

TG2: Recent activities of the group

2 November 2021

Aris Perperoglou

Website: stratos-tg2.org

The screenshot shows a web browser window with the address bar displaying "stratos-tg2.org". The website header includes the "STRATOS INITIATIVE" logo and "Topic Group 2" on the left, and navigation links for "Home", "Members", "Activities", and "Resources" on the right. The main content area features a large blue title: "Selection of Variables and Functional Forms in Multivariable Analysis". Below the title are two columns: "Aim" and "Focus". The "Aim" section states: "Derive guidance for variable and function selection in multivariable analysis." The "Focus" section states: "The main focus of TG2 is to identify influential variables and gain insight into their individual and joint relationship with the outcome. Two of the (interrelated) main challenges are selection of variables for inclusion in a multivariable explanatory model, and choice of the functional forms for continuous variables". At the bottom of the page, there is a small information icon on the left and the text "STRATOS Initiative" in the center.

← → ↻ stratos-tg2.org ☆ 🔴 📺 ⬆️ ⚙️ 👤 ⋮

STRATOS INITIATIVE Topic Group 2 Home Members Activities Resources ▾ 🔍

Selection of Variables and Functional Forms in Multivariable Analysis

Aim

Derive guidance for variable and function selection in multivariable analysis.

Focus

The main focus of TG2 is to identify influential variables and gain insight into their individual and joint relationship with the outcome. Two of the (interrelated) main challenges are selection of variables for inclusion in a multivariable explanatory model, and choice of the functional forms for continuous variables

📘 STRATOS Initiative

Georg Heinze (chair): Institute of Clinical Biometrics, Center for Medical Statistics, Informatics and Intelligent Systems, Medical University of Vienna, Austria. https://www.meduniwien.ac.at/researcher/georg_heinze

Aris Perperoglou (chair): Data Science & AI, AstraZeneca R&D Biopharmaceuticals , Cambridge UK

Willi Sauerbrei (chair): Institute for Medical Biometry and Medical Informatics, Medical Center - University of Freiburg, Germany

Michal Abrahamowicz: Faculty of Medicine, McGill University, Canada

Heiko Becher: Institute for Medical Biometry and Epidemiology, University Medical Center Hamburg-Eppendorf, Hamburg, Germany

Harald Binder: Institute for Medical Biometry and Medical Informatics, Medical Center - University of Freiburg, Germany

Daniela Dunkler: Institute of Clinical Biometrics, Center for Medical Statistics, Informatics and Intelligent Systems, Medical University of Vienna, Austria

Rolf Groenwold: Leiden University Medical Centre, Leiden, The Netherlands

Frank Harrell: School of Medicine, Vanderbilt University Medical Center, USA

Nadja Klein: School of Business and Economics , Humboldt-University of Berlin, Germany

Geraldine Rauch: Institute of Biometry and Clinical Epidemiology, Charité – Universitätsmedizin Berlin , Germany

Patrick Royston: Department of Statistical Science, UCL, UK

Matthias Schmid: Institute for Medicine Biometry, Informatics and Epidemiology, Bonn, Germany

Early career adjunct members

Michael Kammer: Institute of Clinical Biometrics, Center for Medical Statistics, Informatics and Intelligent Systems, Medical University of Vienna, Austria

Edwin Kipruto: Institute for Medical Biometry and Medical Informatics, Medical Center - University of Freiburg, Germany

Kim Luijken Leiden University Medical Centre, Leiden, The Netherlands

Christine Wallisch: Institute of Clinical Biometrics, Center for Medical Statistics, Informatics and Intelligent Systems, Medical University of Vienna, Austria

Publications

Perperoglou A, Heinze G, Sauerbrei W on behalf of STRATOS TG2 (2018): STRengthening Analytical Thinking for Observational Studies (STRATOS): Introducing the Topic Group on Selection of Variables and Functional Forms in Multivariable Analysis (TG2). Biometric Bulletin; 35(3):18-19.

Perperoglou A, Sauerbrei W, Abrahamowicz M, Schmid M on behalf of TG2 of the STRATOS initiative (2019): A review of spline function procedures in R. BMC Medical Research Methodology (19:46). DOI: 10.1186/s12874-019-0666-3

Sauerbrei W, Perperoglou A, Schmid M, Abrahamowicz M, Becher H, Binder H, Dunkler D, Harrell Jr. FE, Royston P, Heinze G for TG2 of the STRATOS initiative (2020). State of the art in selection of variables and functional forms in multivariable analysis - outstanding issues. Diagnostic and Prognostic Research, 4:3, 1-18.

Bach P, Wallisch C, Klein N, Hafermann L, Sauerbrei W, Steyerberg EW, Heinze G, Rauch G, for topic group 2 of the STRATOS initiative (2020): Systematic review of education and practical guidance on regression modeling for medical researchers who lack a strong statistical background: Study protocol. Plos One. <https://doi.org/10.1371/journal.pone.0241427>

Heinze G, Perperoglou A, Sauerbrei W on behalf of Topic Group 2 of the STRATOS initiative (2021): STRengthening Analytical Thinking for Observational Studies (STRATOS): Recent activities of the Topic Group on Selection of Variables and Functional Forms in Multivariable Analysis (TG2). Biometric Bulletin; 38(2):7-8.

Public (virtual) appearances in 2021

- September 20-29, 2021 Invited Session at the IBS 2021, Milan, Italy (Virtual), organized by Federico Ambrogi and Willi Sauerbrei
- September 7-10, 2021 Invited Session at the IBS-ROeS Conference 2021, Salzburg, Austria, organized by Mark Baillie and Georg Heinze
(Workshop on variable selection)
- September 6 - 9, 2021 Invited Session at RSS, Manchester, UK, organized by Aris Perperoglou
(Session on multivariable modelling)
- July 18 - 22, 2021 Mini-Symposium at ISCB 42, Lyon, France

TG2: Overview paper



Diagnostic and
Prognostic Research

[Diagn Progn Res.](#) 2020; 4: 3.

