Statistical analysis of the regression discontinuity design (RDD): Implications of combined or repeated measures

Inaugural Dissertation

zur

Erlangung des Doktorgrades
philosophiae doctor (PhD) in Health Sciences
der Medizinischen Fakultät
der Universität zu Köln

vorgelegt von

Anna Hagemeier

aus Steinheim (Westf.)

Köln

2026

Betreuer:	Prof. Dr. Martin Hellmich
Gutachterin:	Prof. Dr. Elke Kalbe
	Prof. Dr. Nicole Skoetz

Datum der Mündlichen Prüfung: 01.06.2026

The project and new form of care integrated, cross-sectoral psycho-oncology provided the framework for the three publications of this dissertation. The project was funded by the Innovation Fund by the German Federal Joint Comittee (Gemeinsamer Bundesausschuss (GBA); grant number: 01NVF17022).

Abstract

Introduction: With its threshold-based approach, the quasi-experimental regression discontinuity design (RDD) represents an alternative to randomised controlled trials (RCTs). It enables the formation of comparable groups and their examination using regression analysis when random group assignment is not feasible, for example for ethical reasons. However, when evaluating longitudinal data with repeated measures, as it is often the case in health-related fields, the RDD reaches its methodological limits. In order to analyse longitudinal data and time effects in the outcome, linear mixed models (LMMs) for repeated measures are often used. Therefore, combining both approaches merges two methodological challenges and offers a comprehensive and integrated solution.

Aims and Objectives: The central aim of the cumulative dissertation is the methodological development of the RDD approach for the low-bias identification of causal effects in longitudinal data with repeated measures. At the same time, the approach is to be opened up for the health sciences and its practical applicability strengthened. Therefore, the work pursues two interlinked objectives: i) the introduction and application of the classical RDD approach in health sciences (articles 1 and 2) and ii) its extension to analyse repeated measures (article 3).

Methods: The methodological extension of the RDD for repeated measures is based on the standard RDD and its combination with LMMs for repeated measures. Therefore, the first article presented the mathematical fundamentals of the standard RDD approach and guidelines for its implementation using statistical software, as well as a worked example in health research. Furthermore, another section of the methodology part of the dissertation presented the fundamentals and the mathematical notation of LMMs for repeated measures in the context of health sciences.

Applications and Results: Based on the conceptual principles of the univariate standard RDD approach, its application in the clinical setting was demonstrated in the second article using a subgroup analysis within the framework of the new form of care integrated cross-sectoral psycho-oncology (nFC-isPO) in order to highlight the relevance of the method for quasi-experimental designs in health services research. Hence, this analysis highlights the limitations of the standard RDD for longitudinal data and identifies a significant research gap in the application, as repeated measurements cannot be taken into account yet. Based on this finding, a methodological extension was developed in the context of the third article, which combines the RDD approach with an LMM for repeated measures, thereby enabling the analysis of longitudinal data. The consistent mathematical notation of the combination of both approaches was derived in a generalised form. To illustrate its practical relevance and

test the feasibility, the new LMM-RDD approach was applied to the data from the nFC-isPO.

Conclusion: The results indicate that the proposed LMM-RDD increases the precision and interpretability of estimated effects in longitudinal data settings and enables a more nuanced investigation of time-dependent treatment effects compared to the standard RDD. The detailed description of the mathematical notation and implementation steps ensures an easy replication and facilitates its transfer into applied medical research. This work therefore makes a significant contribution to the further development of medical statistics and provides a practical set of tools that will help to achieve reliable, evidence-based results in clinical trials in the long term. In summary, this work contributes to the methodological advancement of the quasi-experimental RDD for repeated measures in health research and to the further development of alternative study designs, including their analysis methods, and opens up new perspectives for their application in clinical research.

Summary in German

Einleitung: Das quasi-experimentelle Regressionsdiskontinuitätsdesign (RDD) bietet mit seinem schwellenwertbasierten Ansatz eine Alternative zu randomisierten kontrollierten Studien (RCTs). Es ermöglicht die Bildung vergleichbarer Gruppen und deren Untersuchung mittels Regressionsanalyse, wenn eine zufällige Gruppenzuweisung - beispielsweise aus ethischen Gründen - nicht möglich ist. Bei der Auswertung von Längsschnittdaten mit wiederholten Messungen, wie sie häufig im Gesundheitsbereich vorkommen, stößt das RDD jedoch an seine methodischen Grenzen. Um Längsschnittdaten und Zeiteffekte auf das Outcome zu analysieren, werden vielfach lineare gemischte Modelle (LMMs) für wiederholte Messungen verwendet. Die Kombination beider Ansätze vereint daher zwei methodische Herausforderungen und stellt somit eine umfassende, integrierte Lösung bereit.

Ziele und Zwecke: Das zentrale Ziel der kumulativen Dissertation ist die methodische Weiterentwicklung des RDD-Ansatzes zur verzerrungsarmen Identifizierung kausaler Effekte in Längsschnittdaten mit wiederholten Messungen. Parallel dazu soll der Ansatz für die Gesundheitswissenschaften zugänglich gemacht und seine praktische Anwendbarkeit gestärkt werden. Aus diesem Grund verfolgt die Arbeit zwei miteinander verknüpfte Zwecke: i) die Einführung und Anwendung des klassischen RDD-Ansatzes in den Gesundheitswissenschaften (Artikel 1 und 2) und ii) dessen Erweiterung zur Analyse wiederholter Messungen (Artikel 3).

Methoden: Die methodische Erweiterung des RDD für wiederholte Messungen basiert auf dem Standard-RDD und dessen Kombination mit LMMs für wiederholte Messungen. Daher befasst sich der erste Artikel zunächst mit den mathematischen Grundlagen des Standard-RDD-Ansatzes, stellt Leitlinien für dessen Umsetzung mithilfe von Statistiksoftware bereit und zeigt anhand eines Arbeitsbeispiels aus der Gesundheitsforschung die Anwendung. In einem weiteren Abschnitt des Methodikteils der Dissertation wurden darüber hinaus die Grundlagen und die mathematische Notation von LMMs für wiederholte Messungen im Kontext der Gesundheitswissenschaften vorgestellt.

Anwendungen und Ergebnisse: Basierend auf den konzeptionellen Prinzipien des univariaten Standard-RDD-Ansatzes wurde dessen Anwendung im klinischen Setting im zweiten Artikel anhand einer Subgruppenanalyse im Rahmen der neuen Versorgungsform der integrierten sektorübergreifenden Psychoonkologie (nFC-isPO) demonstriert, um die Relevanz der Methode für quasi-experimentelle Designs in der Versorgungsforschung hervorzuheben. Diese Analyse verdeutlicht somit die Grenzen des Standard-RDD für Längsschnittdaten und identifiziert eine erhebliche Forschungslücke hinsichtlich der Anwendung, da wiederholte Messungen bislang nicht berücksichtigt werden konnten. Auf Basis dieser Erkenntnis wurde im

Rahmen des dritten Artikels eine methodische Erweiterung entwickelt, die den RDD-Ansatz mit einem LMM für wiederholte Messungen kombiniert und damit die Analyse von Längsschnittdaten ermöglicht. Die konsistente mathematische Notation der Kombination beider Ansätze wurde in einer verallgemeinerten Form abgeleitet. Um die praktische Relevanz zu veranschaulichen und die Durchführbarkeit zu prüfen, wurde der neue LMM-RDD-Ansatz auf die Daten aus dem nFC-isPO-Projekt angewendet.

Schlussfolgerung: Die Ergebnisse zeigen, dass das vorgeschlagene LMM-RDD die Genauigkeit und Interpretierbarkeit der geschätzten Effekte in Längsschnittdaten erhöht und im Vergleich zum Standard-RDD eine differenziertere Untersuchung zeitabhängiger Behandlungseffekte ermöglicht. Die detaillierte Beschreibung der mathematischen Notation und der Implementierungsschritte gewährleistet eine einfache Reproduzierbarkeit und erleichtert die Übertragung in die angewandte medizinische Forschung. Diese Arbeit leistet somit einen wichtigen Beitrag zur Weiterentwicklung der medizinischen Statistik und stellt ein praktisches Instrumentarium bereit, das langfristig dazu beitragen wird, zuverlässige, evidenzbasierte Ergebnisse in klinischen Studien zu erzielen. Zusammenfassend trägt diese Arbeit zum methodischen Fortschritt des quasi-experimentellen RDD für wiederholte Messungen in der Gesundheitsforschung sowie zur Weiterentwicklung alternativer Studiendesigns einschließlich ihrer Analysemethoden bei und eröffnet neue Perspektiven für deren Anwendung in der klinischen Forschung.

Contents

Abstract i

Summary in German iii

List of Figures vii

List of Tables vii

List of Abbreviations ix

1 Introduction 1

1.1 Development of the regression discontinuity design 3

1.2 Clinical background of the dissertation 5

1.3 The nFC-isPO 7

1.4 Repeated Measures 9

2 Aims and Objectives 11

3 Methods 13

3.1 Metholodical Overview of the RDD - 1. Article 14

3.2 Linear mixed models for repeated measures 22

4 Applications and Results 29

4.1 Application of RDD in the nFC-isPO - 2. Article 30

4.2 Extension of RDD to LMM-RDD - 3. Article 39

5 Discussion 51

5.1 Key aspects of the publications 51

5.2 Relevance to the current state of research 53

5.3 Strengths, limitations and open issues 55

5.4 Conclusion 58

References 69

Declaration of contribution 71

Acknowledgement 75

List of further Publications 80

Appendix **81**

Appendix of 1. Article 81

Confirmation of ethics approval 86

Attachment **87**

Curriculum vitae 87

Affidavit (Eidesstattliche Erklärung) 89

List of Figures

1 Experiments vs. quasi-experiments 1
2 Mortality by types of diseases 2023 5
3 The isPO care programme with schematic RDD diagram 8

4 Visual abstract of first article 14
5 Schema of a hierarchical data structure 23
6 Fixed and random effects in LMMs for repeated measures. 24

7 Visual abstract of second article 30
8 Visual abstract of third article 39

List of Tables

3.1 Overview of the different regression models 27

5.1 Aims and key aspects of the publications 52
5.2 Strenghts and Limitations of the RDD and LMM-RDD 55

List of Abbreviations

ACE	Average causal effect
ATE	Average treatment effect
AWMF	Association of the Scientific Medical Societies in Germany (Arbeitsgemeinschaft der Wissenschaftlichen Medizinischen Fachgesellschaften)
CACE	Complier average causal effect
COVID-19	Coronavirus disease 2019
dapo	German Working Group for Psychosocial Oncology
DKG	German Cancer Society (Deutsche Krebsgesellschaft)
GBA	German Federal Joint Comittee, (Gemeinsamer Bundesausschuss)
GFO	Gesellschaft der Franziskanerinnen zu Olpe
HADS	Hospital Anxiety and Depression Scale
ITT	Intention-to-treat
LATE	Local average treatment effect
LMM	Linear mixed models
nFC-isPO	New form of care integrated crosssectoral psycho-oncology
NRW	North Rhine-Westphalia
PS	Propensity score
PSM	Propensity score matching
PSO	Working Group Psycho-oncology
PSR	Psychosocial Risk Assessment Questionnaire
RCT	Randomised controlled trial
RDD	Regression discontinuity design
RDiT	Regression discontinuity in time
RM-ANOVA	Analysis of variance for repeated measurements
WPO	Further education and training in psychosocial oncology e.V. (Fort- und Weiterbildung in der Psychosozialen Onkologie)

1 Introduction

In scientific research, empirical studies are conducted to obtain information, i.e. investigations are carried out to confirm or reject hypotheses and gain new insights into certain processes, e.g. new treatment methods in clinical studies. These so-called experiments are therefore used to investigate or test cause-and-effect relationships, i.e. causality, causal effects or causal inferences [1–3]. For this purpose, Rubin formulated a model in 1974 [4] that mathematically defines causality as the average causal effect (ACE), also known as the average treatment effect (ATE). The causal effect results from the difference between two possible outcomes of an observed individual - the hypothetical ('potential') outcome with treatment and the hypothetical outcome without treatment. Since these two conditions cannot be realised simultaneously for the same individual, the effect is determined across groups of individuals. In other words, experiments on causality analyse the influence of at least one independent variable, e.g. a new drug compared to a placebo, on one or more dependent variables, e.g. laboratory parameters such as blood values. In order to create structural equality, i.e. to limit bias, individual test subjects, e.g. patients, are therefore assigned to the various independent variable characteristics or groups at random. With this approach, the experiments achieve a high degree of internal validity, i.e. the observed effects can actually be attributed to the independent variable and are less likely to be biased by confounding factors (see figure 1).

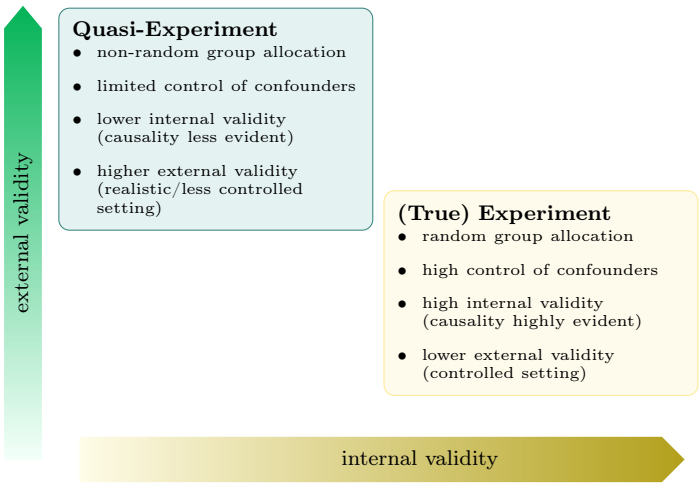


Figure 1: (True) experiments vs. quasi-experiments and internal and external validity [1–3].

Besides internal validity, there is also external validity, which indicates the generalisability of the results beyond the specific controlled experimental setting. External validity depends on the extent to which an controlled experiment, i.e. the study environment including its sample and procedures, represents the broader population to which the conclusions intended

to be applied (see figure 1) [1–3].
The strongest experiments for clinical trials with the highest level of evidence (methodologically high quality) are randomised controlled trials (RCTs), which are considered the practical gold standard for causal conclusions [5, 6]. However, random allocation of patients is often not feasible or even unethical because, for example, there is already an effective treatment that cannot be withheld from a patient group. For this reason, alternative study designs with the highest possible level of evidence are required in addition to RCTs and systematic reviews of RCTs [7].
Quasi-experiments provide a promising approach in some of these settings. These classes of designs are suitable alternatives in real-world settings as there is a high potential to study patient groups that would not or cannot participate in a RCT [8, 9]. Like experiments, they test causal hypotheses, i.e. they investigate the effect of one or more independent variable on one or more dependent variable, but without the comprehensive control of a real experiment, in particular without randomised group assignment. However, in order to control for confounding factors and consequently ensure internal validity, appropriate statistical methods must be used to determine causal relationships with as little bias as possible. Even if confounding factors are adequately controlled, the internal validity of quasi-experimental studies can be high, but does usually not reach the internal validity of a true experiment. However, this does not apply to the external validity of quasi-experiments, which can be even higher, as they are often conducted under more realistic, less controlled conditions (see figure 1) [2, 3, 10, 11].

The foundation for the variety of experimental and statistical conclusions was laid by R. Fisher in 1935 [12]. Around 30 years later, Campbell and Stanly introduced and established the terminology of quasi-experimental study designs for the first time in 1963 [3]. They described the use of comparison groups as in experimental studies, but without randomised group allocation [13]. Due to the lack of randomisation, the formation of comparison groups respectively the control for confounding must be carried out differently [10].
A frequently used method of forming comparison groups is to use historical controls based on data available before the start of the study intervention. However, there may be systematic differences between the observations of the historical control group and the treatment group, e.g. due to differences in the standard of care, or due to incomplete or non-standardised data from the historical controls, which may lead to deviations in the primary endpoint and thus a lower comparability of the data [14]. Another way to create comparable groups is to use matching methods such as propensity score matching (PSM). PSM creates equality between treated and untreated groups by matching subjects with similar propensity scores (PS). A PS is the probability that a subject will receive the treatment based on observed characteristics (e.g. age, gender, health status, etc.). It is usually estimated using logistic regression and is therefore also called a predictive model. Nevertheless, this approach bears the problem that the matching may be biased because important influencing variables were not measured or recorded. Furthermore, it is possible that too few PS pairs overlap, which reduces the sample size [3, 15–17]. The quasi-experimental difference-in-difference approach offers another way to compare groups and estimate causal effects. Treatment effects are estimated by comparing

the treatment and control groups before and after the intervention. For example, groups can be formed in different areas e.g. different hospitals with the same standard intervention, whereby the new intervention is only implemented and carried out in one of them during the study period. Nonetheless, a central assumption of this approach and the way the groups are formed is that the trends of the treatment and the control group under standard intervention run parallel, which is a strong assumption when comparing groups, especially in spatially separated groups in different hospitals. Moreover, possible structural differences must be taken into account in the final evaluation, which is why possible additional sensitivity analysis must be carried out. Also external influences can lead to larger differences in the results, especially if only one group is influenced [18–20].
Another commonly used option is to observe the sample as a whole, rather than forming groups, and then apply regression modelling and include confounding factors as covariates to estimate the impact of these factors on the outcome. With this approach, however, only the measured variables can be controlled, and often not all confounding factors can be measured. Furthermore, it is commonly the case that supposedly unimportant factors are overlooked at the beginning of the study, so that unobserved confounding factors can further distort the results. And even if confounding factors are controlled for, causality in observational studies is not certain - it remains an approximation [21]. Moreover, incorrect modelling assumptions are problematic because if, for example, a linear relationship between variables is wrongly assumed when in reality a non-linear relationship exists, estimates could be biased [22].

So although there are several alternatives to randomised group allocation, each of them has minor or major disadvantages. However, there is another alternative in the field of quasi-experiments, the regression discontinuity design (RDD). This approach is much closer to an RCT than those mentioned above and therefore has significantly fewer of the disadvantages mentioned.

1.1 Development of the regression discontinuity design

The quasi-experimental RDD is an alternative approach to test causal inferences (for definitions, see section 1 Introduction) by using a cut-off on a continuous scale to assign groups, e.g. to divide patients into a treatment and a control group. These groups are then compared using regression models, or more precisely, a regression is estimated below and above the cut-off. If the regression lines show a jump (discontinuity) at the threshold, the difference can be interpreted as a causal effect of the treatment. The most important assumption here is that the two groups are very similar near the cut-off, and that the allocation is therefore quasi-random. This is usually achieved by limiting the comparison to a predefined bandwidth around the cut-off [23–25]. Just like the difference-in-difference approach, it is a method which is primarily known from the field of econometrics. However, history shows that several disciplines (psychology, statistics, econometrics) have worked independently on the development and application of RDD [26, 27]. The first research results for the later RDD were provided by Cambell and Thistlethwaite in the year 1960 in the field of education

and psychology [28]. Both are best known for their research on quasi-experimental designs, particularly their work on evaluating research questions in which thresholds could be used for group allocation and thus measure causal effects when randomisation was not possible. In 1963, Campbell, together with Stanley, continued to work on quasi-experimental designs, especially the RDD, and provided further conceptual explanations focusing on issues such as selection bias and, in particular, local randomisation around a cut-off. However, they could not provide statistical evidence at this time. They only briefly discussed statistical analysis options for measuring effects, such as the use of a t-test or an analysis of covariance directly at the cut-off value. Also, they emphasised the importance of parallel and linear slopes in relation to the RDD between the assignment and the outcome variable on both sides of the cut-off value. Non-linearity, in contrast, was only briefly addressed with suggestions for data transformations or non-parametric regression approaches [26, 29].

In the 1970s Campbell continued to work on the RDD [26], supervising, among others, a dissertation by Joyce Sween [30], who introduced two innovations to the topic of RDD. First she argued that the RDD should not be viewed as a mere observational study but rather as an experiment, because a reliably implemented RDD generates the best knowledge about the group selection process, which in turn gives the experiment inferential power. Furthermore she described the handling of non-linear data in detail and showed within an example that the discontinuity, which occurred using a linear model, disappeared when using a quadratic model. In addition, Campbell and another student, T. Cook, worked more intensively on optimising the internal validity of RDD in comparison to other quasi experiments [26, 31].

Parallel to Campbell, other groups from the fields of psychology and statistics formed in the 1970s to focus on RDD. One of the later members of such groups, especially at North Western University, was William M. Trochim [26]. Trochim expanded the considerations and discussions on internal validity (see figure 1) as part of his research on RDD. In his 1984 book, he also describes the fuzzy RDD in more detail compared to the sharp RDD, thereby contradicting Campbell's statement that RDDs are only applicable in situations in which the groups are clearly assigned. Furthermore, Trochim is increasingly concerned with the practical and empirically testable assumptions of the RDD, thus introducing it to broader evaluation research and popularising it outside psychology [26, 32].

Despite its spread use in psychology and statistics, methodological work on the RDD stagnated in the 1990s [26]. However, from a methodological point of view, there were very similar approaches over the years, though under different designations. For example, there is a publication by Finkelstein et al. from 1996 that describes the construct in a clinical-statistical context under the name 'risk-based allocation method', but without linking it to the RDD described by Campbell or Trochim [26, 33, 34]. Also in the mid-1990s, researchers such as Angrist and Imbens addressed the issue of causal inference in econometrics and coined the term local average treatment effect (LATE) for the first time [26, 35, 36]. At this time, the methodology and interest in RDD, particularly as a tool for analysing causal relationships, came back into the focus of researchers, especially in econometrics [26].

So after the first wave of formalisation in the 1970s, RDD received a second boost in econometrics in the 2000s. Thereby, special attention was paid to the applicability of the design

and the use of thresholds to extract causalities from observational data [27]. Particularly influential in the early 2000s were the works of Hahn et al. and Lee, who demonstrated the high methodological potential of applying an RDD approach, and Imbens et al., who presented a practical guide for empirical economists [37–39]. In addition, the development of statistical software such as R, Stata, SAS and SPSS had helped to increase the applicability of RDD over the years. Packages such as rdrobust, rdd and rddtools, which are the most downloaded packages in R for the RDD approach (search from 24-04-2025)[40–43], enabled a broader community to use RDD not only as a design construct but also as an analysis method [27]. With the high availability through various software implementation solutions to a wider community over the last 15 to 20 years and the whole history of RDD, especially the parallel development of the methodology in different disciplines under various names, it is clear that the design and its validity can hardly be questioned [26, 27].
Nevertheless, despite the strong development, the growing use of RDD and its popularity in econometrics, the design was rarely used for studies in medicine, epidemiology and public health until around 2010. Even after that, the number of publications of studies using RDD rose only slowly compared to other disciplines [24, 27, 44]. Only in the last 5 to 10 years, RDD has been increasingly recognised and used as a suitable alternative to RCTs in the planning of clinical trials and health services research [27].
One of these studies was the *new form of care and integrated cross-sectoral psycho-oncology (nFC-isPO)* which was launched in 2018. This project is specifically concerned with reducing anxiety and depression and improving psycho-oncological psychosocial care for patients newly diagnosed with cancer.

1.2 Clinical background of the dissertation

Annually, about 500,000 people in Germany are diagnosed with cancer for the first time. In addition, the number of new cancer cases is expected to rise by 23% by 2030 due to demographic trends (i.e. the ageing of society) [45]. Apart from the rising incidence, cancer is still the second most common cause of death in Germany at 22% (see figure 2) [46].

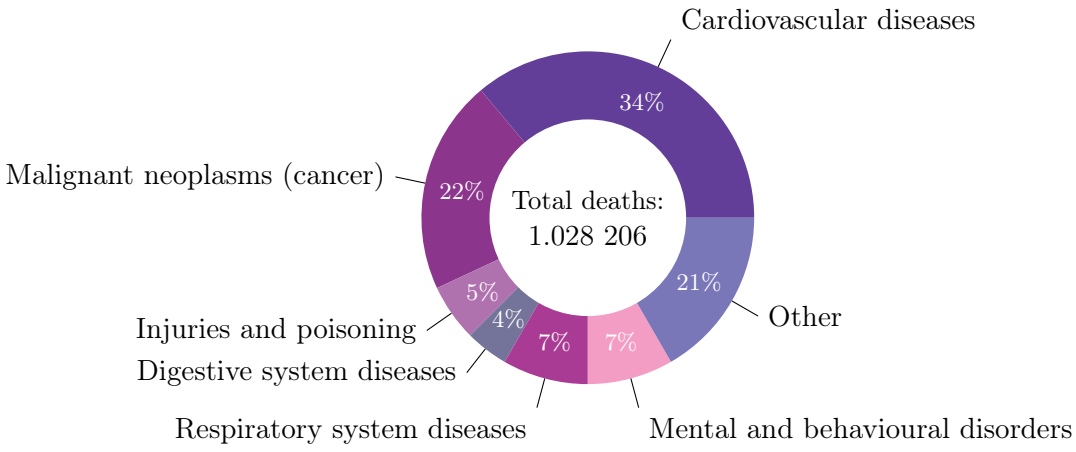


Figure 2: Mortality by types of diseases in 2023 in percentage [46]

And although the mortality rate has fallen in recent years, the relative 10-year survival rate is predicted to be only around 57% for men and 61% for women, despite the constant development of medical care and the latest procedures and treatment methods [45]. This means that, in addition to the physical challenges of the illness and the therapy, patients with cancer also have to contend with a heavy psychological burden due to abrupt changes in everyday life and fears of uncertainty caused by the illness and the topic of death [47]. It is therefore not surprising that almost 35% of all cancer patients suffer from psychological co-morbidity, anxiety and depression during the course of their treatment. This leads to an increase in emotional stress and thus to an increased need for medical and psychosocial care [45, 48]. In fact, around one in two cancer patients feels under severe psychological strain and around one in four suffers from increased depression [49].

For this reason, it is important to include not only physical but also mental health in the treatment and healing process. Based on this knowledge, a relatively young but important sub-field of oncology has developed over the last 50 years, since the 1970s. Psycho-oncology deals with all psychosocial aspects of cancer and is now an integral part of oncological care. It offers many interdisciplinary approaches that integrate psychological support services such as cognitive, systemic, depth psychological or creative therapies into cancer treatment. The well-being of patients and the continuous development of therapies are always at the centre of attention. Current research focuses on improving the quality of life among cancer patients through various coping strategies for anxiety, grief, pain, fatigue and depression [50, 51].

In order to establish and institutionalise psycho-oncology in Germany and create nationwide care, the Working Group Psycho-oncology (PSO) was founded within the German Cancer Society (DKG) in the 1980s [52]. In the 1990s, the PSO and the German Working Group for Psychosocial Oncology (dapo) began to offer psycho-oncological training and advanced education courses in order to improve the range of professional development courses available nationwide. This led to the formation of the association 'Further education and training in psychosocial oncology e. V.' (WPO) in the mid-2000s [52–54]. Shortly after the WPO was founded, the Federal Ministry of Health, together with cooperation partners such as the DKG, adopted the National Cancer Plan in 2008. Four specific fields of action were identified as part of this plan, with the second field of action 'Further development of oncological care structures and quality assurance' pursuing, among other things, the goals of ensuring that all cancer patients receive cross-sectoral, integrated oncological care (goal 7) and, if necessary, appropriate psycho-oncological care (goal 9) [55]. Based on this decision, the PSO, together with the Association of the Scientific Medical Societies in Germany (AWMF), published the interdisciplinary S3 'Evidence-based Guideline Psycho-oncological Diagnostics, Counselling and Treatment of Adult Cancer Patients' for the first time in 2014 as an important result of the National Cancer Plan. These guidelines serve as a valuable decision-making aid for doctors, other healthcare professionals and patients and contribute to improve the quality of medical care [56, 57]. In spite of the National Cancer Plan and the evidence-based S3 guideline on psycho-oncology, there was no adequate, low-threshold, prompt and comprehensive psycho-oncological care in the years that followed, as there was a lack of regulated funding in both the outpatient and inpatient sectors [49, 58]. However, several studies have shown that

psychosocial and psychotherapeutic interventions significantly reduce patients' psychological distress. But these studies have mainly focussed on individual cancer entities and showed a lack of systematic assessment of individual needs and the creation of integrated cross-sectoral structures [58].

