studies: The TARMOS framework
MS2.6	16.35 - 17.00	Anne-Laure Boulesteix	A replication crisis in methodological research? On the design of comparison studies
	17.00 - 17.15	General discussion

Simulation Panel