PMCID: [PMC7114804](#)

Published online 2020 Apr 2. doi: [10.1186/s41512-020-00074-3](#)

PMID: [32266321](#)

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

- 7 methodological issues identified

Research required

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. Adaptation of procedures for **very large sample sizes** needed?

Projects update:

3.20	General Intro on TG2		
3.23	Level-1 Videos	Guidance Level-1	Rolf Groenwold
3.26	Shiny-app lessons learned	Guidance Level-1	Christine Wallisch
3.29	Variable selection workshop	Guidance Level-2	Georg Heinze
3.32	Splines-project	Research	Daniela Dunckler
3.35	Literature Review of model building in Covid-19 prediction models	Research	Georg Heinze
3.38	Collaboration with other TGs (TG3, TG6, TG7)	Guidance and Research	Georg Heinze
3.40	Discussion		



Leids Universitair
Medisch Centrum

Using videos for guidance

STRATOS meeting November 2021

STRATOS
INITIATIVE

Rolf H.H. Groenwold

TG 2 / knowledge translation panel



“To provide accessible and accurate guidance in the design and analysis of observational studies. The guidance is intended for applied statisticians and other data analysts with varying levels of statistical education, experience and interests.”

STRATOS videos ?

STRATOS
INITIATIVE



How things started

- Short message (70 secs = ~250 words)
- Aimed at clinicians
- Animated powerpoint + voice-over
- Approx. 25 different topics
 - P-value
 - Confidence intervals
 - Publication bias
 - Confounding
 - Regression to the mean
 - Missing data
 - etc

How things started

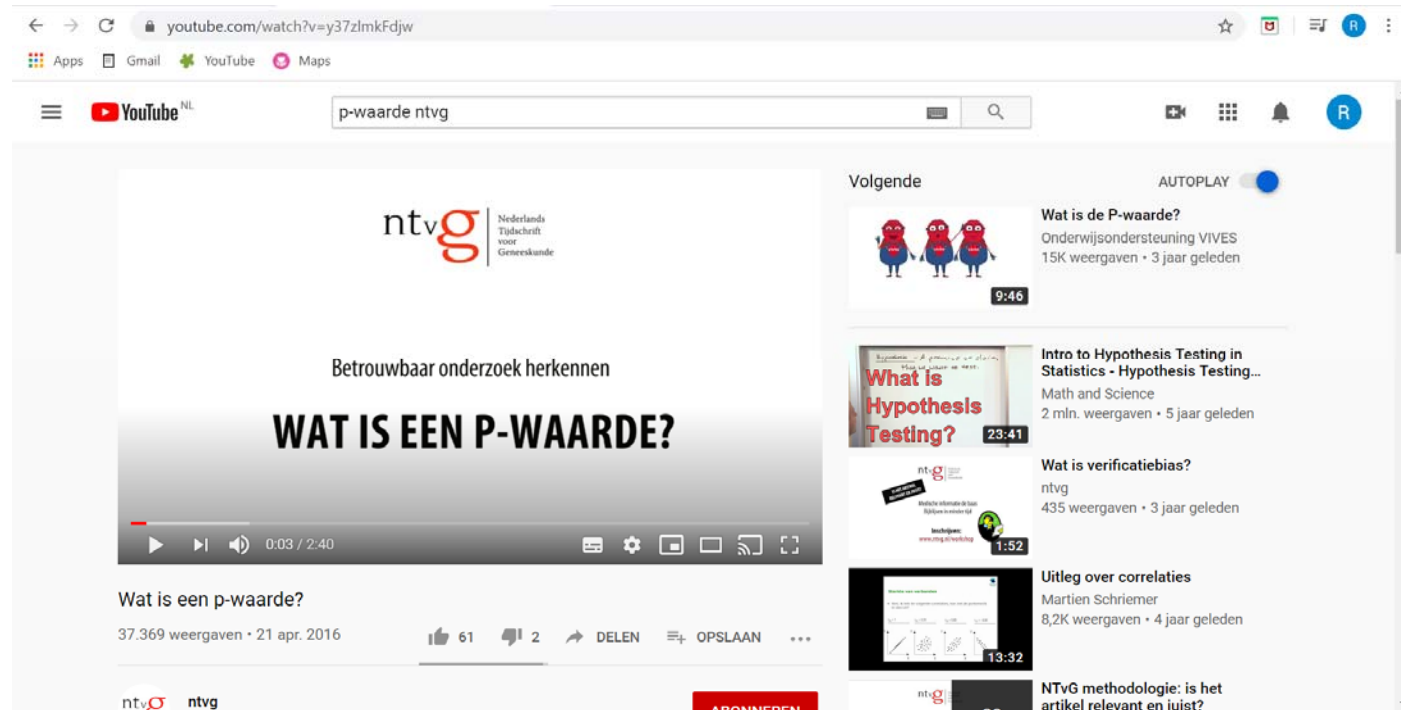
Views:

P-value: >56K

Confidence interval: >32K

Confounding: >17K

(note, videos are in Dutch)



The screenshot shows a YouTube video player interface. The main video is titled "Wat is een p-waarde?" by the channel "ntv g" (Nederlands Tijdschrift voor Geneeskunde). The video has 37,369 views and was uploaded on April 21, 2016. The video player shows the video has just started (0:03 / 2:40). To the right of the video player, there is a "Volgende" (Next) section with an "AUTOPLAY" toggle. The recommended videos are:

- "Wat is de P-waarde?" by Onderwijssteuning VIVES, 15K views, 3 years ago, 9:46 duration.
- "Intro to Hypothesis Testing in Statistics - Hypothesis Testing..." by Math and Science, 2 min. views, 5 years ago, 23:41 duration.
- "Wat is verificatiebias?" by ntv g, 435 views, 3 years ago, 1:52 duration.
- "Uitleg over correlaties" by Martien Schriemer, 8,2K views, 4 years ago, 13:32 duration.
- "NTvG methodologie: is het artikel relevant en juist?" by ntv g, 0 views, 0 years ago, 0:00 duration.