These gaps were addressed with the above-mentioned nFC-isPO study in order to implement the demand of the Federal Government's National Cancer Plan for the further development of 'oncological care structures and quality assurance', including needs-based psycho-oncological care nationwide.

1.3 The nFC-isPO

The nFC-isPO project was funded by the Innovation Fund of the German Federal Joint Committee (GBA) for 54 months, from October 2017 to March 2022. It was carried out in four care networks in North Rhine-Westphalia (NRW) (GFO Kliniken Troisdorf, Johanna-Etienne-Krankenhaus Neuss, Kliniken Maria Hilf GmbH Mönchengladbach and University Hospital Cologne), including the involvement of the outpatient sector of general practitioners. The aim of the nFC-isPO support programme was to specifically determine the individual psychological stress of patients due to a new cancer diagnosis and to combat anxiety and depression as part of a 12-month, staged intervention programme. In addition to taking individual needs into account at patient level, the aim was to create an integrated, cross-sectoral care structure. This is also intended to achieve progress for psycho-oncological care in Germany at the level of the healthcare system with regard to the National Cancer Plan and to enable the programme to be included in standard care by the statutory health insurance funds.
In order to prove the effectiveness of the programme and the efficacy of the intervention, the project was scientifically monitored. isPO is a staged psycho-oncological care programme with 3 main levels. The enrolment in the programme took place at the hands of the treating oncologist, who issued a recommendation certificate for the newly diagnosed patients with cancer. The recommendation is based on the patient's decision to take part in the programme and the individual needs identified in the screening, which are determined using different questionnaires. The Hospital Anxiety and Depression Scale (HADS) [59] is used as the main allocation tool and determines which level of the isPO programme a patient is assigned to. In addition, the Psychosocial Risk Assessment Questionnaire (PSR) is used for a differentiated breakdown of psychosocial support needs within the level allocation [60].
Figure 3 shows the rough structure of the isPO care programme (part A) and where proof of efficacy is provided (part B).
First, all patients receive all necessary information about the isPO programme and the organisation of care within the programme (level 0, Case Management). At level 1, all patients then have the opportunity to discuss all cancer related topics and offers of help with a trained person, e.g. by the DKG, who has suffered from cancer themselves (Onco-Guide). Based on the HADS score (range: 0-42 points), a decision is then made as to whether a patient should be treated at level 2 (HADS ≤ 14) or level 3 (HADS ≥ 15), with higher HADS scores being

associated with a greater need for support. Alongside the established standard care, the patients at level 2 receive psychosocial support from psychosocial specialists. Level 3 patients receive psychotherapeutic care in addition to the regular standard care. Furthermore - and this is where the PSR becomes relevant - there is a subdivision within level 3 into patients with (PSR ≥ 6) and without (PSR ≤ 5) psychosocial support needs [58, 61].

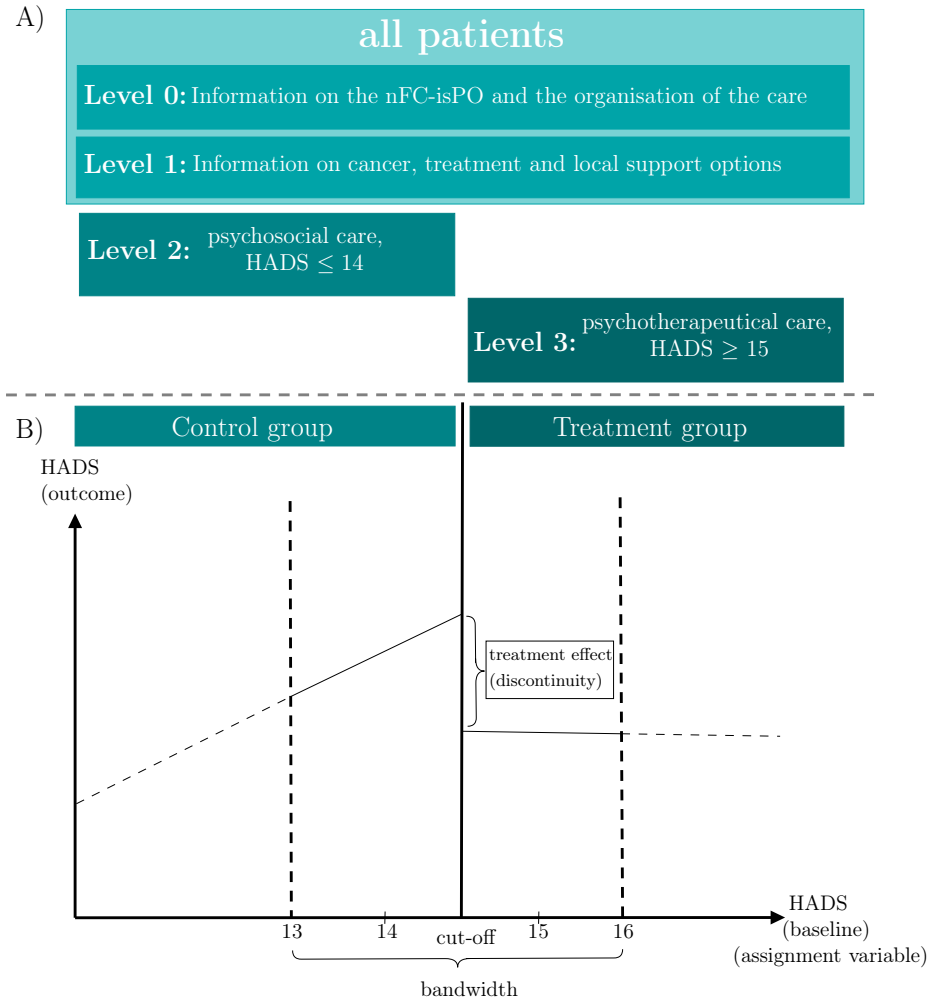


Figure 3: Representation of the isPO care programme including a schematic regression discontinuity design (RDD) diagram

The assignment to levels 2 and 3 as well as the proof of effectiveness was provided using the RDD as the study design and analysis method for the scientific monitoring of the project (see figure 3). Random allocation, as used in an RCT, was not possible for ethical reasons, among others, as all patients should be guaranteed care tailored to their individual needs. Therefore, allocation to needs-based treatment was based on a threshold value on the continuous HADS scale, which was determined in advance on the basis of clinically relevant aspects. This division led to the formation of the two study groups, with patients at level 2 forming the control group and patients at level 3 of the isPO programme forming the treatment group (figure 3).

In accordance with the different care in the two study groups, the effectiveness of complex

psycho-oncological psychotherapeutic care, which has not been routinely funded to date, was compared with the previous standard care, including care by a psychosocial specialist in isPO. In such comparisons, the RDD offers the special feature, like an RCT, of measuring causal inference close to the allocating threshold within a defined bandwidth (see figure 3). To test the effectiveness, the HADS was used again 12 months after baseline to measure patients' psychological distress and to show the change, ideally a reduction, in anxiety and depression as a result of the complex intervention (primary endpoint) [58, 61].
Since the measurement after 12 months offered a large gap to the baseline measurement, the HADS was also reassessed in isPO after 4 months in order to be able to depict a temporal course of the HADS values in addition to the primary endpoint. However, the change in HADS scores over time could not yet be taken into account in the RDD approach and was therefore only investigated descriptively and in secondary sensitivity analyses using alternative analysis methods.

1.4 Repeated Measures

Repeated measures studies such as the nFC-isPO, also known as longitudinal studies, are not uncommon in research, particularly in medical and health-related fields. They are used to measure differences and effects over time [62, 63]. A peculiarity of repeated measurements is the correlation between the respective measurements, as these are much more similar than two measurements from different subjects. This correlation must be taken into account in a statistical analysis by estimating the covariance parameters. [63, 64]. However, a simple linear regression cannot be used for repeated measurements, as it is assumed that the measurements are independent of each other [63]. Usually an analysis of variance for repeated measurements (RM-ANOVA) is used for the evaluation. With this method, significant differences in mean values are analysed, taking into account the correlation between the measures taken from the same individuals [65]. However, RM-ANOVA has some disadvantages because it assumes, among others, sphericity, i.e. the variances of the differences between all combinations of measurement time points must be the same, and this assumption is often violated, especially with more than two measurement repetitions. As mentioned at the beginning, it is important to consider the covariance structure of the parameters in the statistical analysis, but this is only possible to a very limited extent with RM-ANOVA, as fixed covariance structure is specified by the model [64–66]. Furthermore, RM-ANOVA cannot deal with missing values, so that incomplete data records are completely excluded (case-wise deletion), which in turn leads to an immense loss of information and a loss in power. The same applies to unbalanced data i.e. a different number of measurements per person [65–67]. Finally, RM-ANOVA cannot take into account complex structures in the data, such as random slopes, which significantly reduces the probability of detecting an effect [65, 68]. Due to all these disadvantages, researchers have been looking for alternative methods over the decades and have tried to further develop RM-ANOVA.
Theoretical-mathematical approaches for so called linear mixed models (LMM) have existed since the 1950s, but it was only with increasing computing power and the development of suit-

able software in the 1990s and 2000s that they could be increasingly used [64]. Today, LMMs are a popular and frequently used method in medicine for analysing longitudinal studies, as they offer higher flexibility compared to RM-ANOVA. They can deal with missing values, are robust against unbalanced data, and offer various possibilities for modelling covariance structures [63–65].
Due to the many advantages of LMMs and since the integration of repeated measurements has not yet been implemented in the RDD, the data from the nFC-isPO study was only analysed with LMMs in secondary analyses.

For this reason, the nFC-isPO with its longitudinal data and the RDD used in the project led to the question: "How can repeated measurements be included in the RDD approach?". The objectives and individual research questions of this dissertation are presented in more detail in chapter 2. Chapter 3 presents the basic concept of the RDD with a worked example within the first publication. Furthermore, the fundamentals of LMMs for repeated measures are introduced. Chapter 4 contains the application and results of the RDD in its classic, familiar sense for a sub-question within the nFC-isPO and thus demonstrates the application of the worked example presented using study data. This is followed by the main article of the dissertation on the development of the RDD methodology for repeated measurements using the data from the nFC-isPO. Lastly, the three articles are discussed in chapter 5 in the context of the dissertation's research question.

2 Aims and Objectives

The central aim of the cumulative dissertation is to further develop an existing alternative method to the RCT in order to make it usable for analyses with repeated measures and to offer scientists the opportunity to detect causal relationships with longitudinal data with as little bias as possible. In addition, the method is to be made more accessible and attractive to scientists so that the design is considered and applied more frequently in clinical and health science research. Therefore, the dissertation pursues two interlinked objectives. On the one hand, the introduction and application of RDD in the familiar sense in the research field of health sciences and, on the other hand, the methodological extension of the RDD for repeated measurements. The entire work was carried out as part of the nFC-isPO study and resulted in three peer-reviewed journal publications.

Each of the three publications deals with one of the above mentioned aspects. The individual questions and objectives of the articles are briefly explained as follows:

- **Introduction of the RDD:**
  The first article focuses on the methodological introduction of the RDD in the context of health services research. A worked example with medical research context and implementation program code for various popular statistical software is used to provide researchers with the necessary mathematical notations and descriptions of the application [69].

- **Application of the RDD:**
  During the nFC-isP study period, the coronavirus pandemic raised the question of how the influence of the pandemic could be taken into account when evaluating the study. The second article answers this question by applying RDD in the classic, familiar sense by means of a stratified subgroup analysis using the nFC-isPO data [70].

- **Methodological extension of the RDD:**
  The third article reflects the second and thus the main objective of the dissertation and deals specifically with the open problem that repeated measurements are not uncommon in medical research, but have not yet been taken into account in the RDD approach. The article therefore presents in detail a methodological extension of RDD in combination with LMMs for repeated measurements using data from the nFC-isPO [71].

By realising the above objectives, the dissertation contributes to the further development of a study design and analysis method relevant to the health sciences and makes its application accessible to a larger research group, particularly in a medical context.

3 Methods

In the previous chapters, the scientific background as well as the aims and objectives of the dissertation with the associated publications were conceptually contextualised. Building on this, the methodological basis of the work is presented in this chapter. The methods are introduced with the first published article, in which the theoretical and practical principles of RDD in the medical context are presented. Although this article is the first result of this dissertation, it forms the conceptual basis of the work and is therefore assigned to the methods chapter.
In addition, the methodology of linear mixed models for repeated measurements is presented in the course of the chapter. LMMs have a central role in the methodological extension of the RDD and thus form the methodological core of the dissertation and the third article. The integration of LMMs for repeated measures into the RDD approach enables a differentiated modelling of repeated measures within a quasi-experimental design.

3.1 Metholodical Overview of the RDD - 1. Article

Published as:

The regression discontinuity design: Methods and implementation with a worked example in health services research

Anna Hagemeier,
Christina Samel,
Martin Hellmich

Z Evid Fortbild Qual Gesundhwes., 2022 Aug; 172:71-77,
https://doi10.1016/j.zefq.2022.04.014

Visual Abstract:

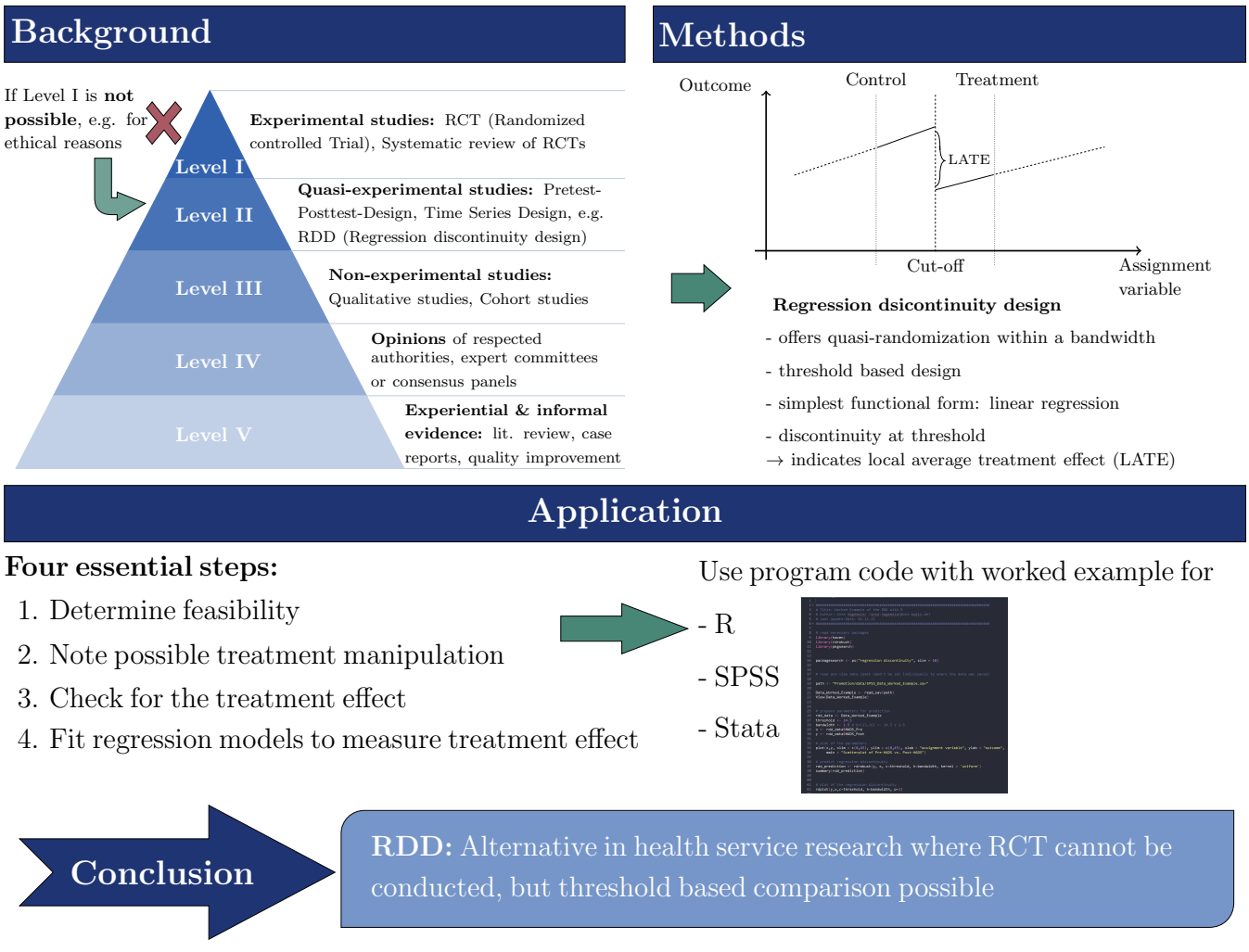


Figure 4: Visual abstract: "The regression discontinuity design: Methods and implementation with a worked example in health services research"

Z. Evid. Fortbild. Qual. Gesundh. wesen (ZEFQ) 172 (2022) 71–77

Contents lists available at ScienceDirect

Z. Evid. Fortbild. Qual. Gesundh. wesen (ZEFQ)

journal homepage: http://www.elsevier.com/locate/zefq

Versorgungsforschung / Health Services Research

The regression discontinuity design: Methods and implementation with a worked example in health services research

Das Regressions-Diskontinuitäts-Design: Methodik und Implementierung mit einem Arbeitsbeispiel in Versorgungs- und Gesundheitswissenschaften

Anna Hagemeier *, Christina Samel, Martin Hellmich

Institute of Medical Statistics and Computational Biology, Medical Faculty, University of Cologne, University Hospital Cologne, Cologne, Germany

ARTICLE INFO

Article History:
Received 14 December 2021
Revised 29 March 2022
Accepted 12 April 2022
Available online: 16 June 2022

Keywords:
Regression discontinuity
Health service research
Medical statistics
Quasi-experimental quasi-randomization
Statistical software

ABSTRACT

Background: The randomized controlled trial (RCT) is the gold standard in evidence-based medicine. However, this design may not be appropriate in every setting, so other methods or designs such as the regression discontinuity design (RDD) are required.
Method: The aim of this article is to introduce the RDD, summarise methodology in the context of health services research and present a worked example using the statistic software SPSS (Examples for R and Stata in the Appendix A). The mathematical notations of sharp and fuzzy RDD as well as their distinction are presented. Furthermore, examples from the literature and recent studies are highlighted, and both advantages and disadvantages of the design are discussed.
Application: The RDD consists of four essential steps: 1. Determine feasibility; 2. Note possible treatment manipulation, 3. Check for the treatment effect, and 4. Fit the regression models to measure the treatment effect.
Conclusion: The RDD comes as an alternative for studies in health service research where an RCT cannot be conducted, but a threshold-based comparison can be made.

ARTIKEL INFO

Artikel-Historie:
Eingegangen: 14. Dezember 2021
Revision eingegangen: 29. März 2022
Akzeptiert: 12. April 2022
Online gestellt: 16. Juni 2022

Schlüsselwörter:
Regressionsdiskontinuität
Versorgungsforschung
Medizinische Statistik
Quasi-experimentelle Methoden
Statistische Software

ZUSAMMENFASSUNG

Hintergrund: Die randomisierte kontrollierte Studie (RCT) ist der Goldstandard in der evidenzbasierten Medizin. Dieses Design kann jedoch nicht in jedem Umfeld eingesetzt werden, sodass andere Methoden oder Designs wie das Regressions-Diskontinuitäts-Design (RDD) erforderlich sind.
Methode: Dieser Beitrag führt in das RD-Design ein, fasst die Methodik im Kontext der Versorgungsforschung zusammen und stellt ein praktisches Arbeitsbeispiel unter Verwendung der Statistiksoftware SPSS vor (Beispiele für R und Stata im Anhang). Vorgestellt werden die mathematischen Notationen der scharfen und unscharfen RDD sowie deren Unterscheidung. Darüber hinaus werden Beispiele aus der Literatur und aktuellen Studien hervorgehoben und die Vor- und Nachteile des Designs diskutiert.
Anwendung: Das RDD besteht aus vier wesentlichen Schritten für die Anwendung: 1. Die Durchführbarkeit bestimmen, 2. Mögliche Manipulation der Behandlung beachten, 3. Behandlungseffekt überprüfen und 4. Die Regressionsmodelle anpassen, um den Behandlungseffekt zu messen.
Schlussfolgerung: Das RDD stellt eine Alternative für Studien in der Versorgungsforschung dar, bei denen keine RCT durchgeführt werden kann, aber ein schwellenwertbasierter Vergleich möglich ist.

* Corresponding author: Anna Hagemeier, M.Sc. Institute of Medical Statistics and Computational Biology (IMSB), University Hospital of Cologne, Kerpener Str. 62, 50937 Cologne, Germany.
E-mail address: anna.hagemeier@uni-koeln.de (A. Hagemeier).

https://doi.org/10.1016/j.zefq.2022.04.014
1865-9217/

Introduction

After many years of exploring, developing and refining various study designs for medical research, the randomized controlled trial (RCT) is still the gold standard. However, even if an RCT is desirable, it cannot be conducted in every setting as the study design needs to address the specific research situation. Coly et al. [1] recommend the choice of design to conduct a study and mention why sometimes, due to limitations such as ethical reasons (e.g. detrimental treatment), financial and feasibility constraints, quasi-experimental designs are a suitable alternative to experimental studies like RCTs.

Just as in experimental designs, quasi-experimental designs examine whether interventions have an objectively measurable effect by testing causal hypotheses, except that group allocation is not random. Instead they offer different allocation mechanisms that can produce a quasi-randomization [2,3]. Which quasi-experimental design should be chosen as an alternative depends on the setting since there is not "the one" design for evaluation [1]. In healthcare research, the quasi-experimental regression discontinuity design (RDD) can be well suited for medical and health service research in situations where the allocation of an intervention is regulated by a threshold [4,5]. However, Moscoe et al. [6] mentioned it as underutilized and underrated in those fields.

The RDD originates from econometrics and is a threshold-based study design. Means, allocation is regulated by a threshold, i.e. group allocation (treatment or control) depends on whether the value of the assignment variable falls below or above a specific threshold. It was first described by Campbell and Thistlethwaite in 1960 [7] and has recently been rediscovered for the medical and health services research field, although there is little relevant literature in statistics and healthcare research. The most popular example beyond medicine or health research are the results of the SATs (Scholastic Assessment Test as assignment variable). Students above a predefined threshold (cut-off value c) will receive a scholarship for college and success (outcome) with or without the scholarship is measured by their income after finishing college [8].

This article aims to describe and summarize the methodological basis of the RDD in the context of medical and health service research. Furthermore, strengths and weaknesses are discussed, possibilities for statistical implementation are presented, including a worked example and a prospective view of open issues is given.

The Design

The appropriate design to answer a given research question is not necessarily obvious and definite. However, if an intervention is already established and cannot be assigned purely at random, e.g. for ethical reasons, but a threshold criterion could be included in the study design, RDD is a possible evaluation strategy. This threshold rule defines whether a participant receives the intervention or control treatment, therefore it specifies the treatment allocation. For allocation, any continuous variable, which can be split at a certain point, can be chosen as the assignment variable (X_i). The treatment allocation rule then needs to be verified, usually based on a scatterplot with the assignment variable on the x-axis and the outcome variable on the y-axis. Inclusion of the cut-off value to the plot can show the vertical separation line between the intervention and control group. If the allocation is defined in a strictly binary fashion with assigned probabilities of 0 or 1 for control or intervention, respectively, the clearly defined separation of the groups is known as the sharp design. Figure 1 shows the schematic draft of the sharp RD-Design with fitted regression lines. When a strictly binary assignment is impossible, a fuzzy assignment or fuzzy RDD is given (Figure 2). In a fuzzy RDD, the above-mentioned allocation probability is not sharply disjunct around the cut-off c. This can lead to bias, e.g. unobserved covariates and confounders (see section 'Sharp vs. fuzzy designs'). Hence, the sharp design is always preferable and should be aimed for [3]. The main idea behind this assignment procedure is the comparability just around the threshold. It is obvious that there is an overall difference between participants above and below the threshold. Still, close to the threshold, this difference is very small and one can measure the causal effect by comparing the mean difference between the outcomes within intervals to the left and right of the threshold. Hence, it is similar to randomization in an RCT but only limited to a pre-defined area (bandwidth b) near the threshold value. Getting a precise prediction without bias within this bandwidth is not trivial. If the bandwidth is large, estimation of

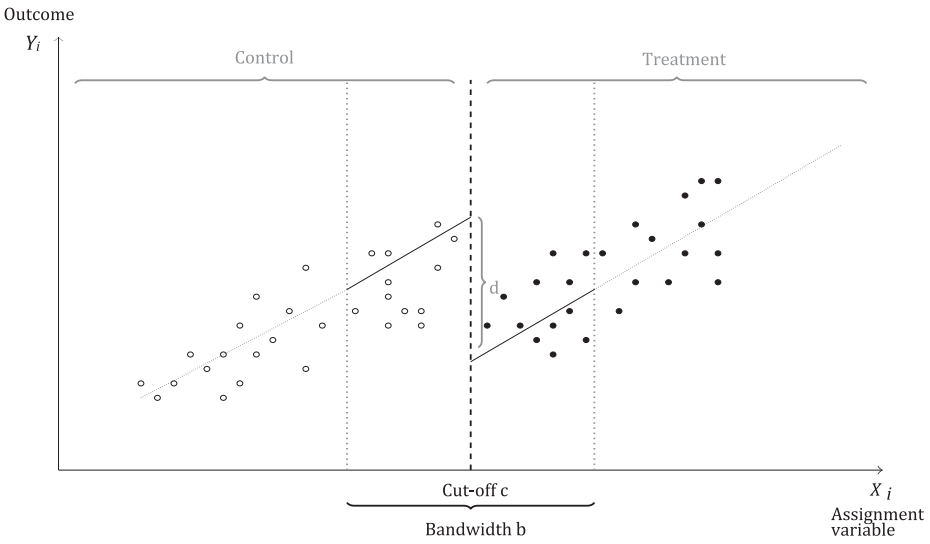


Figure 1. Schematic draft of the sharp regression discontinuity design with a displayed treatment effect 'd' as the discontinuity. (Unfilled points represent the observations in the control group, filled ones the patients in the treatment group.)

A. Hagemeier et al./Z. Evid. Fortbild. Qual. Gesundh. wesen (ZEFQ) 172 (2022) 71–77

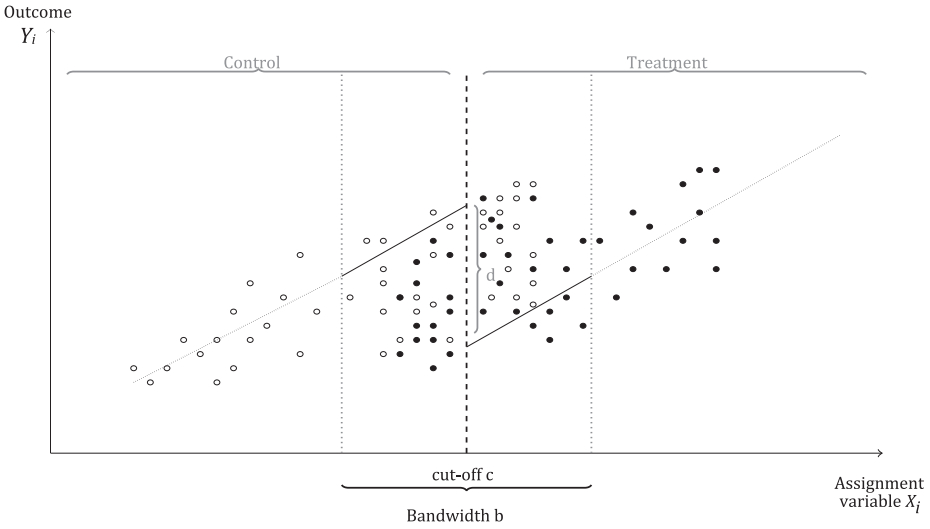


Figure 2. Fuzzy RDD without sharp treatment allocation. Patients can or cannot receive treatment although their assignment value is above or below the cut-off value. The decision is therefore not strictly based on one assignment-variable and can be changed due to other circumstances. (Unfilled points represent the observations in the control group, filled ones the patients in the treatment group.)

the treatment effect might be more precise, but the risk that observations within the bandwidth below and above the cut-off value no longer match in their basic characteristics and are therefore no longer comparable, is high. To determine the optimal width for the comparison interval, an iterative approach, e.g. cross-validation, should be used [9,10]. However, in medicine or health science, the width is often a predefined value, which is added to or subtracted from the cut-off value ($c - b \leq X_i \leq c + b$). If not pre-defined, the distance should be extended until the samples below and above the cut-off value lose comparability. After finding the optimal bandwidth, regression lines are fitted on both sides of the cut-off value between the bandwidth. One can also continue the regression lines for the entire data but only to visualize a trend in the data. For comparability, one can only use the sample limited to the determined bandwidth (local "as if" randomization idea). Hence, a common problem is the necessity of a large sample size to obtain sufficient subjects within the bandwidth [2,3].

