

Statistical Strategies to Address Main Research Questions of the MiSpEx Network and Meta-Analytical Approaches

Statistische Strategien zur Beantwortung der Hauptfragestellungen des MiSpEx-Netzwerks und metaanalytische Herangehensweisen

ACCEPTED: June 2018

PUBLISHED ONLINE: July 2018

DOI: 10.5960/dzsm.2018.339

Hönning A, Stengel D, Güthoff C.

Statistical strategies to address main research questions of the MiSpEx network and meta-analytical approaches. Dtsch Z Sportmed. 2018; 69: 236-239.

1. BG KLINIKUM UNFALLKRANKENHAUS
BERLIN GMBH, Zentrum für Klinische
Forschung, Berlin

Summary

- › **In the MiSpEx project**, the statistical analysis of main research questions accompanied by the aggregation of diverse data obtained in the individual studies represents a particular challenge. Research goals relating to the effectiveness and efficiency of training interventions in order to reduce the risk and degree of back pain may be addressed by meta-analytical approaches. This requires a sufficient similarity of the underlying studies and, in case of applying a meta-analysis of individual data, uniformly structured data sets.
- › **Approaches of supervised statistical learning** should be applied to research goals concerning the diagnosis of back pain, and the prognosis either of further course of symptoms and of the reaction to preventive and curative interventions. The various underlying data sets should enable a step-by-step approach with several development and validation steps, likewise provided that sufficient comparability exists.
- › **The assessment** of moderating effects of the influencing factors pain experience, training status, psychophysical distress, and level of care on treatment success is of high importance.

KEY WORDS:

Aggregation, Analysis Strategy,
Prevention, Back Pain

Zusammenfassung

- › **Im MiSpEx-Projekt** stellen die statistische Auswertung der übergreifenden Hauptfragestellungen und die mit ihr verbundene Frage nach der Möglichkeit der Aggregation der in den einzelnen Studien gewonnenen Daten eine besondere Herausforderung dar. Für die Fragestellungen hinsichtlich Effektivität und Effizienz der Trainingsinterventionen zur Prävention oder Reduktion von Rückenschmerzen sind vor allem metaanalytische Herangehensweisen in Betracht zu ziehen. Diese setzen eine hinreichende Ähnlichkeit der zugrundeliegenden Studien voraus und erfordern bei einer Metaanalyse von Individualdaten einheitlich strukturierte Datensätze.
- › **Zur Bearbeitung der Fragestellungen** zur Diagnostik von Rückenschmerzen, der Prognose des weiteren Symptomverlaufs sowie der Prädiktion der Reaktion auf präventive und kurative Interventionen sollen Ansätze des supervidierten statistischen Lernens herangezogen werden. Hierbei sollen –wieder unter Voraussetzung hinreichender Vergleichbarkeit – die verschiedenen Datensätze ein stufenweises Vorgehen mit mehreren Entwicklungs- und Validierungsschritten ermöglichen.
- › **Von besonderer Bedeutung** ist die Prüfung des moderierenden Effekts der Einflussfaktoren Schmerzerleben, Trainingszustand, Psychophysischer Stress und Versorgungskontext auf den Behandlungserfolg.

SCHLÜSSELWÖRTER:

Aggregation, Analysestrategie,
Prävention, Rückenschmerzen

Introduction

The Medicine in Spine Exercise (MiSpEx) network includes a variety of working groups and individual studies with diverse designs including randomized controlled trials, quasi-experimental investigations, and cross-sectional or longitudinal series conducted in an open-label or blinded manner (refer to the previous article for a comprehensive description of the consortium). Several studies address (to

a greater or lesser extent) one or more of the main research questions of the MiSpEx project. In this article, we describe strategies how to integrate and synthesize findings and data from these individual studies with mixed methods and datasets. To address the main research questions of the project listed below, we sorted them corresponding to the chosen analytical approaches in an optimal way.



Article incorporates the Creative Commons
Attribution – Non Commercial License.

<https://creativecommons.org/licenses/by-nc-sa/4.0/>



QR-Code scannen
und Artikel online
lesen.