Youtube search

statistics - YouTube

youtube.com/results?search_query=statistics

YouTube NL

statistics

SIGN IN

FILTERS

- Home
- Explore
- Subscriptions
- Library
- History

TEACH ME statistics IN 30 MINS 42:09

Teach me STATISTICS in half an hour!
876K views • 2 years ago
zedstatistics

THE CHALLENGE: "teach me statistics in half an hour with no mathematical formula" The RESULT: an intuitive overview of ...
CC

Statistics made easy 12:50

Statistics made easy !!! Learn about the t-test, the chi square test, the p value and more
966K views • 2 years ago
Global Health with Greg Martin

Learning statistics doesn't need to be difficult. This introduction to stats will give you an understanding of how to apply statistical ...

Statistics Khan Academy

Statistics: The average | Descriptive statistics | Probability and Statistics | Khan Academy • 12:35
Statistics: Sample vs. Population Mean • 6:42
VIEW FULL PLAYLIST

CBSE IX ONE SHOT STATISTICS

STATISTICS in One Shot (Complete Chapter) CBSE Class 9 Math Chapter 14 [Term 1 Exam] NCERT Vedantu
Scheduled for 9/4/21, 9:30 AM
Vedantu Class 9 & 10

This session brings you a Statistics in One Shot (Full Chapter) on CBSE Class 9 Math Chapter 14 to revise the important

Statistics - A Full University Course on Data Science Basics

youtube.com/watch?v=xpc-HPKN28

YouTube NL

statistics course

Section 1.1
What is Statistics?

0:03 / 8:15:04 • What is statistics

23K 295 SHARE SAVE

962,586 views • Jun 12, 2019

Excel in math and science

Learn To Think brilliant.org

Deze cryptomunten gaan knallen
Wees de massa voor met deze Crypto knallers. Download snel deze gratis Crypto gids.
marktgevoel.nl
VISIT SITE

TEACH ME statistics IN 30 MINS 42:09

Teach me STATISTICS in half an hour!
zedstatistics
876K views • 2 years ago

Statistics and Probability Course || Statistics For Data Science
Geek's Lesson
420K views • 11 months ago

STRATOS videos ?

Idea:

- Short message (< 3 minutes)
- Aimed at level 1 audience

Format:

- Short
- Informal
- Motion / continuous visual flow

First Topics:

- Categorisation of continuous predictors (TG 2)
- Modelling continuous predictors (TG 2)
- Measurement error (TG 4)

Our experiences

1. Think about the audience!! (and ask for feedback throughout development process)
2. Script everything
3. The shorter the better
4. Don't mind redoing your recordings (*another reason why you want to keep it short ;-)*)
5. Don't underestimate time investment
6. It's fun and rewarding and even apparently simple topics may spark a discussion

Next steps

The screenshot shows a YouTube search results page for the query 'statistics'. The browser address bar shows 'youtube.com/results?search_query=statistics'. The page features a sidebar with navigation options: Home, Explore, Subscriptions, Library, and History. The main content area displays several search results, each featuring the 'STRATOS INITIATIVE' logo. The results include:

- Teach me STATISTICS in half an hour!** by zedstatistics (876K views, 2 years ago). Description: 'THE CHALLENGE: "teach me statistics in half an hour with no mathematical formula" The RESULT: an intuitive overview of ...'
- Statistics made easy !!! Learn about the t-test, the chi square test, the p value and more** by Global Health with Greg Martin (966K views, 2 years ago). Description: 'Learning statistics doesn't need to be difficult. This introduction to stats will give you an understanding of how to apply statistical ...'
- Statistics** by Khan Academy. Playlist including 'Statistics: The average | Descriptive statistics | Probability and Statistics | Khan Academy' (12:35) and 'Statistics: Sample vs. Population Mean' (6:42).
- STATISTICS in One Shot (Complete Chapter) CBSE Class 9 Math Chapter 14 [Term 1 Exam] NCERT Vedantu** by Vedantu Class 9 & 10. Scheduled for 9/4/21, 9:30 AM.

r.h.h.groenwold@lumc.nl

STRATOS
INITIATIVE



The End

Review of guidance on regression modeling

- Aspects: general, functional forms, variable selection
- 23 out of 47 series of statistical tutorials were selected
→ 57 topic-relevant articles
- Methodological gaps:
 - Detailed info on non-linear modeling methods
 - Pros & cons for variable selection methods
 - Software and code
- Currently under revision
- Protocol paper was published in PlosOne (Bach et. al, 2020)

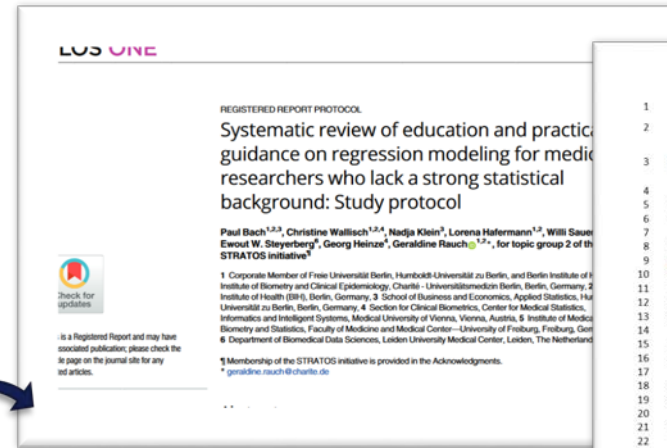
Regression Modeling in Medical Statistics
Review of Statistical Series

Case Report Form - Article Screening

Name of rater: _____
Date of rating: _____

Journal: _____
Statistical series: _____
Title of the article: _____
Rank of the article within the series: _____
Author(s) of the article: _____
Year of publication of the article: _____

Aspect number	Relevant? (y/n)	Extent of explanation (N/A/L)	Example provided? (y/n)	Software advice given? (y/n)	Recommendation given? (y/n)	Warning issued? (y/n)	Base comment? (y/n)
1							
1.1							
1.2							
1.3							
1.4							
1.5							
1.6							
1.7							
1.8							
2							
2.1							
2.2							
2.3							
2.4							
2.5							
2.6							
2.7							
2.8							
2.9							
2.10							
2.11							
2.12							
2.13							
2.14							
2.15							
2.16							