For mathematical estimation of the causal treatment effect within the bandwidth instrument variables or indicator functions (*Ind*) are needed to disjoin the control group from the treatment group. The most commonly used causal effect estimator in connection with the RDD is the local average treatment effect (LATE), also called the complier average causal effect (CACE). LATE/CACE is about estimating the effect of an intervention on subjects actually receiving the treatment. In settings with strict sharp group assignment, there are only "compliers," i.e., each subject was correctly assigned and receives the correct treatment (full compliance). This corresponds to the described assignment with binary probabilities 0 and 1. In fuzzy RDD the probabilities of receiving treatment or control differ by the threshold. Therefore, researchers must address the problem of incomplete compliance, as some subjects receive or do not receive the treatment as assigned. In this regard, there are similarities in the use of LATE/CACE to the role of the per-protocol (PP) analysis in clinical trials. LATE/CACE may be estimated by performing an intention-to-treat (ITT) analysis and then dividing the obtained result by the difference in proportion (probability) of participants who have been treated but were assigned to the control group or were assigned to the treatment group but did not receive treatment. Usually this is done by a two-stage least squares method [10]. Because this paper aims to declare the simplicity of the RDD, present it as a good alternative to RCTs and show its implementation in the medical field, the reader can receive more detailed information on LATE/CACE and their calculation within the recommended articles [11–13].

According to the literature, there are many functional possibilities to perform causal inference with the RDD. The most common and simplest one is linear regression [14]. Function (3) represents the case where the regression lines have the same slope on each side of the cut-off value (c)

$$Y_i = \beta_0 + \beta_1 x + \beta_2 Ind(x > c) + e \tag{3}$$

The notation used for the mathematical description of the linear model is defined as follows [5,15]:

Y_i : continuous outcome with $i = 1, 2, \ldots$, n observations, β_0 : intercept, β_1 : slope, x: covariate of X_i (continuous assignment variable) with $i = 1, 2, \ldots$, n observations, β_2: measures the treatment effect ("jump" at the cut-off value, equals d in Figure 1 and means the discontinuity as the difference between the two regression functions), *Ind* : indicator function or treatment indicator (sharp design)= 0 if x < c (control) or 1 if x ≥ c *(treatment), c :* cut-off value, e : random error

If equation (3) would be extended to allow for a change in slope at the threshold, the regression model would become:

$$Y_i = \beta_0 + \beta_1 x^* + \beta_2 Ind(x > c)x^* + \beta_3 Ind(x > c) + e \tag{4}$$

The equation is extended by the following terms [3,6]:

x^*: x - c; distance to the cut-off, β_3 : measures the treatment effect ("jump" at the cut-off value, equals d in Figure 1 and means the discontinuity as the difference between the two regression functions), β_1 :slope below the cut-off, $\beta_1+\beta_2$: slope above the cut-off

The expansion to a polynomial is possible, but this will not be considered further here.

The RDD in Figure 1 is an example of an effect based on the treatment (no interaction, regression lines run parallel). The treatment effect can be identified around the threshold by e.g. simply calculating the difference of the intercepts of both regression lines on each side of the cut-off or graphically compare the regression lines and scatterplots. To clarify the simple idea behind the RDD, the researcher can restructure the study data at the cut-off value

into two subsets, conduct linear regressions on both subsets and compare the intercepts of both regression functions. The difference between the intercepts indicates the discontinuity and thus the treatment effect, if lines run parallel. If the slopes of the regression lines differ (interaction effect) the researcher can also simply calculate the differences of intercepts but needs to shift the assignment variable by the cut-off value to zero (x-c). Another possibility would be to conduct a simple *t*-test on the groups within the bandwidth b around the cut-off value [c-b; c+b], but the treatment effect then might be underestimated [3,7,16].

Since the assignment variable can be any continuous measure, many possible uses of the RDD in medicine or health service research are imaginable. For example, Maciejewski & Basu [16] mentioned the use of the RDD in conjunction with the neonatal intensive care unit (NICU). New-borns underweight (<1500g) more often need intensive care then heavier infants. Thus, one can compare the difference and effect of intensive care at a cut-off value where infants just around the cut-off should be almost identical and weight would be the assignment variable. Another well-known assignment variable for implementing the RDD in the medical and health service research field with application to psychology is the hospital anxiety depression scale (HADS). The HADS consists of 14 questions, 7 of which relate to depression and 7 to anxiety, and has a total score ranging from 0 to 42 [17]. A recent example using the RDD and the HADS as the instrument for group allocation (assignment variable) is the psycho-oncological study isPO – (integrated cross-sectoral psycho-oncology), where the threshold is set between 14 and 15 score points. Every patient receives either standard psychological care (below threshold) or individualized/special treatment after cancer diagnosis. After 12 months, evidence of efficacy will be provided by re-interviewing patients with the HADS (outcome) [18].

Sharp vs. fuzzy designs

As previously described, to get close to a randomized experiment, a sharp RD-design is the desired approach. However, individual and complex scenarios or incorrect treatment allocation can lead to fuzziness regarding the groups (Figure 2).

As a result, some patients receive treatment even though they are below the cut-off according to the allocation variable and some patients do not receive treatment while exceeding the threshold. Bor et al. [3] describe such a violation of treatment allocation using the example of HIV infections where there are two possibilities to receive treatment. Either the CD4 count drops below a predefined threshold or the symptoms indicate the severity of their disease. Another very simple example comes from Maciejewski & Basu [19]: hemoglobin values at a certain threshold are used to decide whether a patient needs a blood transfusion or not. But sometimes, although the hemoglobin value is in normal range, the patient needs a transfusion because of other clinical factors. Of course, the decision can also be made in the other direction. Thus, the allocation is based partly on a threshold rule and partly on clinical judgment, allowing switching from treatment to no-treatment and vice-versa [12,20]. As a result, an RDD with imprecise threshold allocation is strictly not a real RDD (fuzzy RDD). Linear regression approaches or non-parametric methods can be used for both the sharp and fuzzy RDD versions. Overall, however, one should always seek for a solution that leads to a sharp RDD, e.g. try to control for observed covariates [21] or, if necessary, opt for a different design [2,22].

Possible effect scenarios

Considering the sharp RD-designs with equal assumptions about the patients' characteristics at both sides of the cut-off, there are four theoretical outcomes for treatment effect (Figure 3). First option is panel A where no treatment effect is apparent. This would also be the case when no treatment was performed. The regression lines would be congruent. In panel B an interaction effect is apparent. This is shown by a change in the slope of the regression line. This means that the effect of treatment increases or decreases with increasing values of the assignment variable above the cut-off value in the treatment group but is still equal (no visible jump)

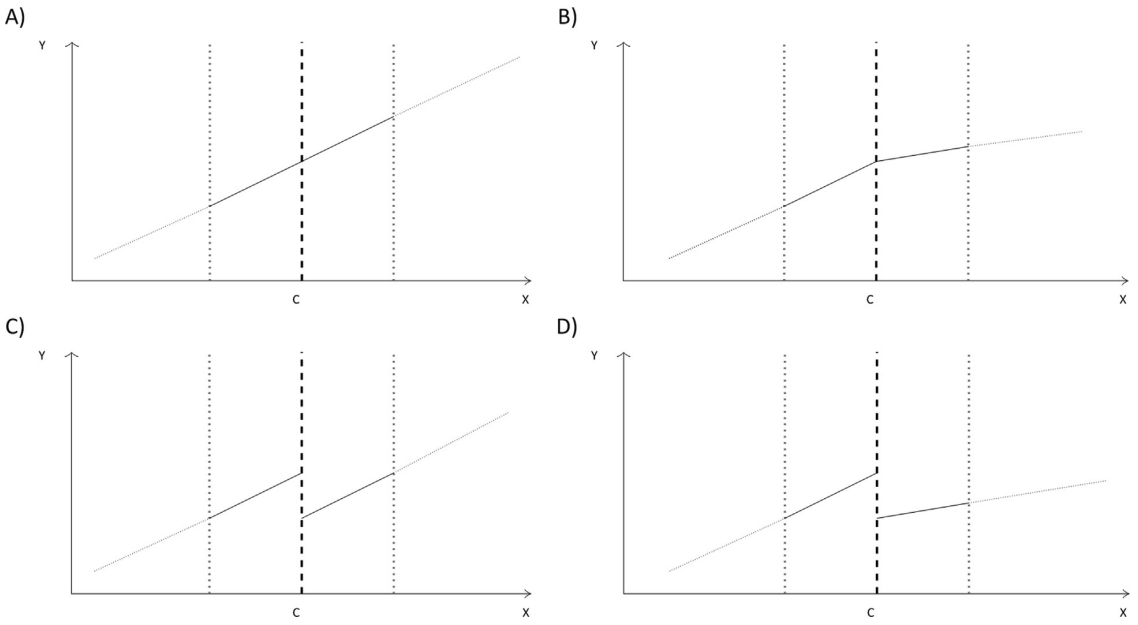


Figure 3. Schematic outcome possibilities in RDDs with X as the assignment variable, C as the Cut-off with bandwidth and Y as the outcome variable; A) No treatment effect, B) no main effect, but interaction effect, C) main effect for treatment and D) main and interaction effect.

to that of the untreated group at the cut-off. Panel C shows the essential RDD idea with only a main treatment effect with a discontinuity (jump) of the regression lines at the cut-off value. Panel D is a combination of panel B and C where an interaction effect (treatment effect only for those with a higher assignment value, change in slope) and a main effect (treatment effect along the entire assignment scale, discontinuity at the cut-off, change in intercept) occur [20].

Strengths and weaknesses of the RDD in theory and practice

When used appropriatly, RDD has several strengths compared to other study designs. The main advantage of RDD is the (local) quasi-randomisation around a specific threshold, which avoids ethical problems in treatment allocation, i.e. treatment is still allocated based on individual needs. However, this requires a specific allocation criterion (threshold). In practice, administrative or secondary health insurance data may be used to evaluate clinical interventions, making the main disadvantage of RDD easier to address. This disadvantage is that the RDD usually requires a large amount of data to precisely estimate the treatment effect near the threshold. However, by using retrospective, already or routinely collected data, a proof of effectiveness in real-life settings may still be possible. In prospective studies the large amount of data required is a major problem, mainly because of the high financial costs entailed by data collection. Note that manipulation with the threshold, e.g. by intentionally giving poorer answers to a questionnaire to achieve a higher score and thus be more likely to get into the treatment group, can lead to a bias in local randomisation, both in retrospective and prospective designs. [1,2,23].

Application of Regression Discontinuity

The following four steps are essential for practical implementation and analysis of the RDD [6].

1. Determine feasibility
   a. Continuous assignment variable (X)
   b. Outcome (Y) observable for both groups (treatment/no treatment)
   c. Clear assignment rule to determine if design is sharp or fuzzy
2. Watch for opportunities for allocation manipulation to control the treatment status. One opportunity to detect manipulation would be a histogram of the assignment variable (X). If the distribution is significantly skewed, that would suggest a way to manipulate the allocation and thus the treatment status. To guarantee the comparability of the groups, covariate balance tests can be performed for the allocation variable X.
3. Check for treatment effect while plotting the result (Outcome Y) against the assignment variable (X). A visible jump near the cut-off value confirms discontinuity and indicates a treatment effect.
4. Depending on the data availability and other factors, fit the regression models to measure the treatment effect
   a. around the cut off with a local linear regression within the bandwidth and/or
   b. over all data.

Implementation with a worked example

Statistical programs like R, Stata or SPSS already offer sound implementations for analysis with the RDD [24–28]. In the following we show a more detailed working example for SPSS. For R and Stata, you will find examples with corresponding program code on the same data basis in the Appendix A. Since the HADS has already been mentioned as an established instrument in the medical context and as an example assignment and outcome variable for the RDD, the worked examples are performed with simulated HADS data in a fictitious study setting. In this setting patients will receive standard psychological care underneath a pre-defined threshold and intensive care above the threshold. The following assumptions are made for the example:

- HADS T1 will be assignment variable and HADS T2 the outcome variable
- Patients in the intensive care group improve their HADS (T2) (decreases from at least 2 up to 7 points)
- Patients in the standard care group improve their HADS (T2) (decreases 0 up to 1 points)
- For simplification, the HADS can also assume non-integer values between 0 and 42 and underlies a uniform distribution.
- The pre-defined cut-off value of $c = 14.5$ and a bandwidth $b =$ [13–16] is assumed.

The program code for generating the data can be found in the Appendix A. The scatter plot of the data from $n = 500$ subjects and the threshold line at $c = 14.5$ is shown in Figure 4.

A worked example with SPSS

SPSS offers two options to implement the RDD analysis.

Strictly speaking, the first option is not an option with SPSS alone, but rather a combination with R. SPSS Version 18 or higher in combination with the 'Essentials for R' and the extension bundle 'STATS_RDD' offers a simple and effective possibility to analyse a regression discontinuity (Use 'STATS RDD /HELP.' for more information in SPSS). The R package used for the extension is 'rdd' implemented by Drew Dimmery [29,30]. For analysis of the given sample data set (Program code 1) follow the path 'Analyze>Regression>Regression Discontinuity' in SPSS or use the syntax (Program code 2) to reproduce the following results. Important, for the comparison with the simple linear regression, the option "Rectangular" must be selected in the package STATS_RDD for the option Kernel.

Table 1 shows the main result table of the performed STATS RDD with SPSS. One can read the significant local average treatment effect (LATE = −4,704, $p = 0{,}001$) within the pre-defined bandwidth ($b = 14.5 \pm 1.5$) as well as the number of cases ($n = 35$). This would mean in the fictitious study setting that patients with a T1-HADS (assignment variable) value > 14.5 improve their anxiety and depression scores significantly.

Second option is the before mentioned approach via simple linear regression (see section 'The Design') in SPSS with specified interaction terms. Use Program code 3 to reproduce the results. The results via simple regression will differ slightly from the results via STATS_RDD function because one needs to pre-define the bandwidth for the simple regression analysis via filtering the data and there might be rounding differences within the different approaches. Because the linear regression does not show the number of cases one should simply calculate the frequency of the grouping variable within the filtered data (HADS<=14.5: $n = 23$ (65.7%); HADS>14.5: $n = 12$ (34.3%); Total: $n = 35$).

Table 2 contains the results from linear regression where one can see the same significant average treatment effect as in the above STATS_RDD results. The coefficient for the model parameter HADS_group (Coefficient = −4,704, $p = -0.004$) is simply the difference of the intercept coefficients of the regression lines at the cut-off and thus the local average treatment effect.

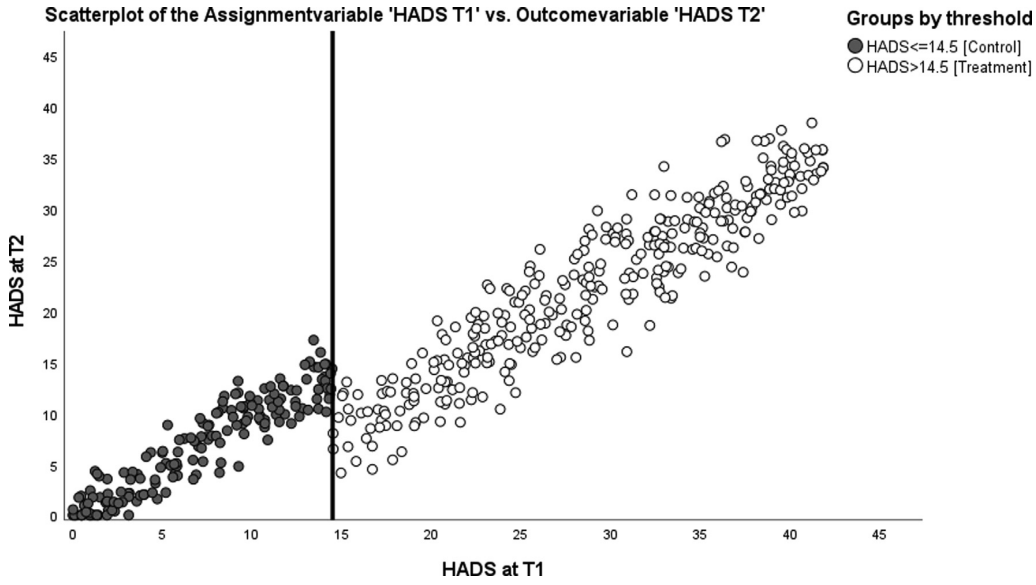


Figure 4. SPSS-Scatterplot of the fictitious/simulated HADS data.

Table 1
Results of regression discontinuity design in SPSS using 'R Essentials STATS_RDD'.

	Bandwidth	Number of Cases (n)	Average Treatment Effect Estimate	SE	p-value
LATE	1,5	35	−4,704	1,410	0,001
Half-BW*	0,75	19	−5,465	1,147	<0,001
Double-BW*	3,0	64	−5,029	1,100	<0,001

* the rdd package tests also half and double of the bandwidth, SE: standard error

Table 2
Results of linear regression with interaction terms to perform an RDD. (for explanation of model parameters see comments within Program code 1).

Model		Coefficient	SE	p-value	95% Confidence Interval
1	(Constant)	42,521	8,219	<0,001	[25,800 to 59,242]
	HADS_Pre	−2,148	0,573	0,001	[−3,314 to −0,982]
2	(Constant)	19,169	16,190	0,245	[−13,852 to 52,189]
	HADS_Pre	−0,432	1,166	0,714	[−2,811 to 1,947]
	HADS_group*	−4,704	1,530	0,004	[−7,824 to −1,583]
	Shift	1,630	1,948	0,409	[−2,343 to 5,603]

Dependent Variable: HADS_Post;
* HADS_group: treatment effect, SE: standard error

Conclusion and open issues

The RDD offers a possibility to compare different interventions and measure treatment effects in a locally randomized sense. Thus, the risk of confounding may be low in contrast to other quasi-experimental study designs. Nevertheless, the design still appears underexplored and rarely used in health sciences. This is probably due to the widespread focus on randomized evidence and the absence of specific thresholds to base the treatment decision on [7,31]. Which also confirms and highlights the need for further research in RDD. For example, the RDD methodology might be extended to the analysis of outcomes measured repeatedly over time and on the same subjects.

Acknowledgement

The preparation of this manuscript was partially funded by the Innovation Fond of the Joint Federal Committee in the form of the project "isPO—Integrated, cross-sectoral psycho-oncology" (grant number: 01NVF17022-IsPo). A big thank you also to Anne Adams, Stefanie Hamacher and Jeremy Franklin for their comments on the manuscript which improved it considerably.

Conflict of interest

The authors declare no conflicts of interest.

CRediT author statement

Anna Hagemeier: Conceptualization, Writing- Original draft preparation, Visualization, Writing – review & editing, Methodology. Christina Samel: Conceptualization, Writing – review & editing, Methodology. Martin Hellmich: Supervision, Conceptualization, Writing – review & editing, Methodology.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.zefq.2022.04.014.

References

[1] Coly A, Parry G. Evaluating complex health interventions: A guide to rigorous research designs 2017:1–27.
[2] White H, Sabarwal S. *Quasi-experimental design and methods.* Methodological Briefs: Impact Evaluation 2014;8:1–16.
[3] Bor J et al. Regression discontinuity designs in epidemiology: Causal inference without randomized trials. Epidemiology 2014;25(5):729–37.
[4] Venkataramani AS, Bor J, Jena AB. Regression discontinuity designs in healthcare research. BMJ 2016;352:i1216.
[5] Choi JY, Lee MJ. Regression discontinuity: Review with extensions. Statistical Papers 2017;58(4):1217–46.
[6] Moscoe E, Bor J, Barnighausen T. Regression discontinuity designs are underutilized in medicine, epidemiology, and public health: A review of current and best practice. Journal of Clinical Epidemiology 2015;68(2):122–33.

[7] Cook TD. "Waiting for Life to Arrive": A history of the regression-discontinuity design in Psychology, Statistics and Economics. Journal of Econometrics 2008;142 (2):636–54.
[8] Thistlethwaite DL, Campbell DT. Regression-Discontinuity Analysis: An Alternative to the Ex-Post Facto Experiment. Observational Studies 2016;2: 119–209.
[9] Imbens G, Kalyanaraman K. Optimal Bandwidth Choice for the Regression Discontinuity Estimator. Review of Economic Studies 2012;79(3):933–59.
[10] Jacob R et al. A practical guide to regression discontinuity. MDRC 2012.
[11] O'Keeffe AG, Baio G. Approaches to the Estimation of the Local Average Treatment Effect in a Regression Discontinuity Design. Scandinavian Journal of Statistics: Theory and Applications 2016;43(4):978–95.
[12] Imbens GW, Lemieux T. Regression discontinuity designs: A guide to practice. Journal of Econometrics 2008;142(2):615–35.
[13] Angrist JD, Imbens GW, Rubin DB. Identification of causal effects using instrumental variables. Journal of American Statistical Association 1996;91 (434):444–55.
[14] Angrist JD, Pischke J-S. Mostly harmless econometrics: An empiricist's companion. Princeton University Press; 2008.
[15] Calonico S, Cattaneo MD, Titiunik R. rdrobust: An R Package for Robust Nonparametric Inference in Regression-Discontinuity Designs. R Journal 2015;7(1):38–51.
[16] Chambers S. Regression discontinuity design: A guide for strengthening causal inference in HRD. European Journal of Training and Development 2016;40(8–9):615–37.
[17] Snaith RP. The hospital anxiety and depression scale. Health Qual Life Outcomes 2003;1:29.
[18] Jenniches I et al. Evaluation of a complex integrated, cross-sectoral psycho-oncological care program (isPO): a mixed-methods study protocol. BMJ Open 2020;10(3):e034141.
[19] Maciejewski ML, Basu A. Regression Discontinuity Design. JAMA 2020;324 (4):381–2.
[20] Cook TD, Campbell DT, Shadish W. Experimental and quasi-experimental designs for generalized causal inference. Houghton Mifflin Boston; 2002.
[21] Frölich M, Huber M. Including covariates in the regression discontinuity design. Journal of Business & Economic Statistics 2019;37(4):736–48.
[22] Arnold M et al. Introduction to Econometrics with R. Retrieved 2019;7:2019.
[23] Tang Y, Cook TD. Statistical power for the comparative regression discontinuity design with a pretest no-treatment control function: Theory and evidence from the National Head Start Impact Study. Evaluation Review 2018;42(1):71–110.
[24] Team, R.C. A language and environment for statistical computing. R Foundation for Statistical Computing; 2019.
[25] StataCorp. Stata Statistical Software: Release 17. College Station, TX: StataCorp LLC; 2021.
[26] Calonico S et al. rdrobust: Software for regression-discontinuity designs. Stata Journal 2017;17(2):372–404.
[27] Cattaneo MD, Idrobo N, Titiunik R. A practical introduction to regression discontinuity designs: Foundations. Cambridge University Press; 2019.
[28] Cattaneo MD, Titiunik R, Vazquez-Bare G. Inference in regression discontinuity designs under local randomization. Stata Journal 2016;16(2):331–67.
[29] Dimmery D. Package 'rdd'. Manual for the statistical software R 2016.
[30] IBM Corp, N. IBM SPSS Statistics for Windows, Version 26. Armonk, NY: IBM Corp; 2019.
[31] Flannelly KJ, Flannelly LT, Jankowski KR. Threats to the internal validity of experimental and quasi-experimental research in healthcare. Journal of Health Care Chaplaincy 2018;24(3):107–30.

3.2 Linear mixed models for repeated measures

Besides the previously described basics of RDD, linear mixed models for repeated measures form the basis for the methodological extension of RDD for repeated measures. They are a further development of classic linear regression models, which model the relationship between a dependent, normally distributed, continuous variable (outcome) and one (univariate) or more (multivariable) independent continuos or categorical variables (predictors). The aim of this modelling is to find an equation that describes the observed data as accurately as possible in order to quantify and predict the relationship between outcome and predictors [72, 73]. For this purpose, simple univariate linear regression represents the basic functional form [74] (see equation 3.1 and figure 6 panel A):

$$y_i = \beta_0 + \beta_1 x_i + \epsilon_i \tag{3.1}$$

Where y_i is the value of the dependent variable for the i-th observation ($i = 1, ..., n$), β_0 is the intercept, β_1 is the coefficient for the slope, x_i is the value of the independent variable for the i-th observation and ϵ are the independent and identically distributed ($\sim N(0, \sigma^2)$) error terms with ϵ_i as the error term of the i-th observation, also known as the residual. The residuals reflect the difference between the observed value and the predicted regression line. In other words, the residual variance is a measure of how spread out the residuals are overall (see figure 6 panel A).

If there is more than one independent variable, equation 3.1 expands to a multivariable linear regression, i.e. equation 3.2 [73–76]:

$$y_i = \beta_0 + \beta_1 x_{i1} + \cdots + \beta_K x_{iK} + \epsilon_i \tag{3.2}$$

The difference to a univariate linear regression consists in the number of independent variables (predictors) x_{ik}, which represent the values of the k-th variable for the i-th observation ($k = 1, ..., K$). The individual β_k estimators, which are assigned to the predictors within the multivariable linear regression, represent the respective partial slope of an influencing factor. As a consequence, they indicate how the outcome (y_i) changes if the respective predictor (x_{ik}) increases by one unit - assuming that all other predictors remain constant. Thus, an isolated effect of an influencing factor on the outcome is shown [75, 76].

However, the standard linear regression approach is limited to the analysis of data from cross-sectional studies and is therefore fundamentally based on the assumption that the observations are independent of each other. In hierarchical data structures, i.e. multi-level structures within the data, this independence of observations is lacking, which means that individual observations correlate with each other. Figure 5 shows a simplified diagram of a hierarchical 2-level data structure for repeated measurements in individual patients. Here, level 1 comprises all measured values, e.g. a score at a time point T_j for $j = 1, ..., J$. These repeated measurements are grouped (nested) within a patient i for $i = 1, ..., n$, thus forming the second level. This means that each patient at the higher (second) level is linked to several

measurements at the lower (first) level [77, 78].

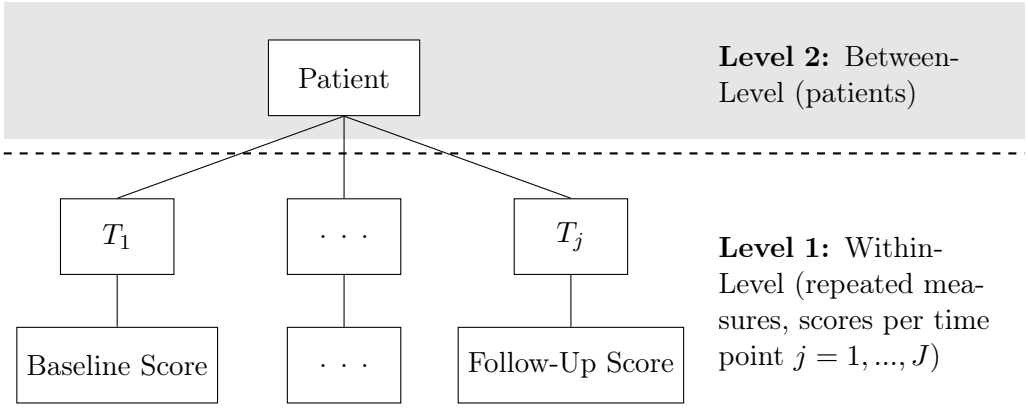


Figure 5: Schema of a hierarchical data structure with two levels

Consequently, modelling hierarchical data structures with simple linear regressions would, for example, lead to an underestimation of the standard errors, as it does not take into account the dependency of repeated measures within the same patient. And that, in turn, leads to an overestimation of the statistical significance [79, 80]. To account for the dependencies between individual observations and possible resulting distortions, such data must be modelled differently. Hierarchical or LMMs are particularly well suited and are often used to estimate hierarchical data structures, such as in longitudinal data with repeated measurements. LMMs for repeated measurements therefore incorporate the correlation of repeated measurements within each patient [63, 74].