CORRESPONDING ADDRESS:

Alexander Hönning, M.Sc.
Zentrum für Klinische Forschung
BG Klinikum Unfallkrankenhaus
Berlin gGmbH
Warener Str. 7, 12683 Berlin
✉: alexander.hoenning@ukb.de

At first, we consider basic data management requirements in the context of meta-analysis. Then three approaches for data analysis corresponding to the following research goals are described: i. identification of minimal physical activity causing adaptational response, ii. construction and calculation of the diagnostic and prognostic indices, and iii. evaluation of the efficacy of sensory-motoric training intervention on reducing the risk and degree of back pain.

The main research questions of the project are defined as follows:

- Q1: Which variables facilitate function diagnostics in chronic unspecific back pain by means of an upper body stability and function index ("Rumpf-Stabilitäts- und Funktions-Index", RSFI)? This RSFI must allow for assigning patients, healthy individuals and / or athletes to preventive or therapeutic interventions with a focus on physical activity.
- Q2: Which physical activity interventions applying to the target group may reduce the risk of back pain in healthy individuals, or reduce back pain in symptomatic subjects? Is it possible to measure the effectiveness of these interventions using an upper body prevention index ("Rumpf-Präventions-Index", RPI) based on training status, pain experience, and psychophysical distress in a reliable and valid way?
- Q3: Is it feasible to apply physical activity interventions to reduce the risk or degree of back pain in different contexts (high-performance versus leisure sports, sports medical versus general health care)? Is it possible to identify common characteristics of such programs?

Besides the main research questions, the following two aspects shall be investigated:

- The minimum physical activity leading to adaptation, and
- The extent and individual response of adaptation to physical activity given moderating factors like pain experience, training status, psychophysical distress, and level of care.

Type of Meta-Analysis and Assumptions

According to Blettner et al. (1) there are four distinct types of meta-analyses. Type I refers to analyses of reviews and is of no relevance here. Type II, which refers to analyses of published single studies also is not feasible in this setting as published information is insufficient to perform a type II meta-analysis yet. Type III and type IV both refer to analyses of individual data of single studies with either a retrospective study selection or a prospective study selection. For these types of meta-analyses, data sets of single studies must be merged in one common data set. This needs standardized labelling of all variables in the common data set to provide unambiguous identifiability of any variable with its respective label. A useful tool is a unified code book describing the variables of the single studies and their value labels.

Identification of Minimal Physical Activity causing Adaptational Response and Dependence of Response to Moderating Factors

Physical activity is defined as the product of the average number and duration of training exercises per week. Changes in parameters of pain and function will be considered relevant in terms of their clinical relevance. Both absolute and relative changes may serve as threshold values. The period from

start of center-based training up to three weeks after start is defined as the empirical basis for calculating the extent of physical activity and changes in parameters.

Pain is operationalized by the Von Korff questionnaire (measuring pain intensity and disability). Parameters of function encompass diverse measures like muscular capacity, postural control, mobility, movement and muscle activity.

There are several approaches in the literature how to determine clinically relevant changes (CRC) which usually fall into one of two categories: consensus based versus data based. (9, 11) In case of consensus, an expert panel (of the MiSpEx network) defines CRC for every parameter based on clinical experience and available evidence.

The data-based approach has two further alternatives: 1. the anchor-based approach relates the parameter of interest to a pre-specified change in another parameter (2), 2. the distribution-based approach measures the minimal detectable effect not attributable to measurement error (7). The latter is feasible in most scenarios, but often lacks a relation to other clinically relevant indices.

All studies within the project investigating sensory-motoric training interventions and showing sufficient similarity are to be pooled. Similarity between studies will be investigated using the PICOT (i.e., population, intervention, comparator, outcome, time) scheme without considering the comparator category. The average physical activity per week must be calculated for each intervention within a study. A binary variable for every parameter of pain and function indicating the achievement or non-achievement of a CRC must be created. These will play the role of status variables in calculating Receiver Operating Characteristic (ROC) curves while the average physical activity per week will function as test variable. ROC curves will be calculated using either a nonparametric or parametric approach depending on distributional properties of the data.