- Systematic review of guidance papers on regression modeling in statistical series of medical journals
- Short title: Systematic review of guidance on regression modeling
- Christine Wallisch^{1,2*}, Paul Bach^{1,3}, Lorena Hafermann^{1,3}, Nadja Klein³, Willi Sauerbrei⁴, Ewout W. Steyerberg⁵, Georg Heinze⁶, Geraldine Rauch^{7,8,9}, on behalf of topic group 2 of the STRATOS initiative
- ¹Charité - Universitätsmedizin Berlin, corporate member of Freie Universität Berlin and Humboldt-Universität zu Berlin, Institute of Biometry and Clinical Epidemiology, Charitéplatz 1, 10117 Berlin, Germany
- ²Medical University of Vienna, Center for Medical Statistics, Informatics and Intelligent Systems, Section for Clinical Biometrics, Spitalgasse 23, 1090 Vienna, Austria
- ³Humboldt-Universität zu Berlin, School of Business and Economics, Emmy Noether Group in Statistics and Data Science, Unter den Linden 6, 10099 Berlin, Germany
- ⁴University of Freiburg, Faculty of Medicine and Medical Center, Institute of Medical Biometry and Statistics, Freiburg, Germany
- ⁵Leiden University Medical Center, Department of Biomedical Data Sciences, Leiden, the Netherlands
- *Corresponding author's email: christine.wallisch@medunivie.ac.at (ORCID: 0000-0003-3943-6234)



BEND YOUR (SP)LINE

by Christine Wallisch, Lena Jiricka, Daniela Dunkler, Georg Heinze

And many thanks to all colleagues for their feedback
during the development phase!

Fractional Polynomials B-Splines Natural Splines

INPUT PARAMETERS

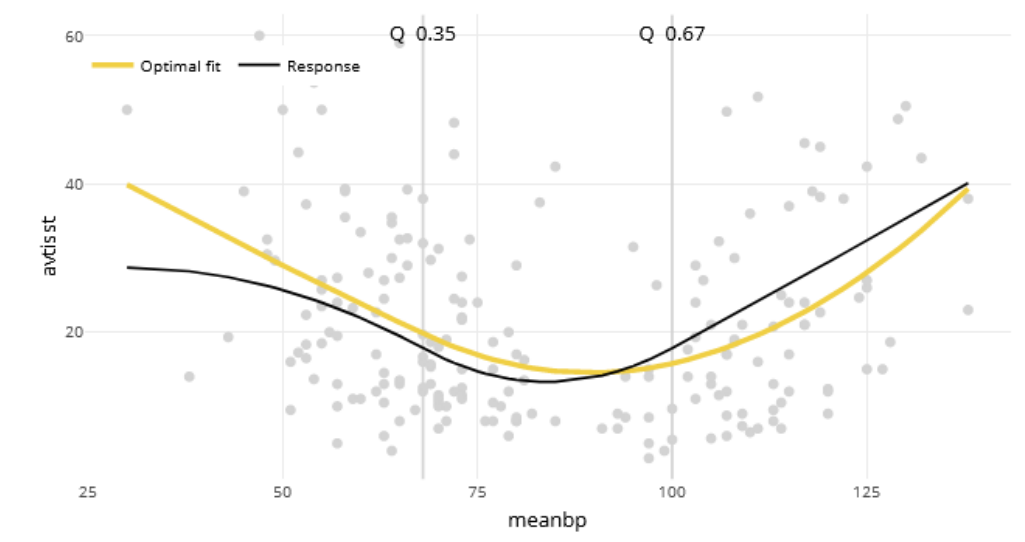
Degree:
 Number of knots:
 Coefficient range:

Position of knot 1:
 Position of knot 2:

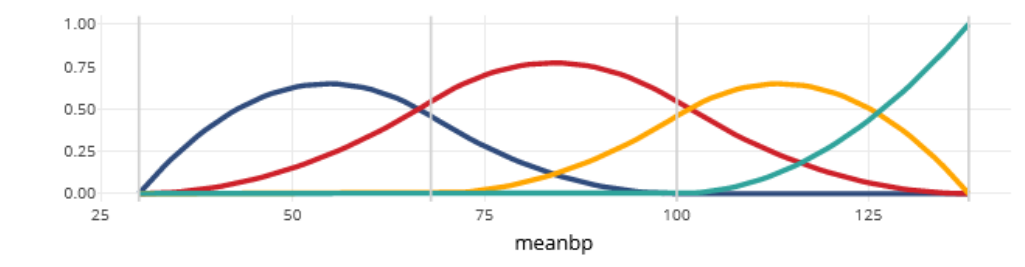
Intercept:

Data points
 LOESS Smoother
 Knot position
 Optimal fit

RESPONSE FUNCTION



SPLINE BASIS FUNCTIONS








EXERCISE

GOODNESS OF FIT

0.104	0.071
R^2	adjusted R^2
0.139	12.053
maximal adj. R^2	prediction error

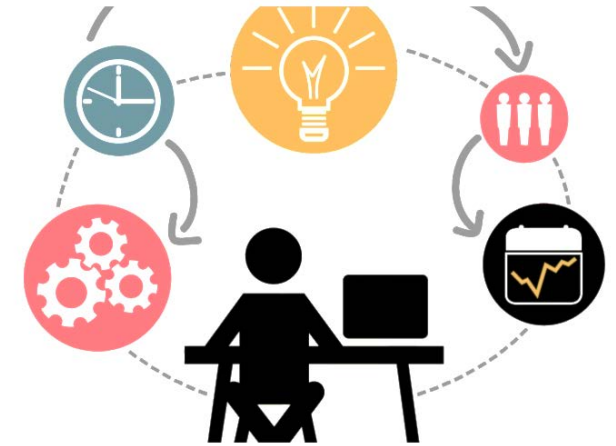
CODE

Results of a pilot study

- Short questionnaire to PhD-students of Georg Heinze
- Valuable feedback from 6 students
- Median time spent exploring the app: 25 min (range: 15-100 min)
- Main insights:
 - Feedback to the exercises: clear and moderately difficult
 - General feedback:
 -  Look and feel
 -  Content
 -  Functionality
 -  Crashes
 -  Speed
- Recommendations for improvement of the layout