For the modelling process, the combination of fixed and random effects is essential. Fixed effects, as they also occur in conventional linear regression, are based on the assumption that certain factors have the same effect on all patients, i.e. the estimates do not vary between patients. They therefore describe a general trend, i.e. they capture systematic influences that apply to all individuals. In the case of random effects, on the other hand, the variation from patient to patient (individual-specific effect) is taken into account in the modelling. In the case of random effects, a further distinction is made between two different components: the random intercept and the random slope. A random intercept is used to model differences in baseline values between patients, i.e. each patient has a different baseline value before treatment. However, a random slope is used to model differences in the effect of a predictor between patients. LMMs, in contrast to conventional linear regression models, always consist of fixed and random effects, which also explains the term 'mixed' model [63, 65, 79, 80].

Figure 6 schematically shows the different combinations of the components of random effects in LMMs compared to simple linear regression. In LMMs, as in the simple linear regression shown in panel A, a (random) intercept denotes the starting value, i.e. the outcome (y-value) when all independent variables (x-values) are equal to zero. However, the LMM allows a separate starting point for each patient, whereby the differences in the starting values are not determined in advance, but modelled as random effects and estimated from the data [80].
Equation 3.2 of the multivariable regression can be extended as follows if a random intercept

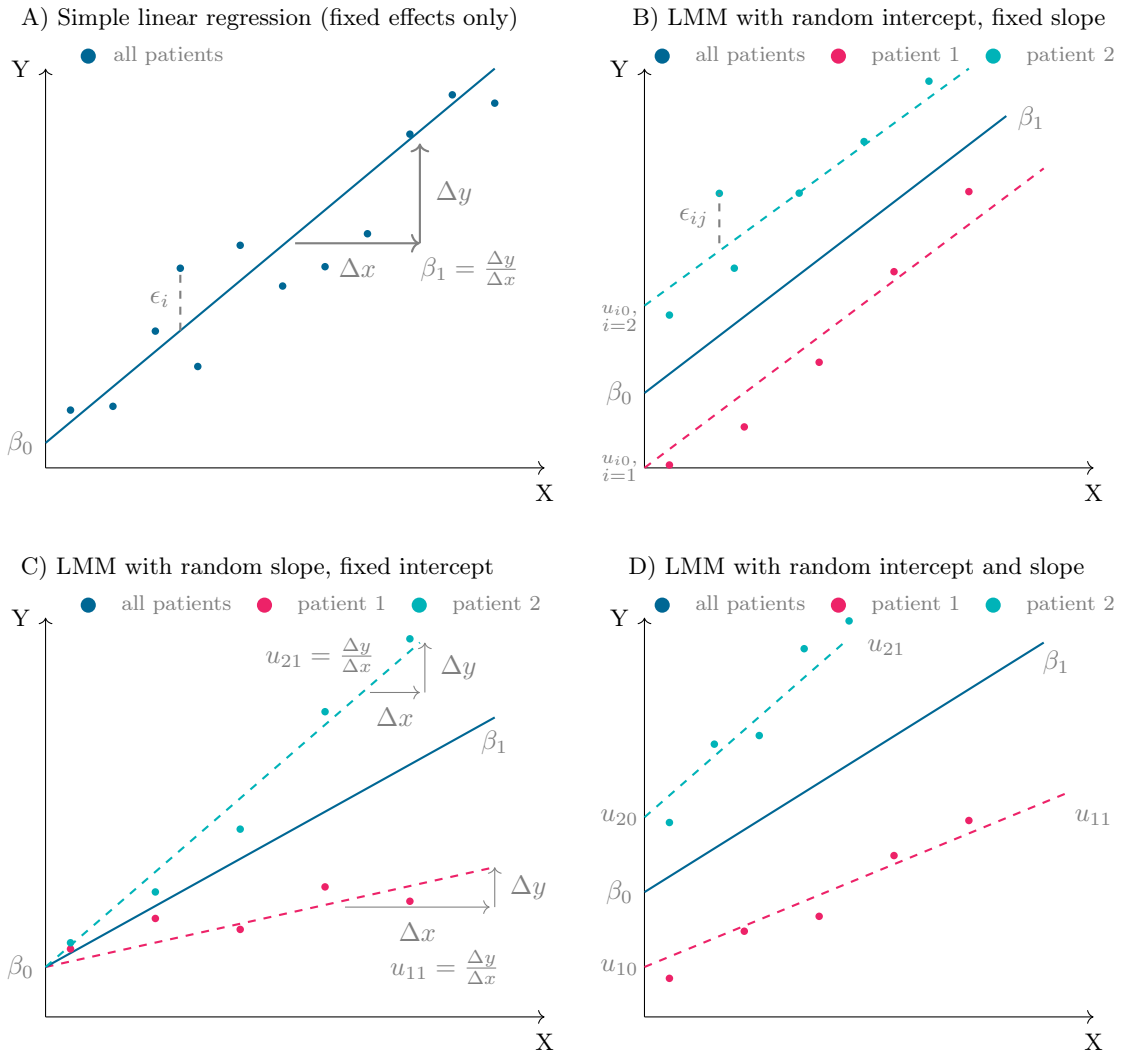


Figure 6: Fixed and random effects in linear mixed model (LMM) for repeated measures. ($\epsilon_i, \epsilon_{ij}$: error terms, residuals; β_0, β_1: fixed effects; u_{i0}, u_{i1}: random effects with u_{i0} as random intercept and u_{1i} as random slope)

(additional model level) is added (see also figure 6 panel B), resulting in equation 3.3, the common functional form for an LMM for repeated measures:

$$y_{ij} = \beta_0 + \beta_1 x_{ij1} + \cdots + \beta_K x_{ijK} + u_{i0} + \epsilon_{ij} \tag{3.3}$$

The dependent variable y_{ij} now corresponds to the value of patient i for $i = 1, ..., n$ (level 2) at time j for $j = 1, ..., J$ (level 1), and x_{ijk} correspond to the k independent predictors for $k = 1, ..., K$. β_k are the corresponding regression coefficients (fixed effects) of these predictors. As with the simple linear or multivariable linear regression, the model also contains the assumption of a normally distributed residual error term $\epsilon_{ij} \sim N(0, \sigma^2)$ with an expected value of 0 and constant variance σ^2, which now, however, indicates the measurement error of a patient i at the respective time j. This extension captures the within-person variance over time (**intra**individual variance) (see figure 6 panel B).

The global intercept (average initial value across all individuals) β_0 is also retained in the

LMM. However, in addition there is an individual deviation u_{i0} from the average intercept (random effect). This random intercept models the different initial levels, the specific distance (shift), of the starting value of a patient i from the global intercept β_0 (see figure 6, panel B), i.e. it takes into account the difference of the repeated measures between persons (**inter**individual variance). In this way, the misleading assumption of independent observations is avoided.
Like the residual error, the random intercept is also subject to the assumption of a normal distribution $u_{i0} \sim N(0, \tau_{00}^2)$, where τ_{00} describes the standard deviation of the random intercept between a respective patient i to the global intercept β_0. For all patients, this results in $u_{i0} \sim N(0, \tau_{00}^2 \cdot I_n)$, where I_n denotes the identity matrix to define the number of random intercepts according to the number of units in level 2 (patients, see figure 5). Thus, the variance for all patients results in a diagonal matrix and the greater τ_{00} is, the more the individuals differ in their average level of the dependent variable. In simple terms, this means that the model separates the total variance into the within-individual variance (σ^2, variations in a patient´s outcome over time) and between-individual variance (τ_{00}, variation of individuals' intercepts around the global intercept β_0) [64, 65, 74, 80].

In general, it is also common to represent the equation 3.3 in matrix notation, as this facilitates the mapping of the additional level 1 and makes the mathematical notation clearer.

$$\underbrace{\begin{bmatrix} y_{i1} \\ y_{i2} \\ \vdots \\ y_{iJ} \end{bmatrix}}_{\mathbf{y}_i} = \underbrace{\begin{bmatrix} 1 & x_{i11} & \dots & x_{i1K} \\ 1 & x_{i21} & \dots & x_{i2K} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & x_{iJ1} & \dots & x_{iJK} \end{bmatrix}}_{\mathbf{X}_i} \underbrace{\begin{bmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_K \end{bmatrix}}_{\boldsymbol{\beta}} + \underbrace{\begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix}}_{\mathbf{Z}_i} \underbrace{\begin{bmatrix} u_{i0} \end{bmatrix}}_{\mathbf{u}_i} + \underbrace{\begin{bmatrix} \epsilon_{i1} \\ \epsilon_{i2} \\ \vdots \\ \epsilon_{iJ} \end{bmatrix}}_{\boldsymbol{\epsilon}_i} = \underbrace{\mathbf{X}_i\boldsymbol{\beta}}_{\text{fixed}} + \underbrace{\mathbf{Z}_i^{[1]}\mathbf{u}_i}_{\substack{\text{random} \\ \text{intercept}}} + \boldsymbol{\epsilon}_i \qquad (3.4)$$

The matrix notation of the random intercept model (Equation 3.4), in particular the now visible design matrix $Z_i^{[1]}$, which links the random effects to the observations at level 1 and thus determines their influence on the modelling of the data structure, highlights that the variance now takes dependencies into account and shows that the variance-covariance structure considers the variability between subjects, which is not the case with simple linear or multiple linear regression models. For this reason, it is also important to examine the distribution of the residuals and random effects more closely [64, 74, 80, 81].

Usually, the distribution of the residuals is specified as $\epsilon_i \sim N(0, R_i)$, where R_i denotes the residual variance-covariance matrix. The R matrix describes the inaccuracies (variance) in the measurements and thus indicates how much individually measured values fluctuate around the true value. Put simply, it provides information about the extent of measurement inaccuracy and whether there are correlations between errors at different measurement times, thereby covering everything that cannot be explained by fixed or random effects.
In the case of the random intercept model and unstructured covariance, there are no specifications in the model, and each measurement can differ from every other by any amount, i.e.,

the residual matrix is estimated from the data, which provides maximum flexibility and does not lead to incorrect model assumptions about correlations. This then leads to the following symmetric matrix R for a patient i with J repeated measurement times:

$$\mathbf{R}_i = \begin{bmatrix} \sigma_1^2 & \sigma_{12} & \cdots & \sigma_{1J} \\ \sigma_{12} & \sigma_2^2 & \cdots & \sigma_{2J} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{1J} & \sigma_{2J} & \cdots & \sigma_J^2 \end{bmatrix}$$

Here σ_j^2 denotes the residual variance (error variance) at time point j and $\sigma_{jj'}$ denotes the covariance between time points j and j' with $j \neq j'$ [64, 74].
In addition to the distribution of the residuals, the distribution of the random effects u_i is also of particular importance and is usually specified as $u_i \sim N(0, G_i)$, where G_i denotes the variance-covariance matrix of the random effects, which reflects the differences between patients at the higher level (level 2, see Figure 5).
In the case of the random intercept model, as already mentioned, the G matrix becomes a diagonal matrix ($u_0 \sim N(0, G)$, with $G = \tau_{00}^2 \cdot I_n$) and explains the variance in the intercept of the individual patients [64, 74].

Besides the random intercept model, as mentioned above, differences in slopes between patients can also be modelled. For this purpose, a random slope model without random intercept, as shown schematically in Figure 6 in panel C and with the following notation, is used.

$$\underbrace{\begin{bmatrix} y_{i1} \\ y_{i2} \\ \vdots \\ y_{iJ} \end{bmatrix}}_{\mathbf{y}_i} = \underbrace{\begin{bmatrix} 1 & x_{i11} & \ldots & x_{i1K} \\ 1 & x_{i21} & \ldots & x_{i2K} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & x_{iJ1} & \ldots & x_{iJK} \end{bmatrix}}_{\mathbf{X}_i} \underbrace{\begin{bmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_K \end{bmatrix}}_{\boldsymbol{\beta}} + \underbrace{\begin{bmatrix} x_{i11} \\ x_{i21} \\ \vdots \\ x_{iJ1} \end{bmatrix}}_{\mathbf{Z}_i} \underbrace{\begin{bmatrix} u_{1i} \end{bmatrix}}_{\mathbf{u}_i} + \underbrace{\begin{bmatrix} \epsilon_{i1} \\ \epsilon_{i2} \\ \vdots \\ \epsilon_{iJ} \end{bmatrix}}_{\boldsymbol{\epsilon}_i} = \underbrace{\mathbf{X}_i\boldsymbol{\beta}}_{\text{fixed}} + \underbrace{\mathbf{Z}_i^{[x]}\mathbf{u}_i}_{\substack{\text{random} \\ \text{slope} \\ \text{for } x_{iJ1}}} + \boldsymbol{\epsilon}_i \quad (3.5)$$

The matrix notation in equation 3.5 of the random slope model differs from the random intercept model only in the design matrix $Z_i^{[x]}$ of the random effects. In the random intercept model, the heterogeneity of the initial level of the outcome variable between patients is modelled by selecting $Z_i^{[1]} = (1, 1, \ldots, 1)^T$. In the random slope model, on the other hand, the influence of a covariate, here $x_{i=1}$, on the outcome y varies between patients due to the selection of $Z_i^{[x]} = (x_{i11}, x_{i21}, \ldots, x_{iJ1})^T$. Thus, Z_i acts on a specific column of the design matrix with the independent predictors X_i and enables patient-specific slopes for this predictor. This is also reflected in the G matrix, which, as in the random intercept model, is a multivariate normally distributed diagonal matrix ($u_1 \sim N(0, G)$, with $G = \tau_{11}^2 \cdot I_n$), whereby in this case the variance of the slopes between patients is modelled and the intercept is constant for all patients (see figure 6, panel C) [64, 74].

Furthermore, it is of course also possible, as in equation 3.6 and schematically shown in figure 6 panel D, to take both random effects, i.e. a random intercept and a random slope, into account in a LMM.

$$\underbrace{\begin{bmatrix} y_{i1} \\ y_{i2} \\ \vdots \\ y_{iJ} \end{bmatrix}}_{\mathbf{y}_i} = \underbrace{\begin{bmatrix} 1 & x_{i11} & \dots & x_{i1K} \\ 1 & x_{i21} & \dots & x_{i2K} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & x_{iJ1} & \dots & x_{iJK} \end{bmatrix}}_{\mathbf{X}_i} \underbrace{\begin{bmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_K \end{bmatrix}}_{\boldsymbol{\beta}} + \underbrace{\begin{bmatrix} 1 & x_{i11} \\ 1 & x_{i21} \\ \vdots & \vdots \\ 1 & x_{iJ1} \end{bmatrix}}_{\mathbf{Z}_i} \underbrace{\begin{bmatrix} u_{0i} \\ u_{1i} \end{bmatrix}}_{\mathbf{u}_i} + \underbrace{\begin{bmatrix} \epsilon_{i1} \\ \epsilon_{i2} \\ \vdots \\ \epsilon_{iJ} \end{bmatrix}}_{\boldsymbol{\epsilon}_i} = \underbrace{\mathbf{X}_i\boldsymbol{\beta}}_{\text{fixed}} + \underbrace{\mathbf{Z}_i^{[1,x]}\mathbf{u}_i}_{\substack{\text{random} \\ \text{intercept} \\ + \text{ slope} \\ \text{for } x_{iJ1}}} + \boldsymbol{\epsilon}_i \tag{3.6}$$

If both the random intercept and the random slope are modelled, the G matrix is expanded accordingly to include the covariance between the intercept and the slope, with

$$\mathbf{u}_i = \begin{bmatrix} u_{i0} \\ u_{i1} \end{bmatrix} \sim \mathcal{N}\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \underbrace{\begin{bmatrix} \tau_{00}^2 & \tau_{01} \\ \tau_{01} & \tau_{11}^2 \end{bmatrix}}_{\mathbf{G_i}} \right).$$

The covariance then indicates to which extent the two random effects change together, i.e. how they correlate with each other [64, 74].

For a quick overview, table 3.1 summarises the most important differences between the individual models in a clear and concise manner.

Table 3.1: Overview of the differences between simple linear, multivariable linear and linear mixed model regression

Model	Basic idea	Mathematical notation	Data structure
simple linear regression	One dependent variable is explained by one predictor.	$y_i = \underbrace{\beta_0 + \beta_1 x_i}_{\text{fixed}} + \epsilon_i$	non-hierarchical
multiple linear regression	One dependent variable is explained by several predictors.	$y_i = \underbrace{\beta_0 + \beta_1 x_{i1} + \dots + \beta_K x_{iK}}_{\text{fixed}} + \epsilon_i$	non-hierarchical
LMM – random intercept	Intercept varies between individuals, same slope for all individuals	$y_i = \underbrace{\mathbf{X}_i\boldsymbol{\beta}}_{\text{fixed}} + \underbrace{\mathbf{Z}_i^{[1]}\mathbf{u}_i}_{\substack{\text{random} \\ \text{intercept}}} + \boldsymbol{\varepsilon}_i$	hierarchical / longitudinal
LMM – random slope	Slope varies between individuals, same intercept.	$y_i = \underbrace{\mathbf{X}_i\boldsymbol{\beta}}_{\text{fixed}} + \underbrace{\mathbf{Z}_i^{[x]}\mathbf{u}_i}_{\substack{\text{random} \\ \text{slope} \\ \text{for } x_{iJ1}}} + \boldsymbol{\varepsilon}_i$	hierarchical / longitudinal
LMM – random intercept & slope	Intercept and slope vary between individuals	$y_i = \underbrace{\mathbf{X}_i\boldsymbol{\beta}}_{\text{fixed}} + \underbrace{\mathbf{Z}_i^{[1,x]}\mathbf{u}_i}_{\substack{\text{random} \\ \text{intercept} \\ + \text{ slope} \\ \text{for } x_{iJ1}}} + \boldsymbol{\varepsilon}_i$	hierarchical / longitudinal

For a better understanding, all combinations were briefly explained above, whereby the notation of the random intercept model (see also figure 6 panel B) is particularly relevant for

the methodological extension of the RDD for repeated measurements in the third article.

4 Applications and Results

In addition to the first publication described in the previous methods chapter, which deals with the conceptual and theoretical foundations of the work, and the detailed methodological explanation of linear mixed models for repeated measures, two further publications will be presented in this chapter. These two articles refer to the application and further development of the methodology described above and therefore contain the empirical results of the dissertation. The first publication focuses on the application of the RDD in the familiar sense within the framework of a subgroup analysis of the nFC-isPO data. The second publication presents the methodological development of the RDD for the use of longitudinal data with repeated measurements, also using the example of the nFC-isPO data. Although both articles are based on the same data set, they differ significantly in their objectives. One focuses primarily on the application of the basic methodology to demonstrate its usefulness in the field of health sciences and make it more attractive to scientists in this field, while the other forms the main part of the dissertation and thus promotes the applicability of the methodology in studies with longitudinal data and is therefore strongly methodologically oriented and focused.

4.1 Application of RDD in the nFC-isPO - 2. Article

Published as:

The impact of COVID-19 on the interpretation of psycho-oncological support trial results: a quasi-experimental approach using the data from the new form of care "Integrated crosssectoral psycho-oncology (nFC-isPO)"

Anna Hagemeier, Anne Adams, Theresia Krieger, Sandra Salm, Natlia Cecon-Stabel, Antje Dresen, Martin Hellmich

BMC Health Serv Res, 2023 May 30, 23(1):556,
https://doi.org/10.1186/s12913-023-09544-y

Visual Abstract:

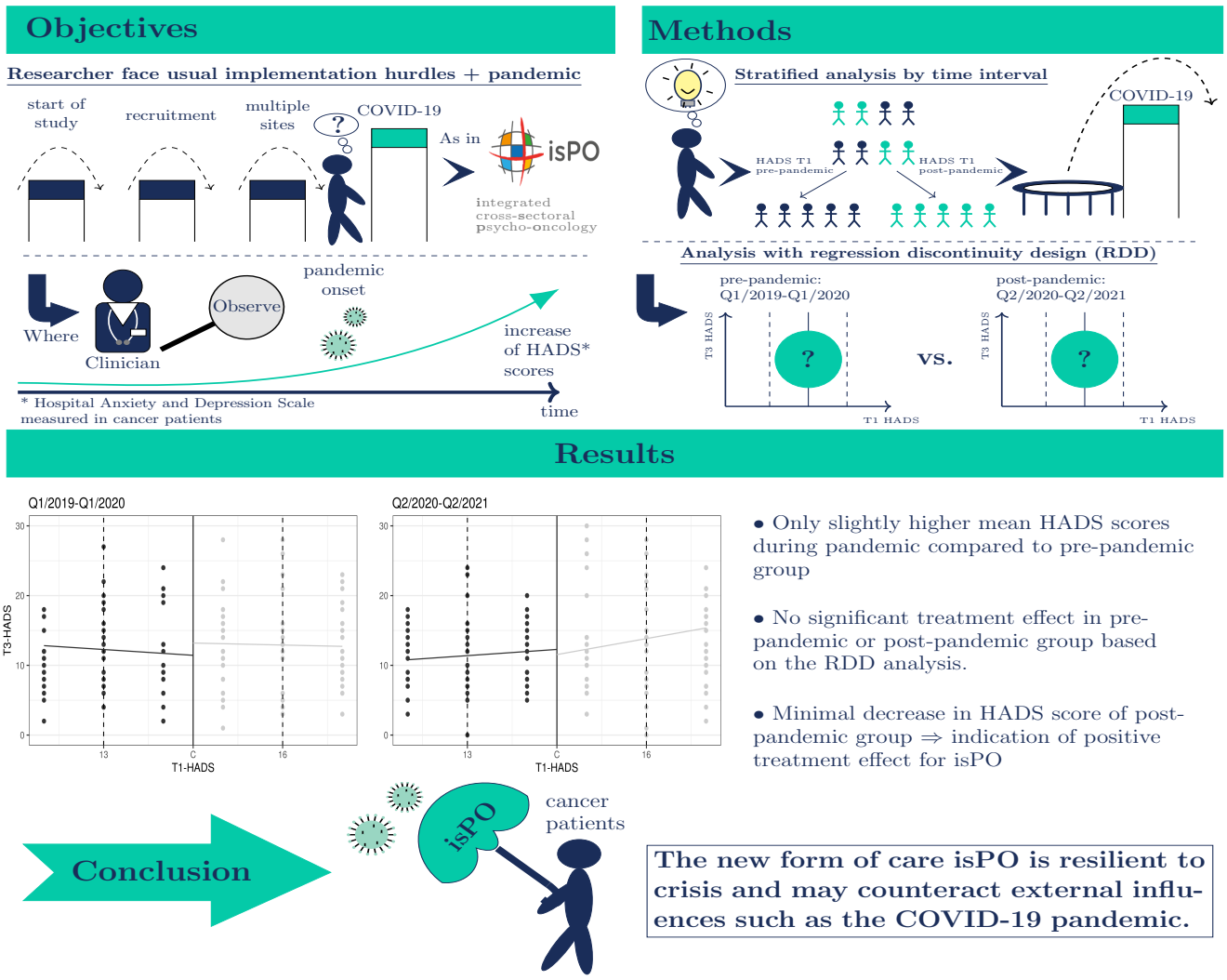


Figure 7: Visual abstract: "The impact of COVID-19 on the interpretation of psycho-oncological support trial results: a quasi-experimental approach using the data from the new form of care "Integrated cross-sectoral psycho-oncology (nFC-isPO)""

Hagemeier et al. BMC Health Services Research (2023) 23:556
https://doi.org/10.1186/s12913-023-09544-y

BMC Health Services Research

RESEARCH Open Access

The impact of COVID-19 on the interpretation of psycho-oncological support trial results: a quasi-experimental approach using the data from the new form of care "Integrated cross-sectoral psycho-oncology (nFC-isPO)"

Anna Hagemeier[1*], Anne Adams[1], Theresia Krieger[2], Sandra Salm[2,3], Natalia Cecon-Stabel[2], Antje Dresen[2] and Martin Hellmich[1]

Abstract

Objective In addition to the common difficulties of ongoing trials, the COVID-19 pandemic posed several challenges to scientists worldwide and created an additional burden for vulnerable patient groups. In the nFC-isPO of individualised treatment for anxiety and depression in newly diagnosed patients with cancer caregivers (e.g. psycho-oncologists) reported elevated HADS scores in newly enrolled patients after the outbreak of the COVID-19 pandemic. Accordingly, the question arises whether the pandemic affected HADS scores. Therefore, stratified analyses by the time of enrolment (T1) were performed for patients with 12 months of care (T3).

Methods Patients with 12 months of care (N = 1,140) were analysed. A comparison within the regression discontinuity design according to the time points at which patients completed the baseline (T1) HADS questionnaire was conducted to examine differences between patients recruited before Q2/2020 (pre-pandemic) and after the coronavirus outbreak. Furthermore, mean HADS scores at T1 and T3 for all quarters during the study were compared.

Results Mean T1 and T3 HADS scores of patients with cancer during the pandemic are only slightly higher than those of the pre-pandemic group. No significant treatment effect was observed in either the pre-pandemic (p = 0.5495, Late = 1.7711) or the post-pandemic group (p = 0.9098, LATE=-0.2933). In contrast, the average local treatment effect in the post-pandemic group suggests a minimal decrease in HADS score in the predefined range and thus a positive treatment effect for isPO. Comparison of mean HADS scores at T1 and T3 did not show a large increase by pandemic-related timepoints, however, a decrease of approximately 2–3 points over each quarter at 12 months compared to baseline is observed.

*Correspondence:
Anna Hagemeier
anna.hagemeier@uni-koeln.de

Full list of author information is available at the end of the article

Conclusion The existing nFC-isPO care is resilient to crisis and may counteract external influences such as the Corona pandemic. Accordingly, the pandemic had little influence on the fears of patients with cancer in the nFC-isPO. This emphasises that psycho-oncology is vital for the reduction of stress, anxiety and depression in patients with cancer.

Trial Registration The study was registered in the German Clinical Trials Registry on 30 October 2018 under the ID "DRKS00015326".

Keywords Anxiety, cancer, COVID-19, Depression, HADS, Mental health, Psycho-oncology, Pandemic, Regression discontinuity design, Stratification

Introduction

COVID-19 has taken a toll on people and health systems around the world, and many studies have examined this novel disease and its effects [1]. In addition to the general impact and the impact on the health system, COVID-19 also affected ongoing trials, e.g. in form of changes in the shape of care supply. These challenges were thus in addition to the usual challenges such as slow recruitment or managing multi-site. However, these problems with trials did not only affect providers, but especially vulnerable patient groups who had a higher risk of infection due to pre-existing conditions and a correspondingly weakened immune system [2–4]. One of these vulnerable groups was patients with cancer, especially those who received their first diagnosis during the pandemic. A cancer diagnosis is already a stressful and drastic event and often leads to stress, depression and especially anxiety. When patients receive such a diagnosis in an exceptional situation such as the COVID-19 pandemic, psychological stress is compounded in many cases [5, 6]. For this reason, holistic psycho-oncological treatment, i.e. in addition to treating the cause of the cancer, the inclusion of mental health and its treatment is an important component [7].