The test variable will be evaluated for every status variable with the area under the curve (AUC) of the ROC curve together with Youden's index ((sensitivity + specificity) -1). While AUC values range between 0 and 1, Youden's index ranges from -1 to +1. Values near 1 of both evaluation parameters indicate a high discriminatory power of the cut-off-value. Parameters for which a high discriminatory power is achieved can contribute to defining a common minimum of physical activity leading to adaptation. Additionally, a stratified analysis will investigate thresholds for top athletes, amateur athletes, and non-athletes.

To examine extent and individual response of the adaptation to physical activity in dependence of the moderating factors, pain experience, training status, psychophysical distress, and health care context will be entered in logistic regression models reflecting the given diagnostic setting as effect modifiers, and interaction terms will be built. A significant interaction term may indicate the presence of a moderating effect.

Index of Upper Body Stability and Function (RSFI) and Upper Body Prevention Index (RPI)

Research questions Q1 and Q2 (second part) aim at searching for the optimal set of diagnostic, prognostic, and predictive markers.

Basically, there are three distinct types of diagnostic studies according to the Fryback-Thornbury hierarchy (6): i. studies of diagnostic test accuracy to investigate the reliability of a certain test to determine the absence or presence of a disease or health-relevant condition, ii. studies of diagnostic

efficacy to investigate the extent a certain test outcome influences subsequent treatment decisions, and iii. studies of diagnostic efficiency to investigate whether the patient outcome is influenced by the deployment of a certain diagnostic test.

Prognostic markers usually indicate how a certain disease will progress independently of treatment, i.e., they correlate with the natural course of the disease.

In contrast, predictive markers allow for predicting the reaction of an individual or organ to a certain treatment.

The outcome of untreated and treated patients with a certain prognostic marker may differ compared to untreated and treated patients without the prognostic marker.

The relevant difference here is the outcome of patients carrying the marker of interest AND undergoing the intervention of interest compared to all three other thinkable groups of subjects.

For selecting a prognostic variable set given the number of theoretically relevant variables and number of test subjects, model selection by a shrinkage method like the Least Absolute Shrinkage and Selection Operator (LASSO) suits best as it avoids overfitting and allows for implementing model estimation and variable selection in a single step (5). The LASSO enables estimation and selection of regression coefficients in the context of high numbers of possible regressors and small sample size. The advantage to other techniques such as ridge regression is that a sparse solution can be obtained by using an absolute instead of a squared value penalty so that small effects can be set to exactly zero (8). This leads ultimately to a valid model with a reasonable number of relevant predictors. In case of more than one dataset it allows for gradual sharpening by stepwise validation.

The RPI, based on training status, pain experience and psychophysical distress, had already been developed and published within the MiSpEx project as RPI-S (10) using a similar approach. To link treatment decisions to future pain outcomes, the latter were regressed on variable sets representing four domains of risk factors: pain, distress, social environment (including training status), and medical care environment, using the lasso for variable selection. The resulting four prognostic variable sets would not only predict the extent of the long-term risk of pain, but also depict the underlying specific deficits of the individual subject. Exceeding the threshold of the prognostic variable set would link the increased risk of future pain to the values in the specific domain and therefore determine the best course of therapeutic action for the participant.

domain as well as across all domains. The impact of bias is investigated with a sensitivity analysis which includes only studies with low risk of bias and compares the resulting effect estimate with that resulting from the main analysis.

In the main analysis, the effect of the intervention will be estimated by the mean difference of gain scores between intervention and inactive control for all outcome criteria and displayed via forest plots separately for single studies. Gain scores will be calculated as differences between baseline values at the start of the study and values obtained 12 weeks later. Heterogeneity will be taken into account by using a random effects model (4), a modification to the inverse-variance approach.

Given a sufficient number of studies to be included in the analyses, several sub group analyses will be performed (e.g., by intervention group, sports activity, chronic back pain, compliance, training frequency, etc.). Otherwise, descriptive comparison of individual subgroups without formal data aggregation within a random effects model will be attempted. ■

Förderung

Das MiSpEx-Netzwerk wird gefördert aus Mitteln des Bundesinstituts für Sportwissenschaft (BiSp) aufgrund eines Beschlusses des Deutschen Bundestages [Förderkennzeichen ZM-VII-080102A/11-18].