Lessons learned

- Keep such apps easy and simple
 - High flexibility → app gets slow and usage gets difficult because of complexity
- Predefine clear purpose, structure and content of the app
 - Structural changes during the development process
 - Affect the architecture of the application
 - May be tedious to implement
 - Required coding time was difficult to estimate
- Full-stack development knowledge would be beneficial



- Interactive app
 - Pretty visualization of functional forms is possible
 - Various parameters can be set
 - Implemented exercises with immediate feedback
 - Reproducibility, e.g. code can be downloaded, open source (github)
 - Easily extendable by other methods

IBS-ROeS conference 2021 pre-conference workshop


Variable selection – a review and recommendations for the practicing statistician

Biometrical Journal →

DOI: 10.1002/bimj.201700067

REVIEW ARTICLE

Variable selection – A review and recommendations
for the practicing statistician

Georg Heinze  | Christine Wallisch | Daniela Dunkler

TG2-endorsed paper

Georg Heinze, Christine Wallisch & Daniela Dunkler
Medical University of Vienna
CeMSIIS – Section for Clinical Biometrics
For TG2 of the STRATOS initiative



georg.heinze@meduniwien.ac.at, christine.wallisch@meduniwien.ac.at, daniela.dunkler@meduniwien.ac.at

Aims of the lecture

- To explain aspects of **variable selection in multivariable regression** analyses of **observational studies**.
- To review **different variable selection strategies and modeling philosophies**.
- To encourage investigations of **model instability** induced by variable selection.
- To illustrate the urgent **need for background knowledge** in statistical modeling.

Interactive presentation of simulation study

~/1Projects/CW 2021-05 VariableSelection/VariableSelection/shinyapp/SimViz - Shiny

http://127.0.0.1:6337 Open in Browser

Publish

Visualization of simulation results: Comparison of variable selection methods

Create plot

Select methods to compare

- Full model (FU)
- Univariate selection, alpha = 0.05 (Uni_005)
- Univariate selection, alpha = 0.20 (Uni_020)
- Forward selection, AIC (FS_AIC)
- Backward elimination, alpha = 0.05 (BE_005)
- Backward elimination, AIC (BE_AIC)
- Augmented backward elimination, alpha = 0.20 (ABE_020)
- Augmented backward elimination, AIC (ABE_AIC)
- Full model approximation (FU_approx)
- Lasso
- Relaxed Lasso (RLasso)

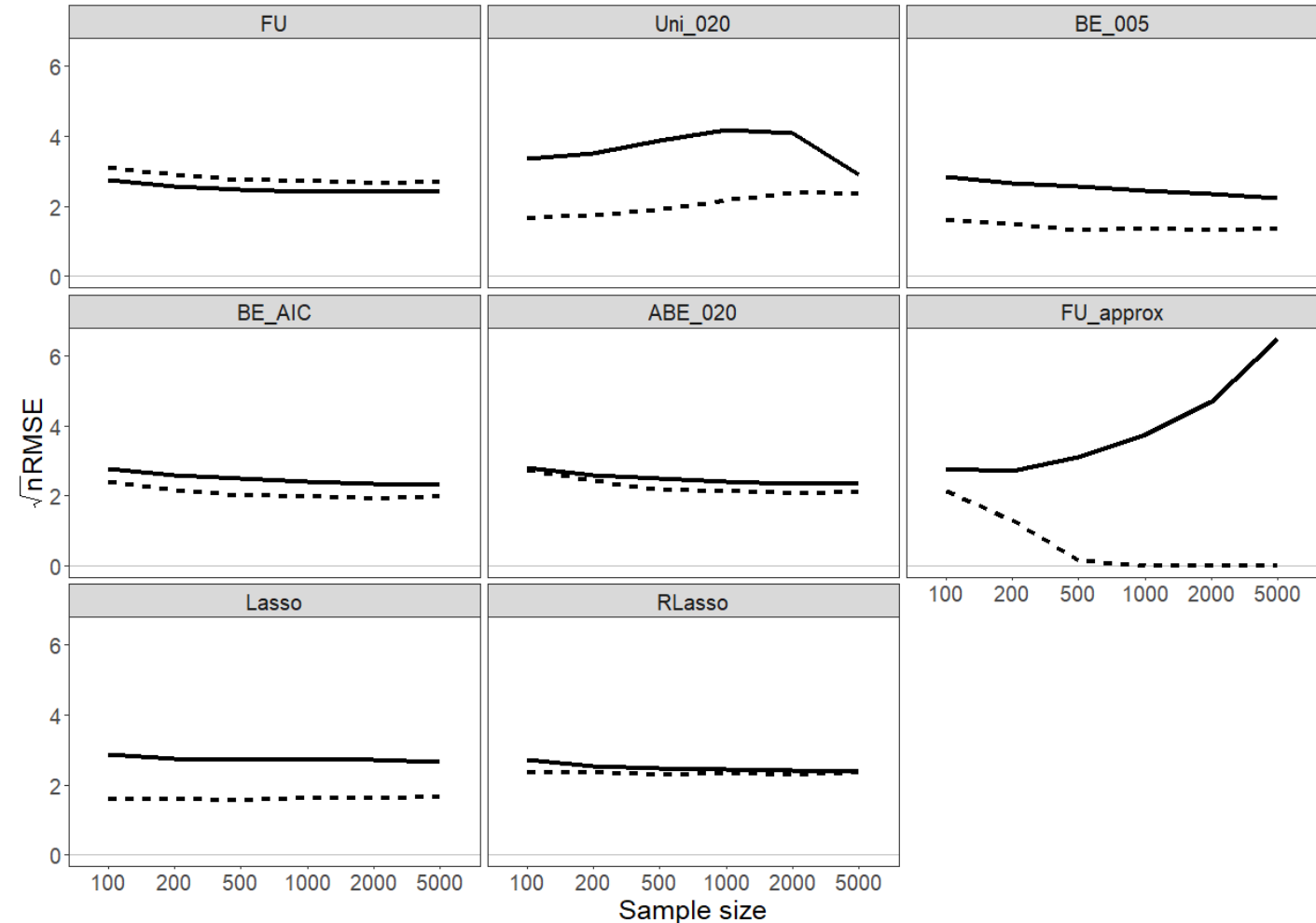
Select the performance measure

- Bias
- RMSE*sqrt(n)
- Coverage
- CI width
- Power/type-1 error
- Selection probability (TPR and FPR)
- True/biased/unbiased model selection rate
- Local bias of y
- Local RMSE of $y * \sqrt{n}$