The project and the new form of care (nFC) "Integrated cross-sectoral psycho-oncology" (isPO) started in 2017 with precisely this intention, namely holistic psycho-oncological care and above all with the aim of creating structured needs-oriented psycho-oncological care in Germany, and was then challenged by the consequences of the pandemic during the project period [8, 9]. isPO was conducted at four sites in North-Rhein Westphalia using a quasi-experimental study design, the regression discontinuity design (RDD) in combination with the Hospital Anxiety and Depression Scale (HADS) to assess individual psycho-oncological needs [10]. The new form of care aims to successfully reduce anxiety and depression within the first year after cancer diagnosis in patients with cancer through psychological, psychotherapeutical and psychosocial measures. Given the study design and psychological treatment, the aim is to determine whether HADS scores decrease in the intervention group. During the study, after the COVID-19 pandemic onset (March 2020) [11], the four isPO care networks provided feedback that higher HADS values were observed in many patients, and the potential for anxiety and depression increased [12]. Therefore, the question arose as to how the influence of the pandemic can be taken into account in the evaluation of the study. This should be considered with regard to effectiveness in terms of an increase in HADS scores, but also with regard to necessary adjustments in care, such as switching from face-to-face care to care via telephone and video telephony. The aim is to discuss whether the higher HADS values observed by caregivers are reflected in the study results by performing stratified analyses (descriptive and using RDD) according to the time of enrolment in the study programme (T1).

Materials and methods

General study information and study design

In isPO, newly diagnosed patients with cancer were offered psycho-oncological care at different levels of care according to their individual needs which are assessed using the HADS. Patients were requested to complete the HADS questionnaire at three different time points (Baseline (T1), after 4 months of care (T2) and after 12 months of care (T3)). T1-HADS scores are used for assignment to the different levels of care and thus for classification into control and intervention groups. This is done as a quasi-randomization using the RDD, which uses a threshold criterion to divide patients into control and treatment groups and measure treatment effect by comparing the intercepts of regression equations above and below a predefined threshold and within a bandwidth [10, 13]. In isPO patients with baseline HADS values ≤ 14 were allocated to the control group (psychosocial intervention). Patients with higher values (≥ 15) received psycho-oncological, psychotherapeutical treatment (intervention group). T2 and T3 (primary endpoint) are used to compare the allocated patients around the threshold and measure their psychological needs after receiving the individual treatment. Therefore, patients with less distress at the first time point were compared to those with higher distress. This procedure aims to record and review the effectiveness of psycho-oncological psychotherapeutical care [8, 9].

Data collection and analysis population

N=1,757 newly diagnosed patients with cancer were recruited in the period 01/2019-03/2021. Data were collected at four isPO care networks (Cologne, Troisdorf, Mönchengladbach and Neuss) in North Rhine-Westphalia (NRW) and recorded in a computer-assisted assistance system (CAPSYS^{2020}) developed for the isPO project. For proper analysis of stratified data and primary end point at T3, only cases followed for 12 months were considered (N=1,140). The main analysis (RDD) compares the HADS scores of N=203 patients within the two strata (T1 pre-pandemic (N=96) vs. T1 post-pandemic (N=107)) within the pre defined bandwidth of 13–16 points at baseline. All patients provided written informed consent.

A detailed overview of the study and analysis population is shown in Fig. 1.

Instruments

The main instrument of the isPO study was the Hospital Anxiety and Depression Scale (HADS). The HADS consists of 14 items, 7 of which are assigned to the depression subscale and 7 to the anxiety subscale with a score from 0 to 3 for each response. The added values of the individual items result in the subscales, whereby values between 0 and 7 are to be interpreted as inconspicuous, values from 8 to 10 as suspicious and values>10 as conspicuous. The total score ranges from 0 to 42. Based on the interpretation of the subscales, 14.5 was chosen as the threshold for group assignment for the regression discontinuity design within the nFC-isPO [14, 15].

Statistics

Simple stratification by time intervals was used to account for the pandemic influence. Therefore, an additional categorical variable in the database was built. The variable considers at what point in the study a patient completed the HADS questionnaire at timepoint T1 (baseline). This variable was then used to stratify patients into a pre- and post-pandemic groups (T1 in Q1/19-Q1/20 vs. T1 in Q2/20-Q2/21), to identify differences between patients´ anxiety and depression levels over time and to explore possible differences during the chronological trend.

First, descriptive analyses were conducted to provide an overview of the study population. Therefore, categorical variables are reported as absolute and relative frequencies in percentage. For continuous variables mean with standard deviation and median with interquartile range are given. Next, RDD analysis were conducted in the two strata to measure the effect of the psycho-oncological treatment. An RDD analysis is based on simple linear regression in which two regression equations are compared below and above a threshold [10]. The difference between the two regression lines or the difference between the two axis intercepts at the pre defined threshold then gives the treatment effect. This is also called the local average treatment effect (LATE) and is graphically

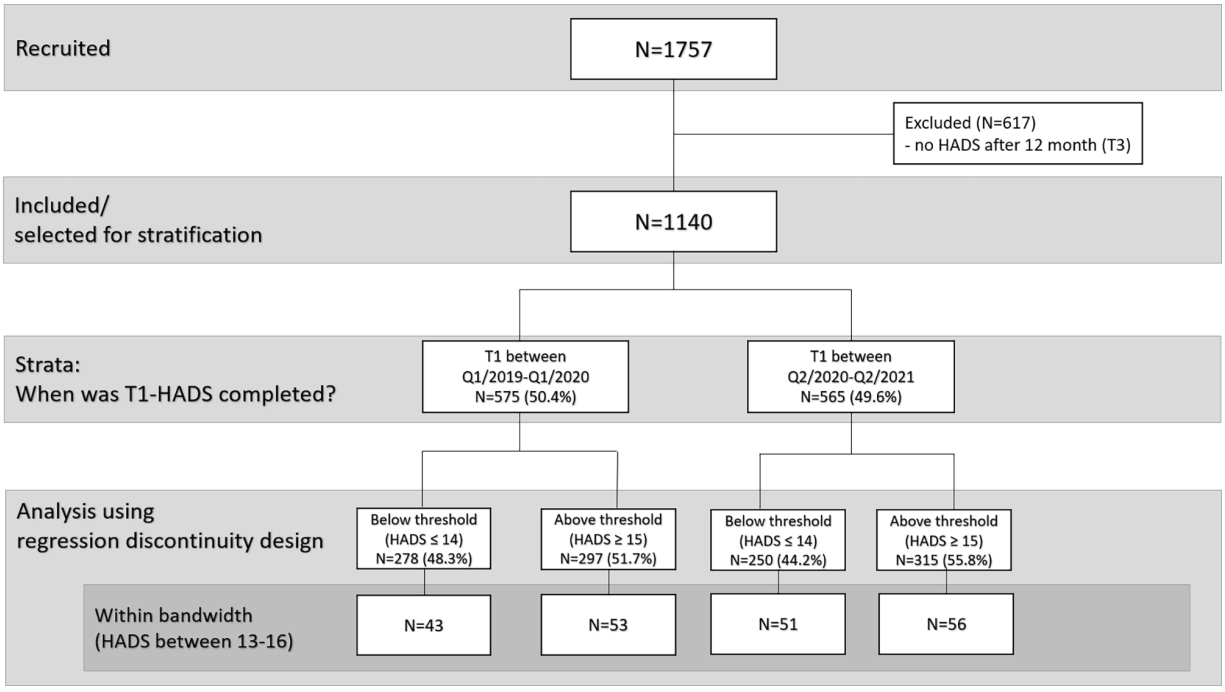


Fig. 1 Flowchart of study and analysis population

represented as a discontinuity (jump) of the regression lines at the threshold [10, 13].

In addition, mean baseline HADS scores were compared with HADS scores after 12 months of care to show changes over time. All analyses were performed using IBM SPSS Version 28 with available R extensions, e.g. regression discontinuity analysis and figures were generated with R Version 4.1.2. [16] Results were considered statistically significant for $p \leq 0.05$ with $\alpha = 0.05$ as the significance level.

Table 1 Sociodemographic and clinical patient characteristics by T1-time intervals

	Q1/19-Q1/20 (N = 575)	Q2/20-Q2/21 (N = 565)	Total (N = 1,140)
Gender			
Female	378 (65.7%)	362 (64.1%)	740 (64.9%)
Male	197 (34.3%)	203 (35.9%)	400 (35.1%)
Age			
Mean ± SD	56.7 ± 12.9	57.0 ± 13.3	56.8 ± 13.1
Median[Q1;Q3]	58 [50; 66]	59 [49;65]	58 [50;66]
Care level			
1	54 (9.4%)	50 (8.8%)	104 (9.1%)
2	224 (39.0%)	201 (35.6%)	425 (37.3%)
3a	94 (16.3%)	101 (17.9%)	195 (17.1%)
3b	203 (35.3%)	213 (37.7%)	416 (36.5%)
Relationship status			
No relationship	117 (21.2%)	134 (24.7%)	251 (22.9%)
In relationship, shared household	390 (70.7%)	377 (69.4%)	767 (70.0%)
In relationship, separate households	45 (8.2%)	32 (5.9%)	77 (7.0%)
ISCED			
Primary education	9 (1.6%)	9 (1.6%)	18 (1.6%)
Lower secondary education	38 (6.8%)	28 (5.1%)	66 (6.0%)
Upper secondary education	389 (69.7%)	385 (70.3%)	774 (70.0%)
Bachelor or equivalent	33 (5.9%)	37 (6.8%)	70 (6.3%)
Master or equivalent	80 (14.3%)	84 (15.3%)	164 (14.8%)
Doctoral or equivalent	9 (1.6%)	5 (0.9%)	14 (1.3%)
Servere disability			
No	393 (69.8%)	386 (69.7%)	779 (69.7%)
Yes	170 (30.2%)	168 (30.3%)	338 (30.3%)
Chronic illness			
No	328 (58.3%)	327 (59.0%)	655 (58.6%)
Yes	235 (41.7%)	227 (41.7%)	462 (41.4%)
ICD10-entities			
C15-C26	74 (13.1%)	75 (13.6%)	149 (13.3%)
C30-C39	49 (8.7%)	49 (8.9%)	98 (8.8%)
C43-C44	41 (7.2%)	29 (5.3%)	70 (6.3%)
C50	177 (31.3%)	158 (28.6%)	335 (30.0%)
C51-C58	26 (4.6%)	42 (7.6%)	68 (6.1%)
C81-C96	65 (11.5%)	54 (9.8%)	119 (10.6%)
Other (each < 5%)	134 (%)	145 (%)	279 (24.5%)
Death			
Yes	11 (1.9%)	1 (0.2%)	12 (1.1%)
No	564 (98.1%)	564 (99.8%)	1,128 (98.9%)

Results

Characteristics of the patients

A total of 1,140 patients completed the HADS questionnaire after 12 months. Of these, 575 completed the HADS for the first time between Q1/19-Q1/20, before the pandemic, and 565 between Q2/20-Q2/21, after the pandemic outbreak. The gender ratio was about two-thirds women to one-third men with a median age of about 58 years (IQR=[50–66], mean=56.8, SD=13.1). The most common diagnoses were breast cancer (30%) and malignant neoplasms of the digestive organs (13.3%). An overview of all sociodemographic and clinical patient characteristics is given in Table 1.

Modification of HADS

The mean and median baseline (T1) HADS scores of the anxiety subscale are minimally higher in the group during the pandemic (9.0±4.4 and 9 [6;12]) than in the group before the pandemic (8.8±4.6 and 9 [5;12]). The same applies to the depression subscale (7.1±4.6 and 6 [4;10] vs. 6.5±4.5 and 6 [3;9]). However, for both groups (before and during the pandemic), it is visible that the T3 HADS decreases to a similar extent, by about 2 points each for the anxiety subscale and by about 1 point each for the depression subscale (see Table 2).

Primary outcome pre- vs. post-pandemic

Scatterplots with all cases of the two groups according to time within the pandemic are shown (Fig. 2). In both groups, i.e. in the control group below the threshold and in the treatment group above the threshold, the higher the T1-HADS score, the higher the T3-HADS score. The second row of Fig. 2 then shows the applied RDD analysis in the previously defined relevant section (bandwidth) of the entire collective. The area of interest was the bandwidth from 13 to 16 points on the HADS scale. The pre-pandemic group consisted in total of N=575 patients (50.4%), of which N=96 patients lie within the pre defined bandwidth. The post-pandemic group had a total N of 565 patients (49.6%) and N=107 patients within the bandwidth. Both groups show a discontinuity of the regression lines in the area around the threshold. However, the local treatment effect (LATE) measured in each case is not significant. In addition, only the local treatment effect of the post-pandemic group points in the right direction (LATE=-0.2933), as it is negative and thus indicates an improvement (reduction) in the HADS values.

Table 2 Summary of the subscales and total HADS scores at baseline (T1) and at the primary endpoint after 12 months of care (T3) for the pre- and post-pandemic groups

	Q1/19-Q1/20 (N = 575)	Q2/20-Q2/21 (N = 565)	Total (N = 1,140)
HADS anxiety T1			
Mean ± SD	8.8 ± 4.6	9.0 ± 4.4	8.9 ± 4.5
Median [Q1;Q3]	9 [5;12]	9 [6;12]	9 [6;12]
HADS anxiety T3			
Mean ± SD	7.2 ± 4.3	7.6 ± 4.3	7.4 ± 4.3
Median [Q1;Q3]	7 [4;10]	7 [4;11]	7 [4;10]
HADS depression T1			
Mean ± SD	6.5 ± 4.5	7.1 ± 4.6	6.8 ± 4.6
Median [Q1;Q3]	6 [3;9]	6 [4;10]	6 [3;10]
HADS depression T3			
Mean ± SD	5.5 ± 4.4	6.2 ± 4.8	5.8 ± 4.6
Median [Q1;Q3]	5 [2;8]	5 [2;10]	5 [2;9]
HADS total T1			
Mean ± SD	15.2 ± 8.5	16.0 ± 8.4	15.6 ± 8.4
Median [Q1;Q3]	15 [8;21]	16 [10;22]	15 [9;21]
HADS total T3			
Mean ± SD	12.7 ± 8.2	13.8 ± 8.5	13.2 ± 8.4
Median [Q1;Q3]	12 [6;18]	13 [7;20]	12 [6;19]
Group by HADS-Cut-off			
HADS < = 14	278 (48.3%)	250 (44.2%)	528 (46.3%)
HADS > = 15	297 (51.7%)	315 (55.8%)	612 (53.7%)

* HADS: Hospital anxiety and depression scale

Mean HADS over time

In addition to the primary outcome, mean baseline and HADS scores after 12 months of treatment are compared in Fig. 3. Across all quarters, the overall mean HADS baseline scores are found to be in a very similar range between 15 and 16 points. 12 months after treatment, the HADS scores are also in a very similar range between 12 and 14 points. A direct comparison of the mean HADS scores at baseline and 12 months after treatment shows a reduction of about 2–3 points.

Comparing the control and intervention groups, the mean pre-treatment HADS scores in the intervention group at baseline (T1) are around 22 points across all quarters, well above the chosen threshold. The mean HADS scores of the control group, on the other hand, were around 8 to 9 points and thus, as with the intervention group, not in the range of 13–16 points relevant for RDD analysis. Comparing the mean HADS scores over time after 12 months of treatment between the two groups (intervention vs. control), it is visible that in the intervention group there was a reduction of the scores from over 20 points at the beginning of treatment to about 15 to 17 points after 12 months. In the control group, on the other hand, a slight deterioration, i.e. an increase in the scores, can be seen.

Discussion

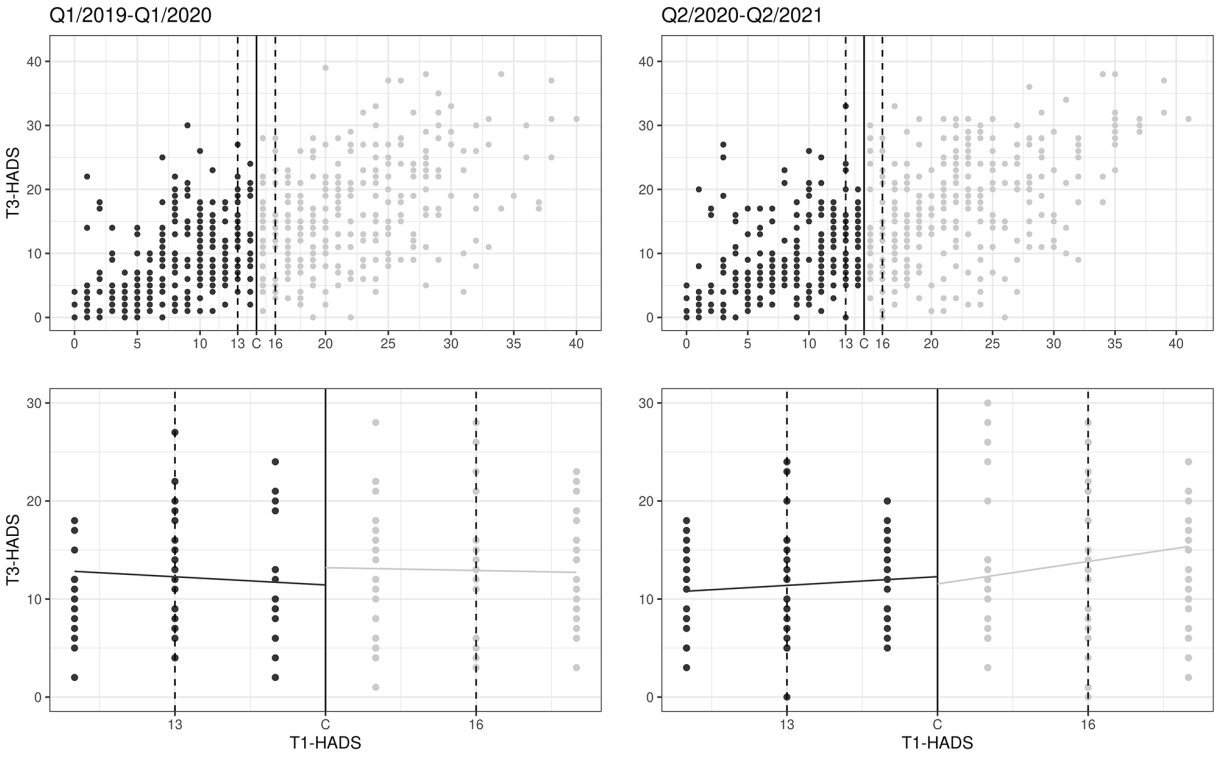
The nFC-isPO aimed to reduce anxiety and depression in patients with cancer as measured by HADS. However, due to the pandemic, it was suspected that patients would be exposed to additional psychological distress on top of the already high psychological distress caused by cancer, which could prejudice the treatment effect. Therefore, stratified analyses for primary outcome- (T3, 12 months of care) were performed in this article to examine the effects of COVID-19 in the isPO trial.

In contrast to the initial feedback from the four isPO care networks, the results of our descriptive analyses over time showed that the HADS scores in the post-pandemic group were only slightly higher on average or median overall than in the pre-pandemic group. This is also shown in other studies, such as Bargon et al.[17] being one of the few studies that examines the association between mental health in relation to cancer and the impact of the pandemic.

Many other studies investigated COVID-19 and the resulting deterioration in mental health [18, 19]. isPO, tried to reduce anxiety and depression in newly diagnosed patients with cancer with a stepped care approach. It was started before and carried on during the pandemic while facing additional challenges as a result of the pandemic. Even before the pandemic, supporting and counseling vulnerable groups, e.g. patients with cancer, was considered as very important [20]. In the face of additional sudden factors, like the pandemic, it is even more important to avoid additional stress[21, 22].

This is also reflected in the results of isPO, as the average anxiety and depression scores before the outbreak of the pandemic hardly changed compared to after. On the other hand, cancer studies without psycho-oncology stepped care approach, show significantly higher anxiety and depression scores due to an external factor such as COVID-19 [23]. It can be concluded that needs-oriented psycho-oncological support has a positive impact and reduces anxiety not only with regard to the cancer diagnosis, but also with regard to other medical factors/problems that have a negative impact on anxiety and depression, such as COVID-19 [24–26]. This implies that because the isPO-patients already received personalized psycho-oncological care during the pandemic, the HADS scores did not increase significantly compared to patients who received a cancer diagnosis during the pandemic without additional psycho-oncological care (no nFC-isPO) [27, 28].

Regarding the results of the RDD analysis, the comparison of the pandemic groups showed that the discontinuity, i.e. treatment effect, increased in the pre-pandemic intervention group instead of decreasing as expected. In comparison, the post-pandemic group showed the expected decrease in HADS scores in the intervention



N Total	N bandwidth	LATE	Std. Error	z value	p-value
575	96	1.7711	2.959	0.5985	0.5495

N Total	N bandwidth	LATE	Std. Error	z value	p-value
565	107	-0.2933	2.59	-0.1132	0.9098

Fig. 2 Scatter plot (four panels) of the primary outcome with the results of the regression discontinuity analyses (local average treatment effect - LATE) for the group before (left) and after (right) the pandemic for the patients' baseline HADS scores (T1) and HADS scores after 12 months of treatment (T3)
* Below Threshold (HADS ≤ 14): control (black); above threshold (HADS ≥ 15): intervention (grey)

group. This outcome might be caused by the fact that the patients who had completed the HADS questionnaire at T1 pre-pandemic almost all only completed the T3 survey during, many also right at the beginning, of the pandemic. That is, the higher HADS scores could be due to additional stress from the pandemic. The group during the pandemic, on the other hand, filled out all questionnaires after the start of the pandemic, so that the pandemic situation was still tense, but no longer completely new. Moreover, by the time the pandemic started, nFC-isPO was already more advanced and matured within the four sites. Therefore, working routines saved the ongoing of the intervention with high quality [29].

Limitations

A major limitation of the study was the high number of cases required for conducting RDD analysis. Originally, a case number of over 3,500 patients was planned in order to have sufficient cases for the RDD analyses (power calculation). After adjustments to the sample size due to poor enrollment, including the pandemic and the complexity of the nFC-isPO, only about half (N = 1,757) of the originally planned sample size could be enrolled. Although this was the minimum number of cases for the relevant RDD analyses, RDDs still benefits from the highest case numbers possible.

Another limitation for the proof of effectiveness was the patient group selected for ethical reasons. In the range between 13 and 16 points of the HADS, the treatment group showed a rather low need for psycho-oncological care. A more targeted examination of more severely stressed cases showed that the care offers were taken up and also showed corresponding effects (see summative study report [12]). Furthermore, the effectiveness of a psychotherapeutic intervention cannot be presented without the effect of time (in this case 12 months) and the effect of individual dose (number of therapeutical talks). Finally, the analyses conducted here cannot conclusively clarify whether the short-term reported increase in HADS scores by isPO-staff during pandemic reflects fear of additional illness and the impact on cancer treatment, or whether scores increased in the short term due to fears of poor accessibility to doctors or other cancer support services.

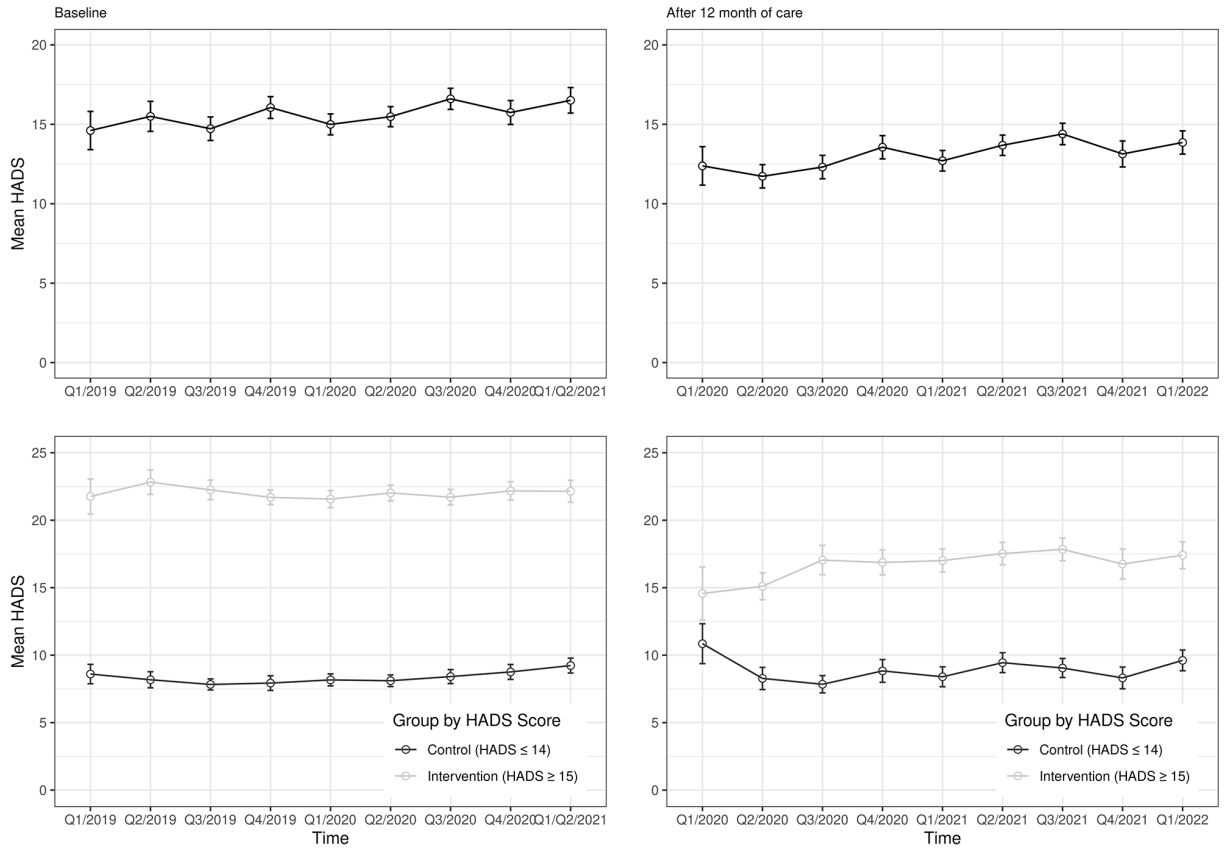


Fig. 3 Mean Baseline and primary outcome HADS values by time. The two upper panels show the unstratified mean HADS scores at baseline (left) and after 12 months (right) over the course of the project. Below this is the stratified representation according to intervention and control group

Conclusion

Patients with cancer are considered as vulnerable group who suffer from anxiety and stress very suddenly [30], regardless of other external factors such as the COVID-19 pandemic presented here. Our results underline how reliable and resilient the nFC-isPO is to external influencing factors in everyday hospital practice. On the one hand, it turned out that HADS scores were not permanently higher, although this was perceived by the isPO care networks at the beginning of the pandemic. For another, despite the pandemic, the nFC-isPO led to a decrease in HADS scores on average over time. Interestingly, this was despite the need to switch from face-to-face care to sometimes exclusively telephone or video-based care [31, 32]. Therefore, the results underline the importance of early and personalized psycho-oncological support in order to reduce stress, anxiety and depression in patients with cancer.

Abbreviations

CAPSYS^{2020}	Computer-assisted assistance system in isPO
ISCED	International Standard Classification of Education
IQR	Interquartile range
isPO	Integrated, cross-sectoral psycho-oncology
HADS	Hospital Anxiety and Depression Scale
LATE	Local average treatment effect
nFC	New form of care
NRW	North Rhine-Westphalia
RDD	Regression discontinuity design
SD	Standard deviation

Acknowledgements

We thank all participating patients and all the clinical staff involved for their cooperation, as well as the cooperating health insurance companies for their support in the isPO project. We also thank all other consortium partners in the isPO project. We acknowledge support for the Article Processing Charge from the DFG (German Research Foundation, 491454339)

Authors' contributions

Conceptualization A.H. and M.H. ; Data preparation and analysis A.H.; Visualization A.H.; Methodology A.H., A.A. and M.H. ; Interpretation A.H., A.A.; Supervision M.H.; Writing - original draft A.H.; Writing - review & editing A.A., S.S., N.C., A.D., T.K. and M.H. All authors read and approved the final manuscript.

Funding

The preparation of this manuscript was partially funded by the Innovation Fond of the Joint Federal Committee in the form of the project "isPO—Integrated, cross-sectoral psycho-oncology" (grant number: 01NVF17022-IsPo). Open Access funding enabled and organized by Projekt DEAL.

Data Availability
The datasets generated and/or analysed during the current study are not publicly available due to ethical and legal restrictions but are available from the corresponding author upon reasonable request.