References

- (1) **BLETTNER M, SAUERBREI W, SCHLEHOFER B, SCHEUCHENPFLUG T, FRIEDENREICH C.** Traditional reviews, meta-analyses and pooled analyses in epidemiology. *Int J Epidemiol* 1999; 28: 1-9.
- (2) **COPAY AG, GLASSMAN SD, SUBACH BR, BERVEN S, SCHULER TC, CARREON LY.** Minimum clinically important difference in lumbar spine surgery patients: a choice of methods using the Oswestry Disability Index, Medical Outcomes Study questionnaire Short Form 36, and pain scales. *Spine J.* 2008; 8: 968-974. doi:10.1016/j.spinee.2007.11.006
- (3) **DEEKS JJ, HIGGINS JPT, ALTMAN DG.** Analysing Data and Undertaking Meta-Analyses. In: *Cochrane Handbook for Systematic Reviews of Interventions.* John Wiley & Sons, Ltd. 2008: 243-296.
- (4) **DERSIMONIAN R, LAIRD N.** Meta-analysis in clinical trials. *Control Clin Trials.* 1986; 7: 177-188. doi:10.1016/0197-2456(86)90046-2
- (5) **FAHRMEIR L, KNEIB T, LANG S, MARX B.** Regression: models, methods and applications: Springer Science & Business Media, 2013.
- (6) **FRYBACK DG, THORNBURY JR.** The Efficacy of Diagnostic Imaging. *Med Decis Making.* 1991; 11: 88-94. doi:10.1177/0272989X9101100203
- (7) **MCGLOTHLIN AE, LEWIS RJ.** Minimal clinically important difference: defining what really matters to patients. *JAMA.* 2014; 312: 1342-1343. doi:10.1001/jama.2014.13128
- (8) **TIBSHIRANI R.** Regression shrinkage and selection via the lasso. *J R Stat Soc B.* 1996; 58: 267-288.
- (9) **WIJEYUNDERA DN, JOHNSON SR.** How Much Better Is Good Enough?: Patient-reported Outcomes, Minimal Clinically Important Differences, and Patient Acceptable Symptom States in Perioperative Research. *Anesthesiology.* 2016; 125: 7-10. doi:10.1097/ALN.0000000000001159
- (10) **WIPPERT P-M, PUSCHMANN A-K, DRIESLEIN D, ARAMPATZIS A, BANZER W, BECK H, SCHILTENWOLF M, SCHMIDT H, SCHNEIDER C, MAYER F.** Development of a risk stratification and prevention index for stratified care in chronic low back pain. Focus: yellow flags (MiSpEx network). *PAIN Reports.* 2017; 2: e623. doi:10.1097/PR9.0000000000000623
- (11) **XU T.** Statistical Development on Determining the Minimum Clinically Important Difference. *Biometrics & Biostatistics International Journal.* 2015; 2. doi:10.15406/bbij.2015.02.00041



Effect of Sensory-Motoric Training Intervention Reducing the Risk and Degree of Back Pain

Article incorporates the Creative Commons Attribution – Non Commercial License
<https://creativecommons.org/licenses/by-nc-sa/4.0/>

Research questions Q2 (first part) and Q3 will be investigated by means of prospective meta-analyses including all studies featuring sensory-motoric interventions. Heterogeneity of studies will visually be assessed by overlapping 95% confidence intervals in forest plots and I^2 -statistic of the main analysis. The following I^2 reference values (3) will be used:

- 0% to 40%: might not be important;
- 30% to 60%: may represent moderate heterogeneity;
- 50% to 90%: may represent substantial heterogeneity;
- 75% to 100%: considerable heterogeneity.

Risk of bias of single studies needs to be categorized as „low risk“, „high risk“ or „unclear risk“ of bias for different domains (i.e., random sequence generation, allocation concealment, blinding of personnel, blinding of outcome assessment, incomplete outcome data, other bias), and discussed for each