- Calculate measures conditional on selection
- Average results of predictors and non-predictors

Update plot

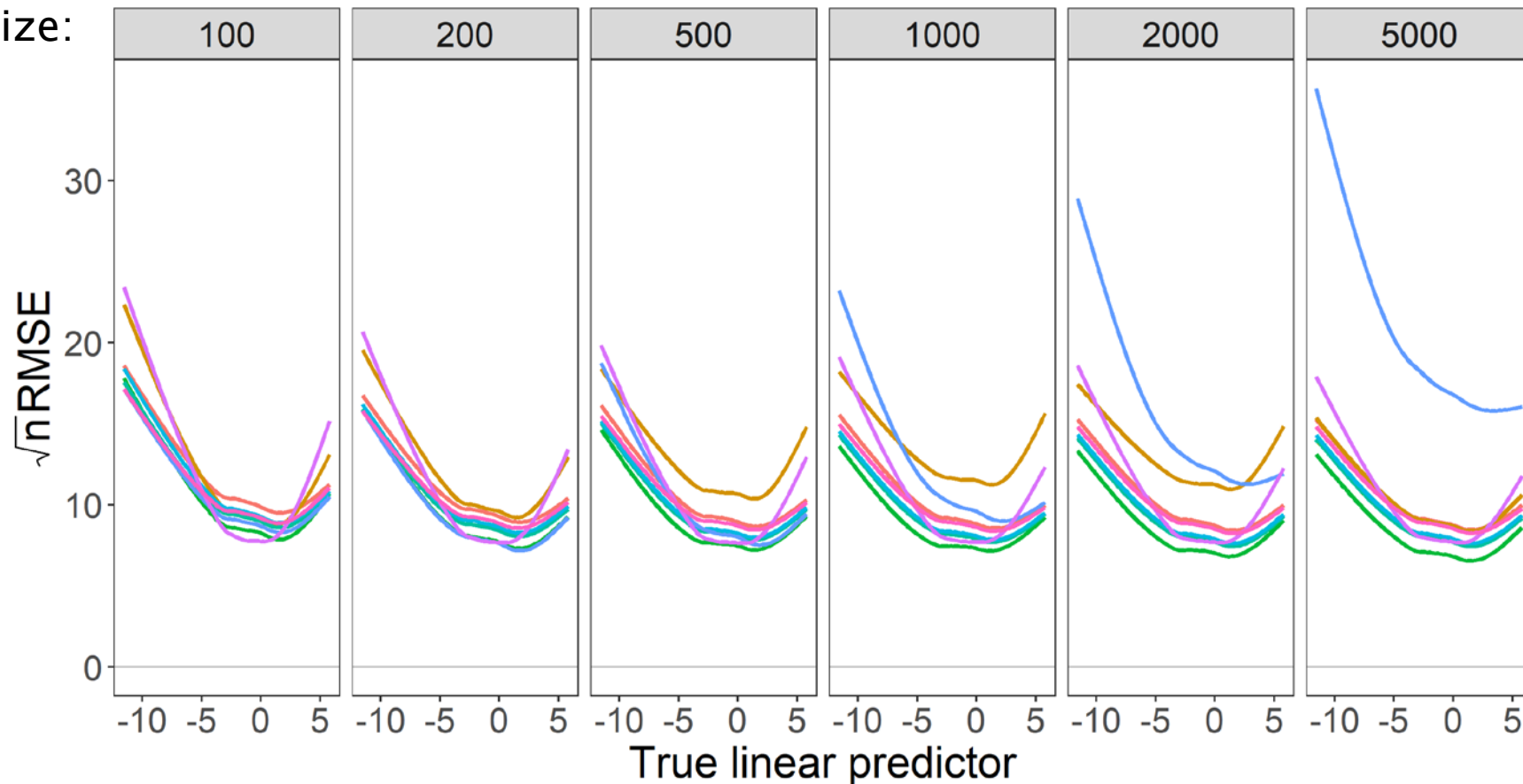
Predictors are represented by a solid line, and noise variables by a dashed line.