Declarations

Ethics approval and consent to participate
The study was conducted in accordance with the Declaration of Helsinki and the Ethics Committee of the Medical Faculty of the University of Cologne has approved the isPO project (18–092). The isPO study is registered in the German Clinical Trials Registry (DRKS00015326). All participants have been informed about the aim of the study and about the voluntary nature of participation and have signed a written informed consent form for participation in the isPO care programme and data processing.

Consent for publication
Not applicable.

Competing interests
The authors declare no competing interests.

Author details
[1]Institute of Medical Statistics and Computational Biology (IMSB), Faculty of Medicine and University Hospital Cologne, University of Cologne, Cologne, Germany
[2]Medicine Faculty, Human Sciences Faculty, Institute for Medical Sociology, Health Services Research, and Rehabilitation Science (IMVR), University of Cologne, University Hospital Cologne, Cologne, Germany
[3]Goethe University Frankfurt, Institute of General Practice, Frankfurt, Germany

Received: 1 February 2023 / Accepted: 14 May 2023
Published online: 30 May 2023

References

1. Teixeira da Silva JA, Tsigaris P, Erfanmanesh M. Publishing volumes in major databases related to Covid-19. Scientometrics. 2021;126(1):831–42.
2. Fleming TR, Labriola D, Wittes J. Conducting Clinical Research during the COVID-19 pandemic: protecting Scientific Integrity. JAMA. 2020;324(1):33–4.
3. Psotka MA, et al. Conduct of clinical trials in the era of COVID-19: JACC Scientific Expert Panel. J Am Coll Cardiol. 2020;76(20):2368–78.
4. Silvestri G, et al. The role of the psycho-oncologist during the COVID-19 pandemic: a clinical breast Cancer Case Report. Behav Sci. 2022;12(7):211.
5. McAndrew NS, et al. Impact of the COVID-19 pandemic on cancer patients and psycho - oncology providers: perspectives, observations, and experiences of the american Psychosocial Oncology Society membership. Psycho - Oncology. 2022;31(6):1031–40.
6. Al-Quteimat OM, Amer AM. The impact of the COVID-19 pandemic on cancer patients. American journal of clinical oncology; 2020.
7. Sanz A, et al. New Challenges for Psycho-Oncology during COVID-19 pandemic. Urgent implementation of Telepsychology. J Cancer Therapy. 2021;12(2):97–105.
8. Jenniches I, et al. Evaluation of a complex integrated, cross-sectoral psycho-oncological care program (isPO): a mixed-methods study protocol. BMJ open. 2020;10(3):e034141.
9. Kusch M, et al. Integrated, cross-sectoral psycho-oncology (isPO): a new form of care for newly diagnosed cancer patients in Germany. BMC Health Serv Res. 2022;22(1):543.
10. Hagemeier A, Samel C, Hellmich M. The regression discontinuity design: methods and implementation with a worked example in health services research. Z Evid Fortbild Qual Gesundhwes. 2022;172:71–7.
11. Cucinotta D, Vanelli M. WHO declares COVID-19 a pandemic. Acta Bio Medica: Atenei Parmensis. 2020;91(1):157.
12. Innovationsausschuss beim Gemeinsamen Bundesausschuss GBA. nVF-isPO – Integrierte, sektorenübergreifende Psychoonkologie. 2023, [20.01.2023]; Available from: https://innovationsfonds.g-ba.de/projekte/neue-versorgungsformen/ispo-integrierte-sektorenuebergreifende-psychoonkologie.103.
13. Thistlethwaite DL, Campbell DT. Regression-discontinuity analysis - an alternative to the Ex-Post-Facto Experiment. J Educ Psychol. 1960;51(6):309–17.
14. Snaith RP. The hospital anxiety and Depression Scale. Health Qual Life Outcomes. 2003;1:29.
15. Mitchell AJ, Meader N, Symonds P. Diagnostic validity of the hospital anxiety and Depression Scale (HADS) in cancer and palliative settings: a meta-analysis. J Affect Disord. 2010;126(3):335–48.
16. Corp IBM. N., *IBM SPSS Statistics for Windows, Version 28*, in *IBM Corp, Armonk, NY*. 2021.
17. Bargon CA, et al. Impact of the COVID-19 pandemic on patient-reported outcomes of breast Cancer patients and survivors. JNCI Cancer Spectr. 2021;5(1):pkaa104.
18. Deng J, et al. The prevalence of depression, anxiety, and sleep disturbances in COVID-19 patients: a meta-analysis. Ann N Y Acad Sci. 2021;1486(1):90–111.
19. Tsamakis K, et al. COVID–19 and its consequences on mental health. Experimental and therapeutic medicine. 2021;21(3):1–1.
20. Leitlinienprogramm Onkologie. S3-Leitlinie Psychoonkologische Diagnostik, Beratung und Behandlung von erwachsenen Krebspatienten. 2014, [18.01.2023]; Available from: http://leitlinienprogramm-onkologie.de/Leitlinien.7.0.html.
21. Murphy L, et al. The impact of the COVID-19 pandemic and its related restrictions on people with pre-existent mental health conditions: a scoping review. Arch Psychiatr Nurs. 2021;35(4):375–94.
22. Inchausti F, et al. Psychological intervention and COVID-19: what we know so far and what we can do. J Contemp Psychother. 2020;50(4):243–50.
23. Kallay E, Medrea F, Degi CL. On top of that all, now Covid-19, too. A scoping review of specificities and correlates of fear of cancer recurrence in breast cancer patients during COVID-19. Breast. 2022;62:123–34.
24. Toquero P, et al. Emotional distress in Cancer Patients during the First Wave of the COVID-19 pandemic. Front Psychol. 2021;12:755965.
25. Bauerle A et al. Mental Health Burden of German Cancer Patients before and after the outbreak of COVID-19: predictors of Mental Health Impairment. Int J Environ Res Public Health, 2021. 18(5).
26. Sigorski D, et al. Impact of COVID-19 on anxiety levels among patients with cancer actively treated with systemic therapy. ESMO Open. 2020;5(5):e000970.
27. Eckford RD, et al. The COVID-19 pandemic and Cancer patients in Germany: impact on treatment, Follow-Up care and psychological Burden. Front Public Health. 2021;9:788598.
28. Moraliyage H, et al. Cancer in Lockdown: impact of the COVID-19 pandemic on patients with Cancer. Oncologist. 2021;26(2):e342–4.
29. Krieger T, Cecon SS, Pfaff N, Dresen H. A., Summative evaluation des projekts isPO (SE 2.0). Health Services Research, and Rehabilitation Science: Cologne: Institute for Medical Sociology; 2022.
30. Hesary FB, Salehiniya H. The impact of the COVID-19 epidemic on diagnosis, treatment, concerns, problems, and Mental Health in patients with gastric Cancer. J Gastrointest Cancer. 2022;53(3):797–804.
31. Burgess R. Moving from face-to-face to telephone assessments with patients during the COVID-19 pandemic. Nurs Stand. 2021;36(10):45–50.
32. Saladino V, Algeri D, Auriemma V. The psychological and social impact of Covid-19: New Perspectives of Well-Being. Front Psychol. 2020;11:577684.

Publisher's Note

4.2 Extension of RDD to LMM-RDD - 3. Article

Published as:

Extension of Regression Discontinuity Designs to evaluate time effects - a quasi-experimental approach using the data from the new form of care "Integrated cross-sectoral psyco-oncology (nFC- isPO)"

Anna Hagemeier, Kerstin Rosenberger, Martin Hellmich

PLoS One, 2026 Feb, 21(2): e0343414.

https://doi.org/10.1371/journal.pone.0343414

Visual Abstract:

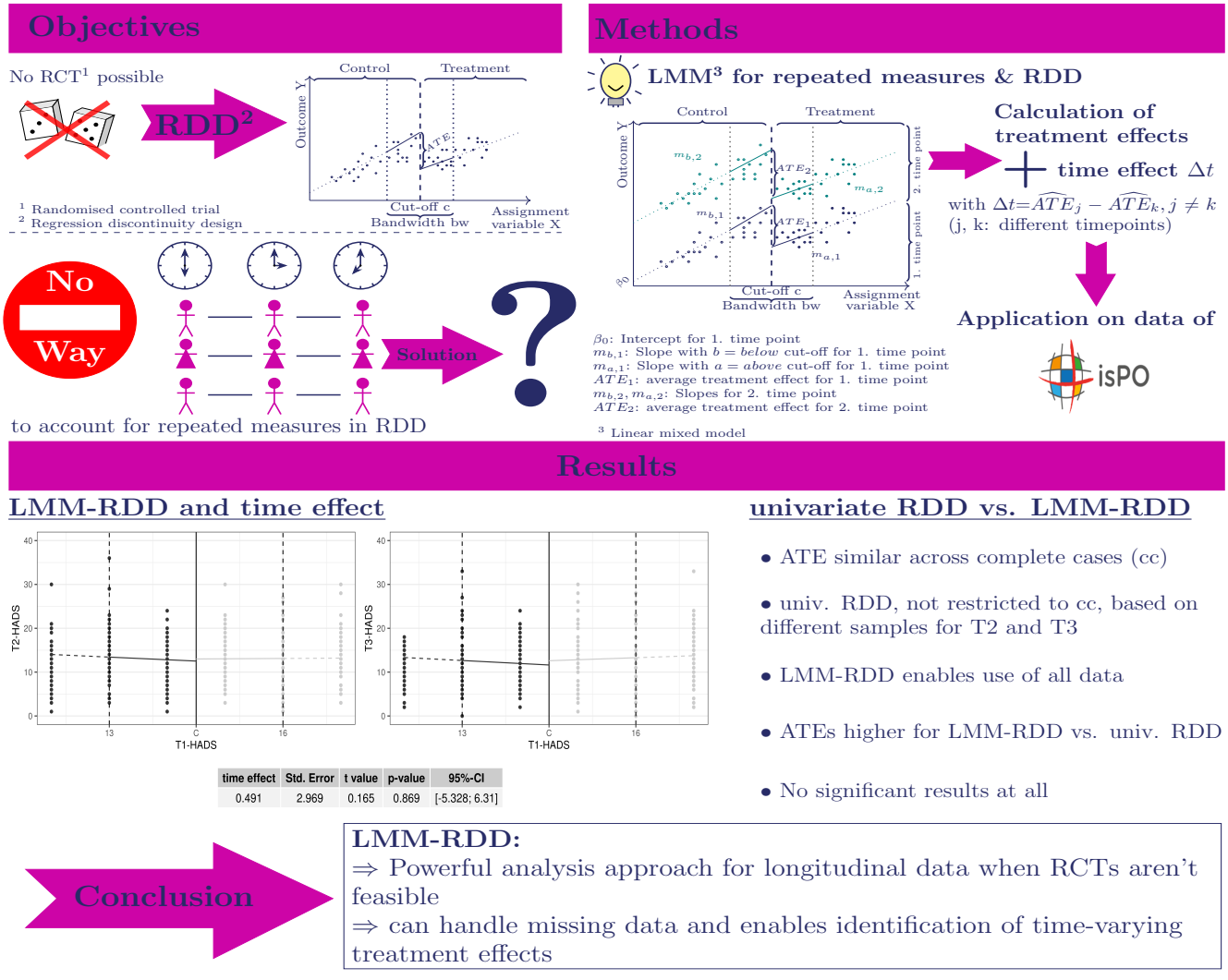


Figure 8: Visual abstract: "Extension of Regression Discontinuity Designs to evaluate time effects - a quasi-experimental approach using the data from the new form of care "Integrated cross-sectoral psyco-oncology (nFC- isPO)""

RESEARCH ARTICLE

Extension of Regression Discontinuity Designs (RDD) to evaluate time effects – A quasi-experimental approach using the data from the new form of care "Integrated cross-sectoral psycho-oncology (nFC- isPO)"

Anna Hagemeier[1*], **Kerstin Daniela Rosenberger**[1], **Martin Hellmich**[1,2]

1 Institute of Medical Statistics and Computational Biology, Medical Faculty, University of Cologne, University Hospital Cologne, Cologne, Germany, **2** Department of Medical Statistics, University Medical Center Göttingen, Göttingen, Germany

* anna.hagemeier@uni-koeln.de

Abstract

Objective

Randomised controlled trials (RCTs) are considered the gold standard in clinical research. However, randomization isn't always feasible, e.g., due to ethical considerations. As a result, quasi-experimental designs, such as the regression-discontinuity design (RDD), have been advocated as valuable alternatives and are increasingly employed in health sciences. So far, however, there is no way to take repeated measures into account in an RDD analysis, although repeated measures are not uncommon in medical research. This article proposes an extension of the RDD methodology to incorporate repeated measures within the statistical analysis.

Methods

The article introduces consistent mathematical notation for combining a univariate RDD with a linear mixed model (LMM) to account for repeated measures. The application is presented using data from the nFC-isPO (N = 1,417), where the Hospital Anxiety and Depression Scale (HADS) was employed to assess anxiety and depression in newly diagnosed cancer patients over a 12-month treatment period. The HADS scores were measured at baseline (T1), after 4 months (T2) and after 12 months (T3). Patients were assigned to control (psychosocial care) or treatment (psycho-oncological-psychotherapeutic care) group based on a predetermined HADS threshold at T1.

Results

The average treatment effect (ATE) was similar in both univariate RDD and LMM-RDD when applied to complete cases. Including all available data (T2 and/or T3),

OPEN ACCESS

Citation: Hagemeier A, Rosenberger KD, Hellmich M (2026) Extension of Regression Discontinuity Designs (RDD) to evaluate time effects – A quasi-experimental approach using the data from the new form of care "Integrated cross-sectoral psycho-oncology (nFC- isPO)". PLoS One 21(2): e0343414. https://doi.org/10.1371/journal.pone.0343414

Editor: Mohammed Misbah Ul Haq, Deccan School of Pharmacy, INDIA

Received: March 1, 2025

Accepted: January 16, 2026

Published: February 26, 2026

Data availability statement: The datasets generated and/or analysed during the current study are not publicly available due to ethical and legal restrictions. Data are available from the Institute of Medical Statistics and

Computational Biology (IMSB) of the Medical Faculty of the University of Cologne and the University Hospital Cologne, Germany (contact via imsb-kontakt@uni-koeln.de) for researchers who meet the criteria for access to confidential data.

Funding: The preparation of this manuscript was partially funded by the Innovation Fonds of the Joint Federal Committee (Innovationsausschuss beim Gemeinsamen Bundesausschuss) in the form of the project "isPO—Integrated, cross-sectoral psycho-oncology" (grant number: 01NVF17022-IsPo). The funder had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript. There was no additional external funding received for this study.

Competing interests: The authors have declared that no competing interests exist.

Abbrevations: ATE, average treatment effect; CAPSYS^{2020}, computer-assisted assistance system in isPO; HADS, Hospital Anxiety and Depression Scale; LMM-RDD, linear mixed model RDD; nFC-isPO, new form of care integrated, cross-sectoral psycho-oncology; NRW, North Rhine-Westphalia; RCT, Randomised controlled trial; RDD, Regression discontinuity design.

univariate RDD analyses were based on different samples for each time point. At T2, the ATE sign reversed from negative to positive, suggesting a change in discontinuity direction. At T3, the ATE magnitude nearly doubled compared to the complete case analysis. LMM-RDD estimated treatment effects were higher than those from univariate RDD. Nevertheless, none reached statistical significance. The time effect (Δt), representing the difference in treatment effects between two time points (complete case: Δt=0.577; T2 and/or T3 available: Δt=0.491), was not significant (p = 0.855; 95%-CI: [−5.589; 6.743]; p = 0.869, 95%-CI: [−5.328; 6.310]).

Conclusion

Extending the RDD to incorporate repeated measures within a linear mixed model (LMM-RDD) provides a more robust and comprehensive analytical approach when an RCT is not feasible. This approach not only addresses the challenges of missing data and an unbalanced number of outcome measures, but also allows for the identification of time-varying treatment effects that may not be discernible using traditional RDD.

Trial registration

The study was registered in the German Clinical Trials Registry on 30 October 2018 under the ID "DRKS00015326".

1. Introduction

Although the randomised controlled trial (RCT) is the gold standard in clinical research, it cannot be done in every setting. For example, if the intervention to be evaluated has already been implemented in standard practice withholding from certain patients may not be an ethical option anymore [1]. However, a quasi-experimental design, mimicking randomization of individuals, might still provide valid information to evaluate the intervention. Specifically, in the regression discontinuity design (RDD), which in its simplest form is based on ordinary linear regression, a predefined threshold is used to assign cases to treatment groups. The control group receives standard care, while the intervention group receives standard care plus extended care. This assignment is based on a continuous parameter measured before the intervention. A quasi-randomisation is created around the threshold within a certain pre-defined range (bandwidth). Within this range, the treatment effect can be estimated, visible as the discontinuity of two regression lines at the threshold [2]. The design has its origins in 1960 [3], however was rarely used in the health sciences until 2010, and thereafter has become increasingly popular, especially in the last five years [4]. Despite the increasing use of this design and the ability to examine causal relationships without random assignment, there is currently no way to account for repeated measures within the RDD. Repeated measures are common in medical research and many studies use them to estimate treatment effects over time [5], including the new

PLOS One

Form of Care and integrated cross-sectoral Psycho-Oncological (nFC-isPO) study. The nFC-isPO ran from 2017 to 2022 and aimed to reduce anxiety and depression in newly diagnosed cancer patients during 12 months of care. The Hospital Anxiety and Depression Scale (HADS) was used to measure the patient's burden at time of diagnosis (baseline, T1), after 4 months (T2) and after 12 months of treatment (primary outcome, T3) [6,7]. The study was conducted and analysed using the RDD where both the continuous assignment variable and the outcome variable are HADS scores measured at different time points. The assignment into control (HADS≤14) and intervention (HADS≥15) group was done using the baseline HADS value (T1, assignment variable) with c = 14.5 as the cut-off value [8,9]. This threshold value is established in psycho-oncological care and was used in the isPO study due to its clinical relevance for the classification of 'moderate to high psychological stress' [7,10,11]. For final evaluation of the study, the time points T2 and T3, were analysed separately, as with the standard RDD approach it is not possible to take into account time effects. In order to obtain additional information on temporal relationships, mixed models for repeated measures were applied in secondary analyses [12]. However, an integrated approach considering time effects within the RDD analysis was not yet described in literature.

The aim of this article is to oppose and formally describe mixed model repeated measures approach to analyse the RDD for a scenario with continuous outcome measured at 3 time points (baseline T1, T2, T3) and to exemplify this novel approach with recent data from the nFC-isPO.

2. Methods

2.1 Model extension

To address the effect over time, the functional form of the RDD, which is based on a simple linear regression, needs to be extended to a linear mixed model for repeated measures.

Defining Y as the continuous outcome for n observations, X as the covariate that is the continuous assignment variable for each observation and c as the cut-off value for treatment assignment based on the covariate X. Furthermore, define $x_i^* = x_i - c$ as the measure of the distance to the cut-off and let 1 denote the indicator function for the treatment allocation, that equals 0 *if* $x_i \leq c$ (control group) or 1 *if* $x_i > c$ (treatment) (the so-called "sharp RDD").

Then the univariate (i.e., considering only one time point) RDD equation [2] is given as

$$\boldsymbol{y_i} = \boldsymbol{\beta_0} + \boldsymbol{\beta_1 x_i}^* + \boldsymbol{\beta_2} \mathbb{1}\,(\boldsymbol{x_i} > \boldsymbol{c})\, \boldsymbol{x_i}^* + \boldsymbol{\beta_3} \mathbb{1}\,(\boldsymbol{x_i} > \boldsymbol{c}) + \boldsymbol{e_i}$$

$$= \boldsymbol{\beta_0} + \boldsymbol{\beta_1 x_i}^* + \begin{cases} \boldsymbol{\beta_2 x_i}^* + \boldsymbol{\beta_3} + \boldsymbol{e_i}\,, & \textbf{if } \boldsymbol{x_i} > \boldsymbol{c} \\ \boldsymbol{e_i}\,, & \textbf{if } \boldsymbol{x_i} \leq \boldsymbol{c} \end{cases} \tag{1}$$

with $i = 1, 2, \ldots, n$ patients and residuals $e_i \sim N\left(0, \sigma^2\right)$ with mean of 0 and variance of σ^2 as the random error. Equation (1) corresponds to Fig 1 where β_0 is the intercept, β_1 denotes the slope (m_b) of the regression line below the cut-off (control), β_2 denotes the change in slope $(m_a = m_b + \beta_2)$ for the regression line above the threshold (treatment) compared to the control group and β_3 denotes the average treatment effect (ATE, discontinuity/ "jump" at cut-off).

A more detailed description of the RDD can be found in [2]. Equation (1) is then extended to a linear mixed model (LMM) for repeated measures over time:

$$y_{ij} = \beta_0 + \beta_1 x_i^* + \beta_2 \mathbb{1}\,(x_i > c)\, x_i^* + \beta_3 \mathbb{1}\,(x_i > c) + \mathbb{S}_{ij} + b_{i0} + e_{ij} \tag{2}$$

$$\textit{with } \mathbb{S}_{ij} = \sum_{j'=2}^{J} D_{j'j} [\beta_{4,j'} + \beta_{5,j'} x_i^* + \beta_{6,j'} \mathbb{1}\,(x_i > c)\, x_i^* + \beta_{7,j'} \mathbb{1}\,(x_i > c)]$$

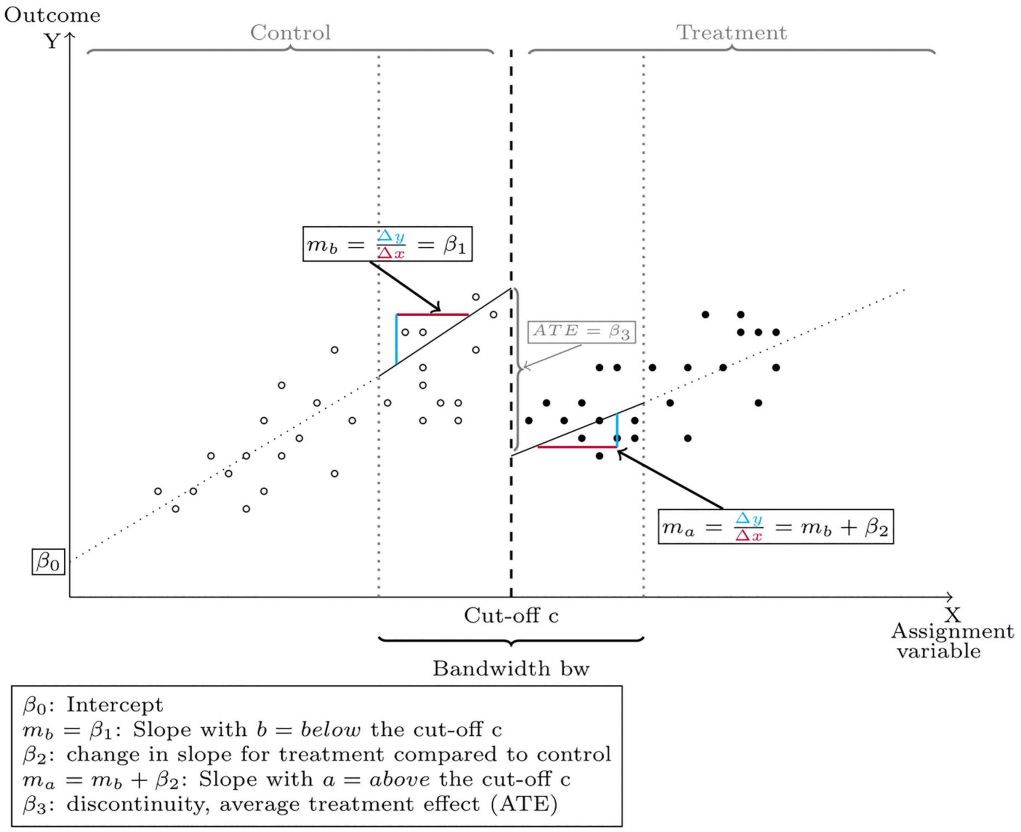


Fig 1. Plot of univariate Regression-Discontinuity-Design.

https://doi.org/10.1371/journal.pone.0343414.g001

where Y is defined as the continuous outcome for patient $i = 1,\ 2,\ \ldots,\ n$ and time points $j = 1, 2, 3, \ldots,\ J$.

$D_{j'j} = \begin{cases} 1, & \text{if } j' = j \\ 0, & \text{otherwise} \end{cases}$ is a dummy variable to model the time effect with running index $j' = 2, 3, \ldots, J$.

$\beta_{4,j'}$ denotes the change in intercept at time point j' compared to time point 1. $\beta_{5,j'}$ *and* $\beta_{6,j'}$ denote the changes in slopes at time point j' compared to time point 1 and $\beta_{7,j'}$ is the change in the treatment effect for time point j' compared to time point 1. $b_{i0} \sim N(0, \sigma_b)$, with a mean of 0 and a variance σ_b, is the random intercept for patient i to account for stochastic dependence over time (repeated measures) and $e_{i,j} \sim N(0, \sigma)$ with a mean of 0 and a variance σ denotes the vector of error terms for patient i at time point j. A random slope could be defined in a similar way for each coefficient, but is not considered further below for the purpose of simplicity.

For illustration, equation (3) shows the example for just two outcome measures, i.e., two time points with $J = 2$ ($j = 1, 2$ and $j' = 2$) for $i = 1, 2, \ldots,\ n$ patients:

$$\begin{aligned} y_{i,j} &= \beta_0 + \beta_1 x_i^* + \beta_2 \mathbb{1}(x_i > c)\, x_i^* + \beta_3 \mathbb{1}(x_i > c) + D_{j',j}[\beta_{4,j'} + \beta_{5,j'} x_i^* + \beta_{6,j'} \mathbb{1}(x_i > c)\, x_i^* + \beta_{7,j'} \mathbb{1}(x_i > c)] + b_{0i} + e_{i,j} \\ &= \beta_0 + \beta_1 x_i^* + \begin{cases} \beta_2 x_i^* + \beta_3 + D_{j',j}[\beta_{4,j'} + \beta_{5,j'} x_i^* + \beta_{6,j'} x_i^* + \beta_{7,j'}] + b_{0i} + e_{i,j}, & \text{if } x_i > c \\ D_{j',j}[\beta_{4,j'} + \beta_{5,j'} x_i^*] + b_{0i} + e_{i,j}, & \text{if } x_i \le c \end{cases} \end{aligned} \tag{3}$$

With $D_{2,1} = 0$ and $D_{2,2} = 1$ a schematic representation of equation (3) is given in Fig 2.

The extended model now considers the correlation between repeated measures for each observation. Moreover, allows the inclusion of cases with only one follow-up outcome measure, assuming that these are missing at random. The extension of the RDD equation to a linear mixed model RDD for repeated measures allows the evaluation of the time effect within an RDD analysis. This is achieved by estimating the difference between two regression coefficients corresponding to the treatment effects at individual time points.

2.2 Calculation of the time effect

To evaluate the time effect, the standard error, the corresponding confidence interval as well as the corresponding t-statistic- and p-value can be calculated for the difference between two regression coefficients [13,14]. First, the difference between

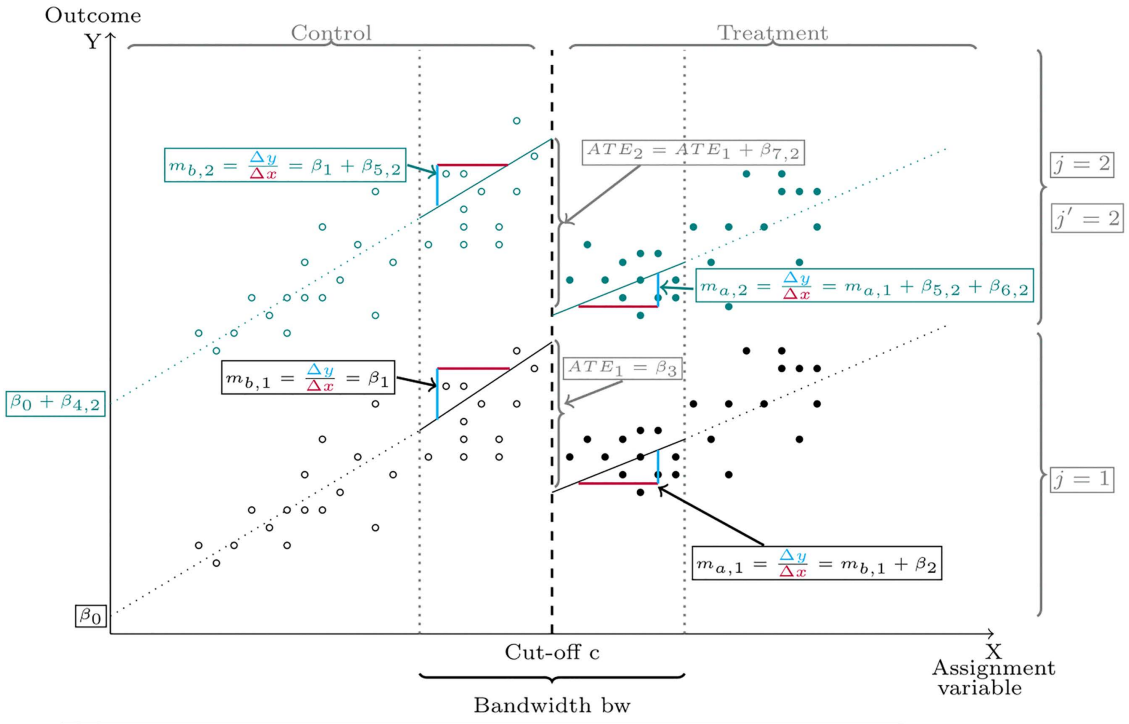


Fig 2. Plot of the LMM-RDD for the special case of j=1,2.

https://doi.org/10.1371/journal.pone.0343414.g002

the coefficients of interest is calculated, where these coefficients represent the treatment effect at different time points $j, k \in \{1, 2, \ldots, J\}$. The difference corresponds to the time effect and is given by $\Delta t = \widehat{ATE}_j - \widehat{ATE}_k$ with $j \neq k$ and where the head notation denotes an estimate. The corresponding standard error can be derived using Bienaymé's theorem [15]:

$$SE(\Delta t) = \sqrt{Var(\Delta t)} = \sqrt{Var\left(\widehat{ATE}_j - \widehat{ATE}_k\right)} = \sqrt{Var\left(\widehat{ATE}_j\right) + Var\left(\widehat{ATE}_k\right) - 2 \cdot Cov\left(\widehat{ATE}_j,\ \widehat{ATE}_k\right)} \tag{4}$$

The standard error and the estimated difference of the coefficients (time effect) can then be used to calculate the corresponding t-value: $t_{\Delta t} = \frac{\Delta t}{SE(\Delta t)}$ and via the z-score (which approximates the t-value for large samples [16]) the p-value can be calculated by $p_{\Delta t} = 2 \cdot (1 - P(Z \leq |z|))$. Finally, a confidence interval (CI) for the difference between the coefficients (time effect) can be determined:

$$95\%\ CI = \Delta t \pm z_{\frac{\alpha}{2}} SE(\Delta t) \text{ with } z_{\frac{\alpha}{2}} = 1.96.$$

In the following, both the simple linear RDD and the LMM-RDD for repeated measures (with unstructured variance-covariance matrix over time) are exemplified using the data from the isPO study. All analyses are exploratory and were done using R 4.4.0 [17]. Results were considered statistically significant at a significance level of $\alpha = 0.05$ when $p \leq 0.05$.

2.3 Data collection and analysis population

Between January 15, 2019 and March 31, 2021, N = 1,756 newly diagnosed cancer patients were enrolled in four isPO care networks (Cologne, Troisdorf, Mönchengladbach and Neuss) in North Rhine-Westphalia (NRW) and documented in the CAPSYS^{2020} computer assistance system developed specifically for the project. Each patient received individualized care and was therefore assigned to different treatment groups based on their needs as determined by their HADS scores (baseline, T1) at study enrolment. Patients with HADS scores ≤ 14 received a psychosocial intervention (control group) and patients with scores ≥ 15 received psychooncological, psychotherapeutical treatment (intervention group). After 4 (T2) and 12 months (T3), the HADS was measured again, and the treatment effect for the primary analysis in isPO was analysed separately for the two measurement time points (T2, T3) using standard RDD within a predefined bandwidth (HADS between 13–16) around the cut-off value [9]. Since a HADS result had to be available at least after 4 or 12 months for evaluation, 339 cases were excluded. This results in an N of 1,417 cases for the analysis population to estimate the treatment and time effects via LMM-RDD (Fig 3). All patients gave their written informed consent.

2.4 Ethics statement

The study was conducted in accordance with the Declaration of Helsinki and the Ethics Committee of the Medical Faculty of the University of Cologne has approved the isPO project (18–092). The isPO study is registered in the German Clinical Trials Registry (DRKS00015326). All participants have been informed about the aim of the study and about the voluntary nature of participation and have signed a written informed consent form for participation in the isPO care programme and data processing.

3. Results

3.1 Characteristics of the patients

N = 1,417 nFC-isPO patients had at least one HADS follow-up outcome available at 4 (T2) or 12 months (T3). Limited to the relevant bandwidth of 13–16 points on the HADS scale at baseline (T1), N = 247 cases remain for primary analysis. Around two thirds of them were female. 61.5% reported to live in a relationship and in the same household with their respective partners. Detailed information about main characteristics of the patients can be found in Table 1.

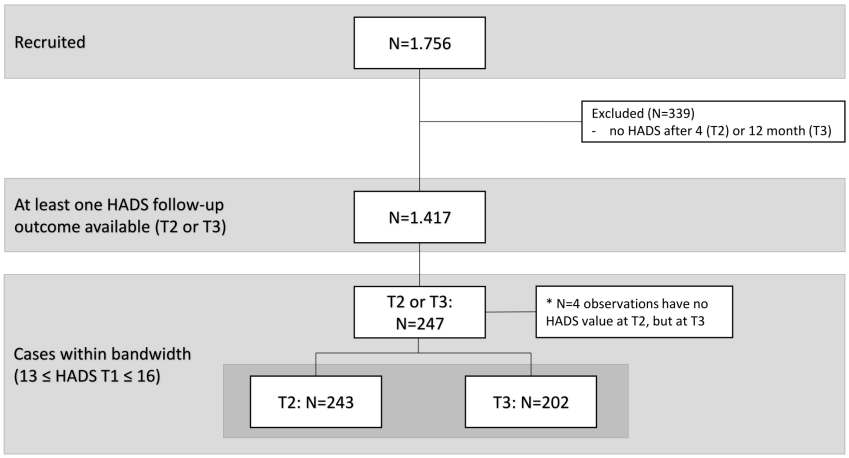


Fig 3. Flowchart of study and analysis population.

https://doi.org/10.1371/journal.pone.0343414.g003

3.2 Results of the univariate and mixed model RDD

Table 2 shows the results of the univariate and linear mixed model RDD for the cases with complete follow-up data and for the case where either T2 and/or T3 are available. To illustrate the difference in results, depending on the data used, both modelling approaches were applied to cases with complete follow-up data to show that the result, the $\widehat{ATE}$, is the same ($\widehat{ATE}_{T2} = -0.213$; $\widehat{ATE}_{T3} = 0.364$) regardless of whether the RDD is performed in its usual univariate or extended linear mixed model form. When the models are applied to patients with at least one follow-up visit (T2 and/or T3 available), univariate RDD estimates are based on (slightly) different samples of cases for each time point. In the univariate case for T2, the sign changes from negative to positive and the discontinuity changes direction. Although the sign of the $\widehat{ATE}$ estimate changes, both estimates were very close to zero with wide confidence intervals. For T3, the $\widehat{ATE} = 0.676$ is almost twice as high as in the model applied to the cases with complete follow-up data, but again the estimate is very small with a wide confidence interval. The treatment effects estimated by the LMM-RDD are even higher than the effects calculated in the univariate models, but still with large confidence intervals. None of the results reached statistical significance.

3.3 Effect over time

Within the mixed model, it is now possible to calculate the time effect as the difference between the two treatment effects at the different time levels. The time effect with cases with complete follow-up data was $\Delta t = 0.364 - (-0.213) = 0.577$ ($p = 0.855$; $95\% - CI: [-5.589; 6.743]$) with a standard error of 3.146 and a t-value of 0.183. Fig 4 is the graphical representation of the LMM-RDD in which the HADS value of T2 or T3 are not missing (see last row of Table 2). In order to obtain an overview of the data, the upper scatter plots show the relationship between the HADS values at time T1 and the respective outcome measurements at T2 and T3 over the entire value range of the HADS, whereby it can be seen that the respective outcome value increases with increasing initial value. Below is the representation of the LMM-RDD within the specified bandwidth with an $\widehat{ATE}$ of 0.481 for T2 and an $\widehat{ATE}$ of 0.972 for T3 (both not significant; s. Table 2). The table below all four plots contains the estimated time effect ($\Delta t = 0.491$), which is also not significant ($p = 0.869$, $95\% - CI: [-5.328; 6.310]$).

PLOS One

Table 1. Sociodemographic and clinical patient characteristics for analysis population and the population within the bandwidth (HADS between 13 to 16) for the primary analysis of the RDD.

	Bandwidth (N = 247)	Total (N = 1,417)
Gender		
Female	160 (64.8%)	902 (63.7%)
Male	87 (35.2%)	515 (36.3%)
Age		
Mean ± SD	56.1 ± 14.0	57.1 ± 13.0
Median [Q1; Q3]	57.0 [49.0;66.5]	58.0 [50.0;66.0]
Care level		
Cancer self-help (care level 1)	7 (2.8%)	122 (8.6%)
Psychosocial cancer counselling (care level 2)	109 (44.1%)	523 (36.9%)
Psycho-oncological psychotherapy (care level 3)	131 (53.0%)	772 (54.5%)
Relationship status		
No relationship	68 (27.5%)	331 (23.4%)
In relationship. shared household	152 (61.5%)	942 (66.5%)
In relationship. separate households	16 (6.5%)	90 (6.4%)
Missing	11 (4.5%)	54 (3.8%)
ISCED		
Primary education	3 (1.2%)	23 (1.6%)
Lower secondary education	11 (4.5%)	90 (6.4%)
Upper secondary education	169 (68.4%)	956 (67.5%)
Bachelor or equivalent	10 (4.0%)	94 (6.6%)
Master or equivalent	43 (17.4%)	199 (14.0%)
Doctoral or equivalent	4 (1.6%)	17 (1.2%)
Missing	7 (2.8%)	38 (2.7%)
Severe disability		
No	168 (68.0%)	959 (67.7%)
Yes	74 (30.0%)	432 (30.5%)
Missing	5 (2.0%)	26 (1.8%)
Chronic illness		
No	140 (56.7%)	815 (57.5%)
Yes	102 (41.3%)	576 (40.6%)
Missing	5 (2.0%)	26 (1.8%)
ICD 10 / Tumor-entities		
Digestive organs (C15-C26)	31 (12.6%)	219 (15.5%)
Respiratory and intrathoracic organs (C30-C39)	18 (7.3%)	135 (9.5%)
Skin (C43-C44)	14 (5.7%)	82 (5.8%)
Breast (C50)	72 (29.1%)	373 (26.3%)
Female genital organs (C51-C58)	13 (5.3%)	83 (5.9%)
Lymphoid, hematopoietic and related tissue (C81-C96)	26 (10.5%)	147 (10.4%)
Other (each<5%)	67 (27.1%)	345 (24.3%)
Missing	6 (2.4%)	33 (2.3%)
Death		
No	236 (95.5%)	1,360 (96.0%)
Yes	11 (4.5%)	57 (4.0%)

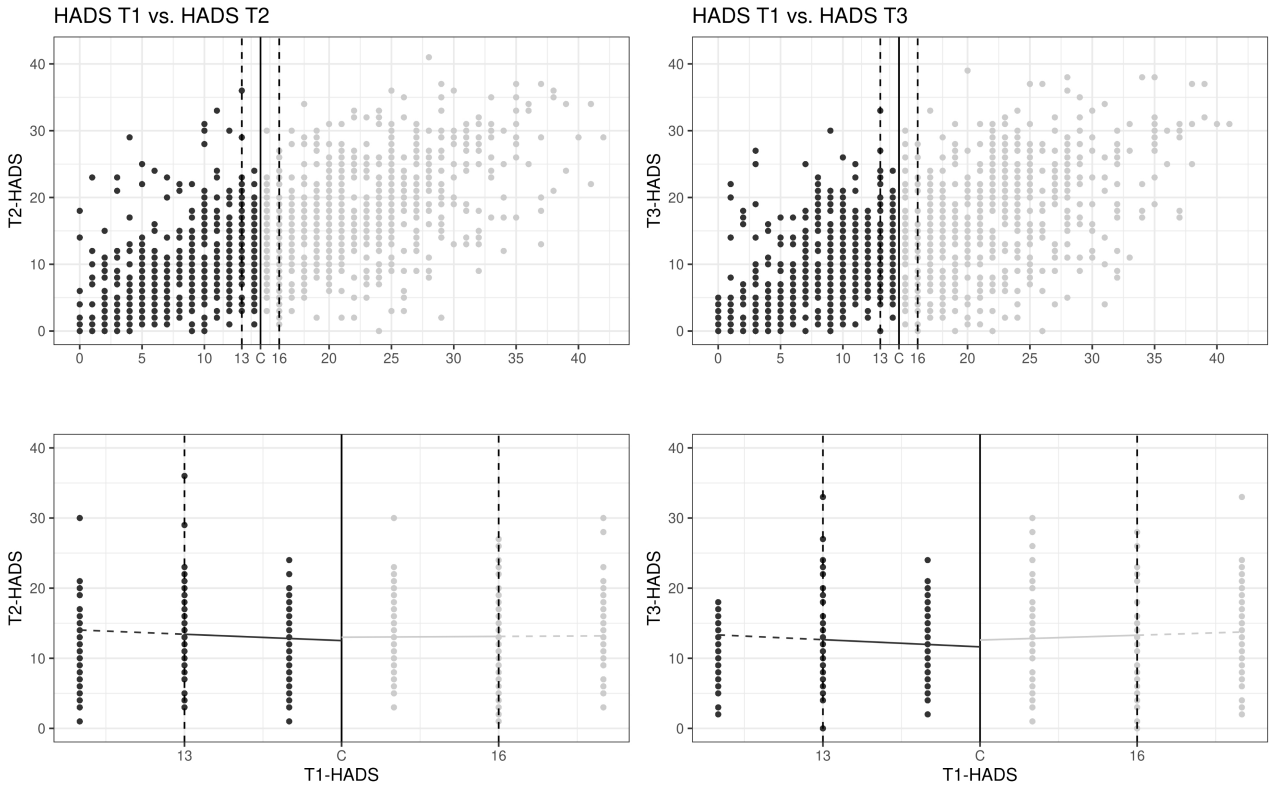
https://doi.org/10.1371/journal.pone.0343414.t001

PLOS One

Table 2. Results of the univariate and the linear mixed model RDD.

	T2						T3					
	N_{bw}	ATE	Std. Err.	t-value	p-value	95% CI	N_{bw}	ATE	Std. Err.	t-value	p-value	95% CI
Complete case												
Univ. RDD	198	−0.213	1.759	−0.121	0.904	[-3.682; 3.255]	198	0.364	2.070	0.176	0.861	[-3.719; 4.447]
LMM-RDD	198	−0.213	1.921	−0.111	0.912	[-3.994; 3.567]	198	0.364	1.921	0.189	0.850	[-3.401; 4.129]
T2 or T3 not missing												
Univ. RDD	243	0.322	1.665	0.193	0.847	[-2.958; 3.601]	202	0.676	2.053	0.329	0.742	[-3.372; 4.724]
LMM-RDD	243	0.481	1.775	0.271	0.786	[-3.01; 3.973]	202	0.972	1.870	0.520	0.603	[-2.693; 4.637]

https://doi.org/10.1371/journal.pone.0343414.t002



time effect	Std. Error	t value	p-value	95%-CI
0.491	2.969	0.165	0.869	[-5.328; 6.31]

Fig 4. Scatter plot (four panels) of the linear mixed model regression discontinuity design and the estimated time effect (Δt) for the comparison of the patients' outcome HADS scores at T2 (after 4 months) and T3 (after 12 months of treatment).

https://doi.org/10.1371/journal.pone.0343414.g004

4. Discussion

The RDD gives a promising alternative for drawing causal conclusions in case random group assignment is not feasible. Previously, it was not possibility to include repeated measures in the RDD. In order to close this gap, the simple linear RDD was extended to a repeated measure linear mixed model RDD by including the factor time and a random intercept (see equation 2). To demonstrated the advantages and differences of the linear mixed model RDD compared to the univariate RDD, data from the nFC-isPO study were reanalysed. First, only cases with complete follow-up data and, second all data (T2 and/or T3 are available) were considered. With the univariate method, a direct comparison between the time points on the basis of the calculated effects for cases with complete follow-up data is not possible. A valid 95% CI needs to account for the statistical dependence of both time points (repeated measures). Therefore, only a graphical interpretation or comparison of the RDD lines is possible. When using all available data, a graphical analysis is still possible, but even more limited, as the calculations of the treatment effect are based on (slightly) different samples. This implies that the database should be limited to cases with complete follow-up data or cases with valid primary endpoint, thus information is deliberately excluded from the analysis. With a linear mixed model a restriction to complete follow-up data is not required. When effects are estimated, particularly due to treatment and time, the correlation structure of repeated measures is accounted for [12].

4.1 Limitations and open issues

A major limitation of conducting a study in the medical field is the often limited sample size, which poses a problem when using an RDD. Simulation studies have shown that RDDs require a larger amount of data than RCTs in order to detect (moderate) treatment effect sizes and achieve comparable statistical power. The reason for this is that only units that are close to the threshold value can be compared – i.e., in the pre-specified range where quasi-randomisation can be assumed [18–20]. This might be one explanation for the overall insignificant results, including the rather wide confidence intervals of the analyses performed. Moreover, with regard to the application of the methods to the nFC-isPO study data, there was the limitation that the cases within the bandwidth were either too similar in terms of the individual need for psycho-oncological care or the burden above the defined threshold was too low to be able to demonstrate a relevant treatment effect.

This article extends a current method and demonstrates its feasibility with empirical data. For clarity, the analysis was limited to a random intercept model; however, further research could incorporate random slopes to gain valuable insights into individual differences over time. Further investigation is also needed to systematically compare the performance of the univariate and LMM-RDD approaches in a simulation study, considering various sample sizes, effect sizes, and missing data patterns.

5. Conclusion

Applying a linear mixed model (LMM) to a regression discontinuity design (RDD) with repeated measures is a powerful approach to analyse longitudinal data in case a RCT cannot be conducted. The presented combined model (LMM-RDD) enables the estimation of time effects, which has not been possible with the existing approaches, yet.

Acknowledgments

We thank all participating patients, all the clinical staff involved, and the cooperating health insurance companies for their support in the isPO project. We also thank all other consortium partners in the isPO project.

Author contributions

Conceptualization: Anna Hagemeier, Martin Hellmich.

Data curation: Anna Hagemeier.

Formal analysis: Anna Hagemeier.

Methodology: Anna Hagemeier, Kerstin Daniela Rosenberger, Martin Hellmich.

Supervision: Martin Hellmich.

Visualization: Anna Hagemeier.

Writing – original draft: Anna Hagemeier.

Writing – review & editing: Kerstin Daniela Rosenberger, Martin Hellmich.

References

1. Bor J, Moscoe E, Mutevedzi P, Newell M-L, Bärnighausen T. Regression discontinuity designs in epidemiology: causal inference without randomized trials. Epidemiology. 2014;25(5):729–37. https://doi.org/10.1097/EDE.0000000000000138 PMID: 25061922
2. Hagemeier A, Samel C, Hellmich M. The regression discontinuity design: methods and implementation with a worked example in health services research. Z Evid Fortbild Qual Gesundhwes. 2022;172:71–7. https://doi.org/10.1016/j.zefq.2022.04.014 PMID: 35718728
3. Thistlethwaite DL, Campbell DT. Regression-discontinuity analysis: an alternative to the ex post facto experiment. J Educ Psychol. 1960;51(6):309–17. https://doi.org/10.1037/h0044319
4. Hilton Boon M, Craig P, Thomson H, Campbell M, Moore L. Regression discontinuity designs in health: a systematic review. Epidemiology. 2021;32(1):87–93. https://doi.org/10.1097/EDE.0000000000001274 PMID: 33196561
5. Cleophas TJ, Zwinderman AH, van Ouwerkerk B. Clinical research: a novel approach to the analysis of repeated measures. Am J Ther. 2012;19(1):e1-7. https://doi.org/10.1097/MJT.0b013e3181ed83b0 PMID: 21048432
6. Snaith RP. The hospital anxiety and depression scale. Health Qual Life Outcomes. 2003;1:29.
7. Mitchell AJ, Meader N, Symonds P. Diagnostic validity of the Hospital Anxiety and Depression Scale (HADS) in cancer and palliative settings: a meta-analysis. J Affect Disord. 2010;126(3):335–48. https://doi.org/10.1016/j.jad.2010.01.067 PMID: 20207007
8. Jenniches I, Lemmen C, Cwik JC, Kusch M, Labouvie H, Scholten N, et al. Evaluation of a complex integrated, cross-sectoral psycho-oncological care program (isPO): a mixed-methods study protocol. BMJ Open. 2020;10(3):e034141. https://doi.org/10.1136/bmjopen-2019-034141 PMID: 32156765
9. Kusch M, Labouvie H, Schiewer V, Talalaev N, Cwik JC, Bussmann S, et al. Integrated, cross-sectoral psycho-oncology (isPO): a new form of care for newly diagnosed cancer patients in Germany. BMC Health Serv Res. 2022;22(1):543. https://doi.org/10.1186/s12913-022-07782-0 PMID: 35459202
10. Kusch M, Labouvie H, Hein-Nau B. Klinische Psychoonkologie. Springer-Verlag; 2012.
11. Vodermaier A, Millman RD. Accuracy of the Hospital Anxiety and Depression Scale as a screening tool in cancer patients: a systematic review and meta-analysis. Support Care Cancer. 2011;19(12):1899–908. https://doi.org/10.1007/s00520-011-1251-4 PMID: 21898134
12. Detry MA, Ma Y. Analyzing repeated measurements using mixed models. JAMA. 2016;315(4):407–8.
13. Clogg CC, Petkova E, Haritou A. Statistical methods for comparing regression coefficients between models. Am J Sociol. 1995;100(5):1261–93.
14. Altman D, et al. Statistics with confidence: confidence intervals and statistical guidelines. John Wiley & Sons; 2013.
15. Loeve M. Probability theory. 3rd ed. Dover Publications; 2017.
16. Hogg RV, Tanis EA, Zimmerman DL. Probability and statistical inference. Boston: Pearson; 2015.
17. Team RC. A language and environment for statistical computing. R Foundation for Statistical Computing; 2021.
18. Schochet PZ. Technical methods report: statistical power for regression discontinuity designs in education evaluations. NCEE 2008-4026. National Center for Education Evaluation and Regional Assistance; 2008.
19. Jacob R, et al. A practical guide to regression discontinuity. MDRC; 2012.
20. Cattaneo MD, Idrobo N, Titiunik R. A practical introduction to regression discontinuity designs: foundations. Cambridge University Press; 2020.

5 Discussion

Although RCTs are still considered the gold standard for conducting and evaluating studies in the field of health services research, they are often unsuitable in clinical situations or study settings with real world data for ethical or practical reasons. At the same time, alternative approaches with the next highest level of evidence, such as quasi-experiments like the RDD, reach their methodological limits when evaluating longitudinal data. Therefore, this dissertation examines an expanded methodological approach that combines the RDD as a study design and analysis method with the LMM approach in order to adequately consider longitudinal data for the first time. To this end, the dissertation comprises three consecutive publications. These describe the path from the methodological fundamentals of the RDD (first article) to the practical application of the RDD approach in the context of health research (second article) to the methodological expansion for longitudinal data (third article). Whereby the latter article forms the methodological core of the thesis.

In order to contextualise the three publications and their findings, first, the most important aspects of each article are summarised below. Building on this, a critical reflection on the methodological strengths and limitations is carried out within this chapter. Furthermore, the results of the dissertation are classified according to the current state of research and open issues are discussed. Finally, the derived conclusions and implications are presented for future research .

5.1 Key aspects of the publications

The three publications highlight different but complementary aspects of the RDD approach. Thereby, the key findings from the articles illustrate how the methodological approach has developed from theoretical fundamentals to clinical application and methodological expansion, and what potential it has for future research.

Table 5.1 summarises the aims and key aspects of the three articles. The first article aims to systematically introduce the standard RDD approach and its methodological fundamentals as a quasi-experimental method and demonstrate its practical applicability in health research. It provides researchers with a step-by-step guide for practical application, including programme code for various statistical software programmes. Thus, the publication emphasises that the RDD approach offers an accessible strategy for presenting causal inference when randomisation is not feasible and group assignment can be based on a threshold. Furthermore, a particular feature of RDD compared to RCT is that it represents both the design and the evaluation method of a study, thus closely linking the methodological logic of the study to

the statistical analysis. Consequently, the article forms the methodological basis of the dissertation, as it examines the theoretical principles of RDD and operationalises them in the health context. Simultaneously, the work establishes an important link between methodology and application of the standard approach by demonstrating how theoretical concepts can be implemented in practice. Overall, the article thus provides the conceptual framework for the dissertation, on which the subsequent publications – application of the standard RDD method in a clinical setting and methodological extension for repeated measurements – are based.

Table 5.1: Aims, contribution and key aspects of the three articles to the dissertation

	1. Article	***2. Article***	***3. Article***
Aims and Contribution	• methodological introduction of the quasi-experimental RDD • explanation of the basic mathematical notation • hands-on demo using an example from health research • presentation of strengths and limitations in practice • methodological and practical guide for use in health sciences	• application of known classic RDD approach in a health science study • analysis of COVID-19 pandemic as confounder in nFC-isPO • demonstrate the application of the method on real world data (nFC-isPO) • increase visibility and popularity within the health science filed • points out the limitations of the standard RDD with repeated measurements	• extension of RDD for repeated measurements • provides precise math notation of extension compared to standard approach • demonstrates application of the extended method with data from nFC-isPO • compares results of standard RDD with results of LMM-RDD • shows that LMM-RDD improves efficiency and interpretability over standard RDD
Data basis	Simulated data with context to nFC-isPO	Data of the nFC-isPO	Data of the nFC-isPO
Sample size	N=500	N=1140	N=1417

In the second article, the methodological approach of standard RDD was applied in a subgroup analysis within the nFC-isPO study, demonstrating its practical applicability and relevance in a real-world healthcare setting.

The subgroup analysis was an additional, previously unplanned analysis, as the COVID-19 pandemic posed an additional challenge for the study personnel during the conduct of the nFC-isPO and represented a confounding factor for the interpretability of the results. Therefore, a stratified RDD analysis was performed to identify potential biases in the results due to the pandemic. However, the publication not only enables a clearer interpretation of the results, but also contributes to the visibility of the RDD approach in clinical and health research, an area in which this design has only been used to a limited extent to date.

Simultaneously, the article also points out a significant limitation of the quasi-experimental

design, as repeated measurements cannot be adequately accounted for in the standard RDD approach, meaning that time effects could still distort the results. Beyond its specific application, the publication therefore points to the methodological necessity of extending the approach to include suitable methods for longitudinal data for repeated measures, such as linear mixed models. Thus, this article highlights the research gap with regard to the RDD approach and, in the context of the dissertation, serves as a transition to the methodological extension in the third article.

Article 3 forms the core of the dissertation and contributes to closing the aforementioned research gap. It presents the methodological extension with clear, precise mathematical notation compared to the standard RDD approach. The comparative presentation of the results of both approaches shows that, on the one hand, the LMM-RDD now allows all cases with at least one follow-up measurement to be taken into account, i.e. the integration of repeated measurements in the RDD approach is enabled by the LMMs, which leads to less data loss, and that, on the other hand, temporal effects can now be calculated directly. These changes to the standard RDD demonstrate that the new approach improves statistical efficiency and adds a new level of interpretability. Furthermore, the application of the methodology in the field of health research using data from the nFC-isPO study is presented, highlighting the practical applicability of the extended method. By presenting the application and mathematical notation, including graphical representations of individual aspects of the formula, other researchers, such as biometricians, epidemiologists or clinicians can easily reproduce the methodology and apply it to their own research questions.