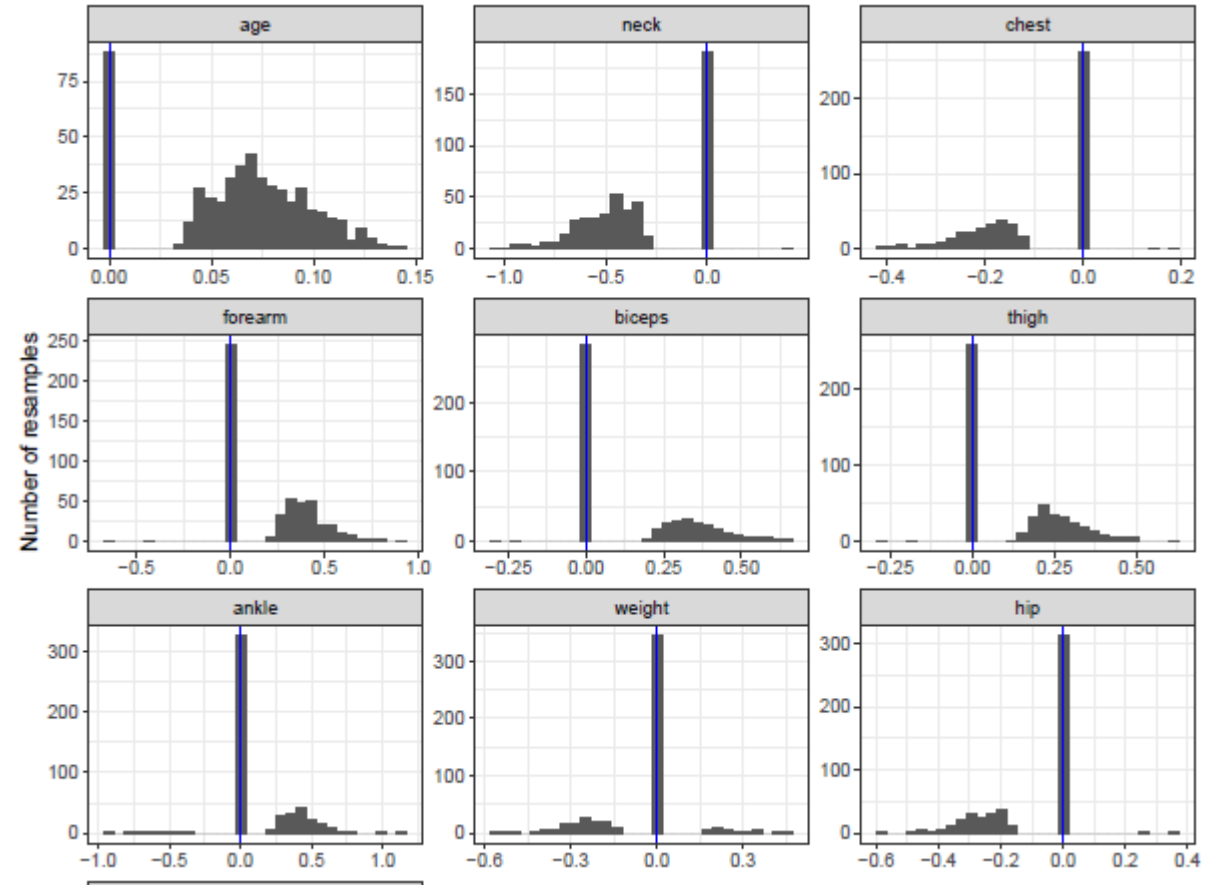
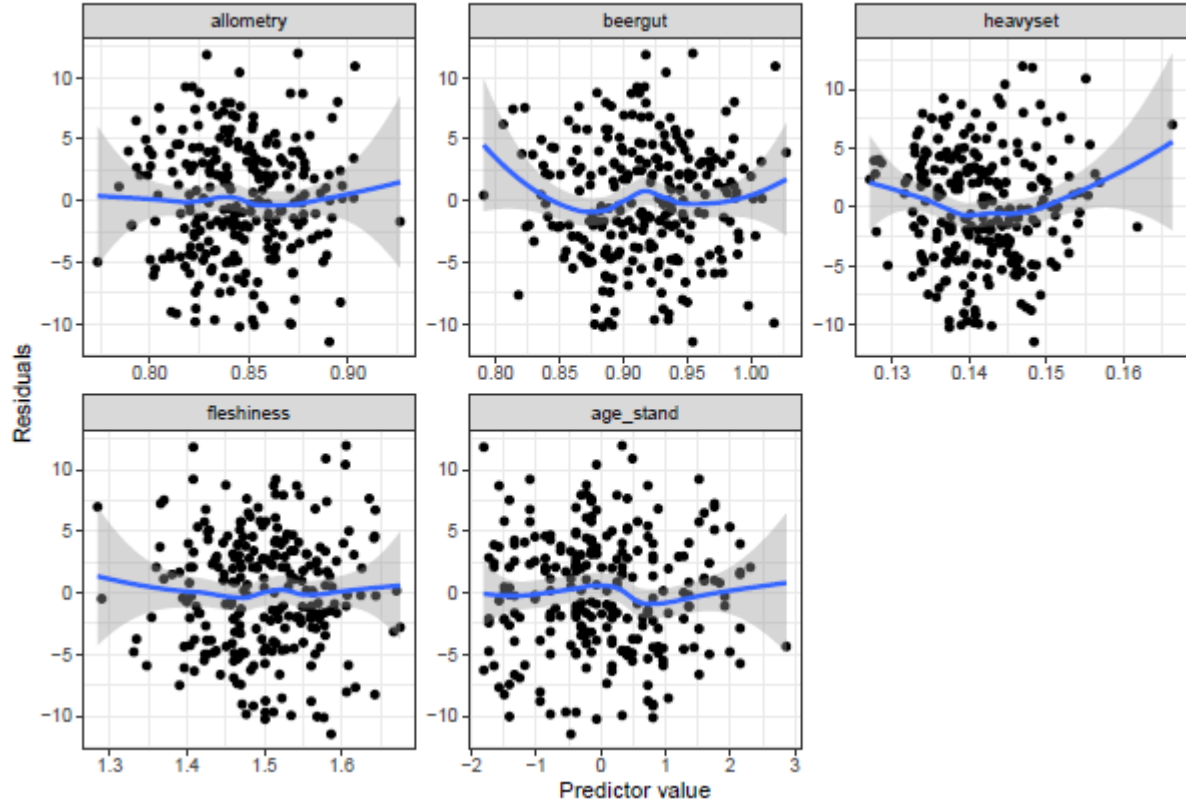
Local RMSE of predictions: sample size as main driver of comparative performance

— FU — FS_AIC — BE_AIC — FU_approx — RLasso
— Uni_020 — BE_005 — ABE_020 — Lasso

Sample size:

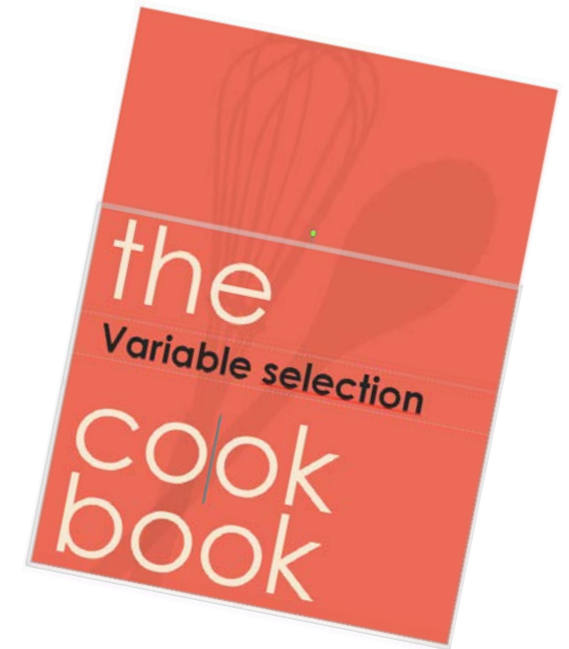


Focus on residual analysis and stability investigations



Recipe for disaster

- Prepare a long list of poorly conceived predictors.
- Add only small n .
- Mix together in an extensive iterative data dredging.
- Select the model with the smallest p -values.
- Present this final model without further considerations.



Bon appétit!

Common spline bases for regression models in practice: a comparative simulation study

Aris Perperoglou, Mathias Schmid, Michal Abrahamowicz,
Daniela Dunkler, Christine Wallisch, Georg Heinze

Nov. 2021

What has happened so far?

Review > [BMC Med Res Methodol.](#) 2019 Mar 6;19(1):46. doi: 10.1186/s12874-019-0666-3.

A review of spline function procedures in R

Aris Perperoglou¹, Willi Sauerbrei², Michal Abrahamowicz³, Matthias Schmid⁴

Affiliations + expand

PMID: 30841848 PMCID: [PMC6402144](#) DOI: [10.1186/s12874-019-0666-3](#)

Abstract

Background: With progress on both the theoretical and the computational fronts the use of spline modelling has become an established tool in statistical regression analysis. An important issue in spline modelling is the availability of user friendly, well documented software packages. Following the idea of the STRENGTHENING Analytical Thinking for Observational Studies initiative to provide users with guidance documents on the application of statistical methods in observational research, the aim of this article is to provide an overview of the most widely used spline-based techniques and their implementation in R.

Methods: In this work, we focus on the R Language for Statistical Computing which has become a hugely popular statistics software. We identified a set of packages that include functions for spline modelling within a regression framework. Using simulated and real data we provide an introduction to spline modelling and an overview of the most popular spline functions.

Results: We present a series of simple scenarios of univariate data, where different basis functions are used to identify the correct functional form of an independent variable. Even in simple data, using routines from different packages would lead to different results.

Conclusions: This work illustrates challenges that an analyst faces when working with data. Most differences can be attributed to the choice of hyper-parameters rather than the basis used. In fact an experienced user will know how to obtain a reasonable outcome, regardless of the type of spline used. However, many analysts do not have sufficient knowledge to use these powerful tools adequately and will need more guidance.

Keywords: Functional form of continuous covariates; Multivariable modelling.

Follow-up project

- Goal: investigate the usability of splines procedures for applied research.
- Audience: average user with limited splines experience.
- Focus on regression modelling with multiple variables of mixed types.
Splines used for modelling functional forms of continuous variables.
- Methods:
 - Natural splines
 - b-splines
 - p-splines
 - thin plate regression splines
 - Default settings and (if sufficiently different) suggestions from experts.
- Neutral comparison study for the univariate and the multivariable case.
- Plan: 2022 - finalize study protocol, simulation study
→ paper maybe with accompanying R tutorial

Long-term plan

- *Research questions:*
 - Can splines correctly identify functional forms?
 - ...
 - What do we gain by adding degrees of freedom in terms of predictive performance?
 - ...
 - **How do splines perform in a variable selection setting?**
 - What variable selection methods are the most promising?
- *Plan:*
 - Apply for third-party funding

Literature review of model building in Covid-19 prediction models

The screenshot shows the top navigation bar of the BMJ website with the logo 'thebmj' and menu items for 'covid-19', 'Research', 'Education', 'News & Views', 'Campaigns', and 'Jobs'. Below the navigation is a 'PEOPLE' section featuring four headshots of the authors: Maarten van Smeden, Carl Moons, Laure Wynants, and Ben Van Calster. The main article title is 'Prediction models for diagnosis and prognosis of covid-19: systematic review and critical appraisal', published in BMJ 2020; 369. The DOI is <https://doi.org/10.1136/bmj.m1328> and it was published on 07 April 2020. The citation information is 'Cite this as: BMJ 2020;369:m1328'.

- ‘37 421 titles were screened, and 169 studies describing 232 prediction models were included.’
- ,All models were rated at high or unclear 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 unclear reporting.

Literature review of model building in Covid-19 prediction models

- Currently Update 4 is finalized and will soon be published
- Data base of all extracted data available

- Issues of model building: What is the current practice regarding
 - Use of statistical vs. algorithmic models
 - Variable pre-selection
 - Variable selection during model building
 - Modeling of continuous variables
 - Shrinkage/regularization

Literature review of model building in Covid-19 prediction models

- Current status:
 - Preparation of protocol, comparison with extraction form of Covid-19 review
 - Initial descriptive analyses with the available data (via a shiny app!)
 - Inclusion/exclusion criteria defined
- Next steps:
 - Finish protocol:
 - Data cleaning of available data (for the purpose of this review)
 - Identify additional data to be extracted from the studies
 - Identify human resources (who wants to take over parts of the workload?)
 - Write paper

Collaboration with other TGs

- With TG3: Initial data analysis before regression analysis
 - Ongoing collaboration: ISCB2020, IBC2022
- With TG6: Model building for prediction models
 - Literature review of model building in Covid-19 prediction models (started)
 - Guidance for model building for prediction?
- With TG7: Model building for counterfactual prediction models
 - Guidance for model building for counterfactual prediction?