5.2 Relevance to the current state of research

Although the RDD has become a central approach for identifying causal effects in health research in recent years and has proven to be a solid alternative to RCTs [82–85], methodological challenges remain that limit its applicability in certain contexts. The publications and the dissertation address these challenges and contribute to the expansion and refinement of the existing RDD approach through the methodological advancement of the RDD for repeated measurements.

The application of the standard RDD approach is mainly limited to settings with cross-sectional data, where only one observation is available for each study unit [86]. Some studies also apply RDD to pre-post designs, but this application of the design essentially represents a special form of the standard RDD approach, as the pre-measurement is interpreted as the group allocation variable with the threshold value and the post-measurement as the outcome variable. The approach therefore clearly reaches its limits when it comes to time effects so that there is a risk of underestimating standard errors or interpreting effects as causal that are actually attributable to existing time trends [86, 87].
The special pre-post form of the standard RDD approach was also used in the nFC-isPO study, whereby an additional second outcome measurement was carried out over time. How-

ever, although repeated measurements of the same patient fundamentally implied a longitudinal data structure, the standard RDD initially only allowed separate analysis of the individual measurement points due to its conceptual limitations, which in fact corresponded to a cross-sectional analysis. As a result, intra-individual correlations and time effects, which often occur in longitudinal data, could not be taken into account in the assessment of the primary endpoint in isPO using the standard RDD approach [70, 88].
Since these existing limitations of the standard RDD led to methodological challenges not only in isPO but also in other studies, considerations were made in the past regarding the methodological further development of the RDD with regard to time-related influences.
A relatively new approach that establishes a link between time and RDD and is increasingly being used is the regression discontinuity in time (RDiT) approach [86, 89]. The RDiT uses time as the continuous assignment variable (running variable). It has its origins primarily in political science and econometrics, where it is used to analyse the effects of political decisions, reforms or institutional changes that are implemented at a clearly defined point in time. The central idea is that a specific event or intervention takes place at a specific point in time, and observations immediately before and after this time are compared in order to determine the causal effect of the action. A typical example is the examination of legislative changes, political interventions or regulatory measures, where the date of implementation serves as a threshold value [86]. However, the RDiT approach has also been increasingly used in the field of health research, for example to evaluate changes in digital health measures [89]. Overall, RDiT therefore follows the same methodological logic as standard RDD. It compares observations immediately before and after a threshold value, but only at a specific point in time. While this demonstrates that the use of time as an allocation variable integrates the time component into the RDD approach, it nevertheless remains an analysis of unique, abruptly dated events [86, 89].
Consequently, the fundamental limitation remains that the design cannot adequately capture complex time-related relationships and effects or intra-individual developments, as is the case with repeated measurements of other continuous variables. Therefore, the challenge remains to consistently identify the causal effect at the threshold while methodically and correctly take repeated measurements over time into account.
This methodological basis emphasises the need to further develop the RDD in order to enable consistent identification of causal effects and to uncover temporal effects in longitudinal data with repeated measurements. The newly developed LMM-RDD approach in this dissertation combines the established standard RDD with the LMM approach for repeated measurements. This allows both the discontinuity – i.e. the treatment effect – at the threshold and time effects to be taken into account and estimated simultaneously, so that LMM-RDD enables more valid estimates of the effects, thereby significantly increasing the informational value of the analysis and the reliability of the conclusions for research. Furthermore, this also leads to a more efficient use of data, as not only one individual measurement per person but also the entire longitudinal structure is included in the estimation [71]. This increases flexibility in the application of the RDD design, particularly in research areas such as medicine, epidemiology and health sciences, where repeated measurements are common.

Therefore, the profound and practical introduction of RDD into health research and the methodological extension to an LMM-RDD presented in this dissertation provide an important contribution to the further development of quasi-experimental study designs as an alternative to RCTs and to the further development of an empirical analysis method for longitudinal data within a quasi-experimental design.

5.3 Strengths, limitations and open issues

The discussion of the relevance of this dissertation in the context of the current research already provides indications of specific strengths considering the developed method. Table 5.2 provides an overview of the strengths and limitations of the RDD and the new LMM-RDD.

Table 5.2: Strenghts and Limitations of the RDD and LMM-RDD

Strenghts	***Limitations***
• Clear and intuitive graphical visualisation of the estimation • RDDs have comparatively mild methodological assumptions • LMM enables incorporation of repeated measures • LMM enables better handling of missing data • Random effects in LMMs increase flexibility and robustness	• Complex interventions hinder assessing individual component effects • Continuous assignment variable with threshold required for RDD approaches • RDD approaches require larger sample size

One strength of the standard RDD, which is also particularly noteworthy in the new LMM-RDD compared to many other statistical methods, is the clearly structured graphical representation of the estimation, which enables a clear and easily interpretable presentation of the effects [71]. Furthermore, an RDD is characterised by comparatively mild methodological assumptions, such as a defined threshold value with a discontinuous change in exposure probability, the absence of systematic manipulation of the allocation variables, the comparability of individuals just above and below the threshold value, and a continuous course of the outcome without intervention. This makes it possible to obtain valid causal estimates under less restrictive conditions compared to other study designs [23, 90]. And although the combination of an RDD and an LMM tends to lead to additional assumptions, it also offers important methodological advantages. LMMs allow for more flexible modelling of repeated measures and thus the use of more data, taking into account a random effect. Against this background, careful implementation of an LMM-RDD can be expected to lead to more robust effect estimates and higher internal validity compared to other non-randomised or quasi-experimental study designs [63, 65, 91].

A major strength of the combination of an LMM with an RDD is therefore the ability to efficiently utilise all available data. In contrast, in the standard RDD approach, the data first had to be reduced complete cases in order to establish correlations between different pre-post comparisons. The integration of LMMs therefore not only allows for consideration of repeated measurements and thus the estimation of temporal effects, but also enables better handling of missing data [65, 79].
One frequently stated limitation of the quasi-experimental RDD is the large sample size required. Therefore, it is often concluded that an RDD is more suitable for retrospective studies, as the effect estimation is restricted to a predefined bandwidth in order to maintain local randomisation [9, 92, 93]. Despite the fact that this bandwidth limitation remains in place in the new LMM-RDD approach, the use of repeated measurements means that more data is available in the bandwidth overall, thereby mitigating this limitation somewhat [68].

While the promising results provide a solid foundation, they also highlight further research questions that should be examined in more detail in future research.
One open issue concerns the underlying functional form of the proposed approach, which combines the logic of RDD with LMMs. For conceptual reasons and due to the introductory nature and development of a new method, a linear relationship was assumed in the dissertation in order to enable the integration of both methods and simplify the interpretation of the results. Nevertheless, it cannot be ruled out that the assumption of linearity does not adequately reflect the underlying relationship or even leads to bias. In the standard RDD context, which does not take repeated measures into account, higher-degree polynomials are often used to more accurately map non-linear relationships between the continuous variable and the outcome. Considering this, it would be beneficial to extend this approach to nonlinear mixed models in future work in order to better capture more complex nonlinear relationships within the RDD in conjunction with longitudinal data and repeated measurements, thereby further increasing the informative value and flexibility of the new approach.
Moreover, an LMM for repeated measures with a random intercept component was specified to demonstrate the fundamental performance of the approach under controlled, simplified conditions. In further analyses, the consideration of random slopes could be a useful extension to even better model potential interindividual differences in effects or time-dependent changes within individuals. This could also contribute to a more differentiated representation of the variance structure of the data and thus further increase the accuracy of the estimates and the inferential statistical significance compared to the standard RDD approach. Simultaneously, however, the current model structure would remain straightforward to interpret and would be suitable for medical research questions involving complex data structures.
A further consideration arises from the approach pursued within the framework of the RDD. As mentioned in the methodological introduction at the beginning, the literature often distinguishes between sharp and fuzzy regression discontinuity designs, with the latter allowing for probabilistic assignment to the intervention or control group. The focus of this work was on the sharp approach based on nFC-isPO with its unique group assignment. Nevertheless, future work could investigate extending the new LMM-RDD to a fuzzy design in order to

further expand its scope of application. Such an adaptation could further increase the flexibility of the method.
Finally, it should be emphasised that the method presented here should be understood as the first systematic introduction of the combined LMM-RDD approach. For a more comprehensive evaluation, it is recommended to benchmark the performance of LMM-RDD against standard RDD under controlled conditions. Different sample sizes, different assumptions for handling missing values (e.g., missing completely at random, missing at random, or missing not at random), and different effect sizes should be considered to ensure the robustness of the estimates and the sensitivity of the method. Such analyses would further substantiate the methodological validity of the approach and assess its applicability to clinical studies with repeated measurements and complex data structures.

Beyond all this, both the aggregation of multidimensional allocation criteria and an extension to the field of multivariate regression would also be intriguing approaches. Both approaches can be described as 'combined measures', but refer to different levels of the analysis model – the level of allocation to intervention and the outcome level with simultaneous modelling of multiple endpoints. At the allocation level, the approach of combined measures refers to the conventional single-cutoff design of the RDD, which is extended to a multidimensional treatment frontier, as treatment eligibility depends on several independent criteria [94, 95]. In healthcare settings, access to an intervention, for example, may be linked to a certain symptom severity and additional risk factors or care characteristics. Instead of a clear single threshold, this results in several criteria (a set of boundaries) that together determine the allocation of treatment, so that patients may be close to the threshold for treatment eligibility in terms of several decision-making factors at the same time. In other words, the transition between untreated and treated patients in this case does not represent a single threshold, but rather a range in which very similar clinical profiles exist on both sides of the treatment decision. In the nFC-isPO, assignment to intervention was primarily based on a cut-off value for the HADS total score (HADS-T ≥ 15). Initially, this dissertation project also considered using the subscales for anxiety (HADS-A) and depression (HADS-D) as allocation criteria, so that eligibility for the intervention would have been based not solely on the HADS-T but on a differentiated clinical stress profile. Conceptually, such an approach would correspond to an RDD with multiple assignment variables, in which the relevant jumps at all cut-off points - HADS-A and / or HADS-D ≥ 7 - are used together to identify the causal effect. This would have shifted the decision boundary from a one-dimensional threshold into a multidimensional space in which patients who are just above or below the treatment threshold have very similar clinical profiles.

To analyse the effects in the context of this complex assignment structure, the various assignment criteria are often combined into a one-dimensional composite measure using a binding score or centring approach. This measure describes a patient's distance from the nearest treatment boundary and serves as a new running variable for the RDD. On this basis, local average treatment effects can be estimated for patients who differ only minimally in clinical

terms but have been treated differently, while maintaining the central assumption of continuity of potential outcomes near the boundary.
However, within the nfC-isPO framework, the inclusion of the HADS subscales in the classification logic appeared methodologically unproductive, as the subscales are strongly correlated components of the total score and do not represent independent clinical decision-making dimensions. An additional combination would have represented the same clinical condition multiple times without offering any significant added value.

In parallel to this, the second approach of combined measurements reflects the other consideration of the dissertation project, namely the extension to the field of multivariate regression. The aim here was to examine whether the analysis of the nFC-isPO data would require an extension of the RDD for repeated measurements or, alternatively, a multivariate analysis of combined outcome measures within the RDD. Initially, modelling of the two subscales (HADS-A and HADS-D) as separate endpoints was considered in addition to the HADS-T score. Within this context, the question emerged as to whether an analysis at the level of a combined measure in the multivariate sense would be useful.
However, such an approach of combined measure does not correspond to the mere summation or parallel consideration of the subscales, but rather refers in multivariate statistics to a linear combination of correlated outcomes derived from the statistical model itself. Such data-based linear combinations – as generated in multivariate regression approaches or multivariate analysis of variance – serve to optimally map common variance structures, examine effects on a synthetic, mathematically weighted target variable, and thereby uncover overarching stress constructs or clinically relevant symptom clusters [96, 97]. The HADS subscales, on the other hand, are construct-bound, psychometrically defined and clinically distinct measurement domains (anxiety and depression) with fixed item weights and therefore do not represent a combined target variable or a combined result in the multivariate mathematical sense.
This fact, as well as the longitudinal structure of the nFC-isPO data, the growing interest in temporal effects during the project progress, and the identified research gap regarding the inclusion of repeated measures in the standard RDD, motivated this dissertation to extend the RDD approach for repeated measures.

5.4 Conclusion

This dissertation contributes to the methodological advancement of the quasi-experimental RDD in health research, particularly in relation to complex data structures. To this end, the theoretical foundations, empirical application with data from practice, i.e. the nFC-isPO, and the extension for LMMs for repeated measures of the RDD were systematically examined.
The results show that, with appropriate modelling, RDD can be a robust alternative to randomised trials when ethical, logistical or practical constraints preclude randomisation. Furthermore, it was established that the standard RDD approach requires methodological adjustments for longitudinal data and repeated measures in order to adequately account for

time effects. Based on these findings, this dissertation extends the RDD by combining it with LMMs for repeated measures. By combining a robust causal inference framework with random effects structures, the proposed method increases the precision and interpretability of estimated effects in longitudinal data settings and enables a more nuanced investigation of time-dependent treatment effects. The detailed description of the mathematical notation and implementation steps ensures that the new method is easy to replicate and facilitates its transfer into applied medical research. This work therefore not only makes a significant contribution to the further development of medical statistics, but also provides a practical set of tools that will help to achieve reliable, evidence-based results in clinical trials in the long term.

Overall, this work contributes to the further development of alternative study designs, including their analysis methods, and opens up new perspectives for their application in clinical research, particularly in dealing with repeated measures.

References

[1] R. Westermann, "Experiment im Dorsch Lexikon der Psychologie," 2022. [Online: https://dorsch.hogrefe.com/stichwort/experiment; Accessed: 15.05.2025].

[2] J. Bortz and N. Döring, *Forschungsmethoden und Evaluation: für Human- und Sozialwissenschaftler, 4. Auflage*. Springer-Lehrbuch, Springer Berlin Heidelberg, 2013.

[3] T. D. Cook, D. T. Campbell, and W. Shadish, *Experimental and quasi-experimental designs for generalized causal inference*. Houghton Mifflin Boston, 2 ed., 2002.

[4] D. B. Rubin, "Estimating causal effects of treatments in randomized and nonrandomized studies.," *Journal of educational Psychology*, vol. 66, no. 5, p. 688, 1974.

[5] D. Dang, S. Dearholt, K. Bissett, J. Ascenzi, and M. Whalen, *Johns Hopkins Evidence-Based Practice for Nurses and Healthcare Professionals: Model and Guidelines, Fourth Edition*. Sigma Theta Tau International, 2021.

[6] M. D. Cattaneo, L. Keele, and R. Titiunik, "A guide to regression discontinuity designs in medical applications," *Statistics in Medicine*, vol. 42, no. 24, pp. 4484–4513, 2023.

[7] S. F. Jackson, N. Fazal, and N. Giesbrecht, "A hierarchy of evidence: Which intervention has the strongest evidence of effectiveness?." PDF available via ResearchGate, Accessed: 01/2026, 2010. Previously posted on the Canadian Best Practices Portal for Health Promotion and Chronic Disease Prevention.

[8] R. McDermott, "Encyclopedia of social measurement: Experiments, political science," in *Encyclopedia of Social Measurement* (K. Kempf-Leonard, ed.), vol. 1, pp. 901–909, San Diego, CA: Elsevier Science, 2005.

[9] M. L. Maciejewski, "Quasi-experimental design," *Biostatistics & Epidemiology*, vol. 4, no. 1, pp. 38–47, 2020.

[10] H. White and S. Sabarwal, *Quasi-Experimental Design and Methods*. Better Evaluation / UNICEF Office of Research – Innocenti & International Initiative for Impact Evaluation, 2014. Methodological brief on quasi-experimental design and methods in impact evaluation.

[11] M. A. Wirtz, "Quasi-Experiment im Dorsch Lexikon der Psychologie," 2022. [Online: https://dorsch.hogrefe.com/stichwort/quasi-experiment; Accessed: 15.05.2025].

[12] R. A. Fisher, *The Design of Experiments*. New York: Hafner Publishing Company, 9 ed., 1971.

[13] A. Coly and G. Parry, "Evaluating complex health interventions: A guide to rigorous research designs." Online resource, June 26 2017. Available from AcademyHealth, https://academyhealth.org/evaluationguide.

[14] J. D. Marion and A. D. Althouse, "The use of historical controls in clinical trials," *JAMA*, vol. 330, no. 15, pp. 1484–1485, 2023.

[15] O. Kuss, M. Blettner, and J. Börgermann, "Propensity Score – eine alternative Methode zur Analyse von Therapieeffekten," *Deutsches Ärzteblatt International*, vol. 113, no. 35–36, pp. 597–603, 2016.

[16] D. Kabata, Y. Gon, and A. Shintani, "Matching on propensity and prognostic scores can lead to different estimates of heterogeneous treatment effects: a case study and simulation," *Health Services and Outcomes Research Methodology*, vol. 24, no. 2, pp. 227–238, 2024.

[17] D. Bottigliengo, I. Baldi, C. Lanera, G. Lorenzoni, J. Bejko, T. Bottio, V. Tarzia, M. Carrozzini, G. Gerosa, P. Berchialla, and D. Gregori, "Oversampling and replacement strategies in propensity score matching: a critical review focused on small sample size in clinical settings," *BMC Medical Research Methodology*, vol. 21, p. 256, 2021.

[18] M. Bertrand, E. Duflo, and S. Mullainathan, "How much should we trust differences-in-differences estimates?*," *The Quarterly Journal of Economics*, vol. 119, pp. 249–275, 02 2004.

[19] C. Wing, K. Simon, and R. A. Bello-Gomez, "Designing difference in difference studies: Best practices for public health policy research," *Annual Review of Public Health*, vol. 39, no. 1, pp. 453–469, 2018.

[20] A. Bilinski and I. Ganguli, "Methodological considerations for difference-in-differences," *JAMA Internal Medicine*, vol. 184, no. 10, pp. 1161–1162, 2024.

[21] C. A. Hosman, B. B. Hansen, and P. W. Holland, "The sensitivity of linear regression coefficients' confidence limits to the omission of a confounder," *The Annals of Applied Statistics*, vol. 4, no. 2, pp. 849–870, 2010.

[22] N. A. Schuster, J. J. Rijnhart, L. C. Bosman, J. W. Twisk, T. Klausch, and M. W. Heymans, "Misspecification of confounder-exposure and confounder-outcome associations leads to bias in effect estimates," *BMC Medical Research Methodology*, vol. 23, no. 1, p. 11, 2023.

[23] J. Bor, E. Moscoe, P. Mutevedzi, M.-L. Newell, and T. Bärnighausen, "Regression discontinuity designs in epidemiology: causal inference without randomized trials," *Epidemiology*, vol. 25, no. 5, pp. 729–737, 2014.

[24] E. Moscoe, J. Bor, and T. Bärnighausen, "Regression discontinuity designs are underutilized in medicine, epidemiology, and public health: a review of current and best practice," *Journal of clinical epidemiology*, vol. 68, no. 2, pp. 132–143, 2015.

[25] A. S. Venkataramani, J. Bor, and A. B. Jena, "Regression discontinuity designs in healthcare research," *BMJ*, vol. 352, 2016.

[26] T. D. Cook, "'Waiting for life to arrive': a history of the regression-discontinuity design in psychology, statistics and economics," *Journal of Econometrics*, vol. 142, no. 2, pp. 636–654, 2008.

[27] M. Villamizar-Villegas, F. A. Pinzon-Puerto, and M. A. Ruiz-Sanchez, "A comprehensive history of regression discontinuity designs: An empirical survey of the last 60 years," *Journal of Economic Surveys*, vol. 36, no. 4, pp. 1130–1178, 2022.

[28] D. L. Thistlethwaite and D. T. Campbell, "Regression-discontinuity analysis: An alternative to the ex post facto experiment.," *Journal of Educational psychology*, vol. 51, no. 6, p. 309, 1960.

[29] D. T. Campbell, "Experimental and quasi-experimental designs for research on teaching," *Handbook of research on teaching*, vol. 5, pp. 171–246, 1963.

[30] J. A. Sween, *The experimental regression design: An inquiry into the feasibility of non-random treatment allocation*. PhD thesis, Northwestern University, Evanston, IL, 1972.

[31] D. T. Campbell and T. D. Cook, "Quasi-experimentation," *Chicago, IL: Rand Mc-Nally*, vol. 1, no. 1, pp. 1–384, 1979.

[32] W. M. K. Trochim, *Research design for program evaluation: The regression-discontinuity approach*, vol. 6 of *Contemporary Evaluation Research : a series of books on applied social science*. Beverly Hills, CA: SAGE Publications, 1984.

[33] M. O. Finkelstein, B. Levin, and H. Robbins, "Clinical and prophylactic trials with assured new treatment for those at greater risk: I. a design proposal.," *American Journal of Public Health*, vol. 86, no. 5, pp. 691–695, 1996.

[34] M. O. Finkelstein, B. Levin, and H. Robbins, "Clinical and prophylactic trials with assured new treatment for those at greater risk: II. examples.," *American Journal of Public Health*, vol. 86, no. 5, pp. 696–705, 1996.

[35] J. D. Angrist and G. W. Imbens, "Identification and estimation of local average treatment effects," Tech. Rep. 118, National Bureau of Economic Research, Cambridge, MA, February 1995.

[36] J. D. Angrist, G. W. Imbens, and D. B. Rubin, "Identification of causal effects using instrumental variables," *Journal of the American statistical Association*, vol. 91, no. 434, pp. 444–455, 1996.

[37] J. Hahn, P. Todd, and W. Van der Klaauw, "Identification and estimation of treatment effects with a regression-discontinuity design," *Econometrica*, vol. 69, no. 1, pp. 201–209, 2001.

[38] D. S. Lee, *The Electoral Advantage to Incumbency and Voters' Valuation of Politicians' Experience: A Regression Discontinuity Analysis of Elections to the U.S. House*. Cambridge, MA, USA, 2001.

[39] G. W. Imbens and T. Lemieux, "Regression discontinuity designs: A guide to practice," *Journal of econometrics*, vol. 142, no. 2, pp. 615–635, 2008.

[40] M. Stigler and B. Quast, "Package 'rddtools'," 2022. R package manual.

[41] S. Calonico, M. D. Cattaneo, and M. H. Farrell, "Package 'rdrobust'," 2023. R package manual.

[42] D. Dimmery, *Package 'rdd'*, 2016. R package manual.

[43] R Core Team, *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria, 2024.

[44] M. H. Boon, P. Craig, H. Thomson, M. Campbell, and L. Moore, "Regression discontinuity designs in health: a systematic review," *Epidemiology*, vol. 32, no. 1, pp. 87–93, 2021.

[45] C. Ronckers, C. Spix, C. Trübenbach, A. Katalinic, M. Christ, A. Cicero, J. Folkerts, J. Hansmann, K. Kranzhöfer, B. Kunz, K. Manegold, U. Meyer zum Büschenfelde, A. Penzkofer, G. Vollmer, S. Weg-Remers, B. Barnes, N. Buttmann-Schweiger, S. Dahm, M. Franke, I. Schönfeld, K. Kraywinkel, and A. Wienecke, *Krebs in Deutschland für 2019/2020*. Berlin, Germany: Robert Koch-Institut, 2023.

[46] Statistisches Bundesamt, "Destatis – Statistisches Bundesamt: Todesursachen," 2025. `https://www.destatis.de/DE/Themen/Gesellschaft-Umwelt/Gesundheit/Todesursachen/_inhalt.html#235880`, Accessed: 03.04.2025.

[47] A. Künzler, A. Zindel, H. Znoj, and M. Bargetzi, "Psychische Belastungen bei Krebspatienten und ihren Partnern im ersten Jahr nach Diagnosestellung.," *Praxis (16618157)*, vol. 99, no. 10, 2010.

[48] M. Kusch, H. Labouvie, V. Schiewer, N. Talalaev, J. C. Cwik, S. Bussmann, L. Vaganian, A. L. Gerlach, A. Dresen, N. Cecon, S. Salm, T. Krieger, H. Pfaff, C. Lemmen, L. Derendorf, S. Stock, C. Samel, **A. Hagemeier**, M. Hellmich, B. Leicher, G. Hultenschmidt, J. Swoboda, P. Haas, A. Arning, A. Gottel, K. Schwickerath, U. Graeven, S. Houwaart, H. Kerek-Bodden, S. Krebs, C. Muth, C. Hecker, M. Reiser, C. Mauch, J. Benner, G. Schmidt, C. Karlowsky, G. Vimalanandan, L. Matyschik, L. Galonska, A. Francke, K. Osborne, U. Nestle, M. Baumer, K. Schmitz, J. Wolf, and M. Hallek, "Integrated, cross-sectoral psycho-oncology (isPO): a new form of care for newly diagnosed cancer patients in germany," *BMC Health Services Research*, vol. 22, no. 1, p. 543, 2022.

[49] A. Mehnert and S. Koranyi, "Psychoonkologische Versorgung: eine Herausforderung," *DMW-Deutsche Medizinische Wochenschrift*, vol. 143, no. 05, pp. 316–323, 2018.

[50] D. Seitschek, N. Sulz-Lehar, and C. Sebesta, "Psychoonkologie–Bestandteil einer umfassenden onkologischen Behandlung," *Wiener Medizinische Wochenschrift*, vol. 173, no. 9, pp. 206–208, 2023.

[51] P. Herschbach, "Psychoonkologie: Anspruch für die nächsten Jahre," *Der Onkologe*, vol. 25, no. S1, pp. 96–101, 2019.

[52] Arbeitsgemeinschaft für Psychoonkologie in der Deutschen Krebsgesellschaft e. V., "PSO - Arbeitsgemeinschaft für Psychoonkologie," 2025. `https://www.pso-ag.org/struktur-und-organisation/geschichte-der-pso`, Accessed: 25.04.2025.

[53] Weiterbildung Psychosoziale Onkologie e. V., "WPO - Weiterbildung Psychosoziale Onkologie," 2025. `https://www.wpo-ev.de/ueber-uns/geschichte-der-wpo/`, Accessed: 25.04.2025.

[54] J. Weis, G. Blettner, A. Lehmann-Laue, A. Ihrig, and S. von Saalfeld, "Professionalisierung in der Psychoonkologie–die Bedeutung der Fort-und Weiterbildung," *Die Onkologie*, vol. 29, no. 10, pp. 905–912, 2023.

[55] Bundesministerium für Gesundheit, "Nationaler Krebsplan: Handlungsfelder, Ziele, Umsetzungsempfehlungen und Ergebnisse," 2017.

[56] Arbeitsgemeinschaft der Wissenschaftlichen Medizinischen Fachgesellschaften (AWMF), "Awmf-leitlinienportal," 2025. `https://www.awmf.org/leitlinien`, Accessed: 26.04.2025.

[57] Arbeitsgemeinschaft der Wissenschaftlichen Medizinischen Fachgesellschaften e.V., "S3-Leitlinie Psychoonkologische Diagnostik, Beratung und Behandlung von erwachsenen Krebspatient*innen," 2023. `https://register.awmf.org/de/leitlinien/detail/032-051OL`, AWMF-Registernummer: 032-051OL. Accessed: 26.04.2025.

[58] I. Jenniches, C. Lemmen, J. C. Cwik, M. Kusch, H. Labouvie, N. Scholten, A. Gerlach, S. Stock, C. Samel, **A. Hagemeier**, , M. Hellmich, P. Haas, M. Hallek, H. Pfaff, and A. Dresen, "Evaluation of a complex integrated, cross-sectoral psycho-oncological care program (isPO): a mixed-methods study protocol," *BMJ Open*, vol. 10, no. 3, p. e034141, 2020.

[59] C. Hermann-Lingen, U. Buss, and R. P. Snaith, *Hospital Anxiety and Depression Scale–German Version (HADS-D)*. Göttingen, Germany: Hogrefe, 2011.

[60] S. Bussmann, L. Vaganian, A. L. Gerlach, M. Kusch, H. Labouvie, and J. C. Cwik, "Screening and assessment for cancer patients' psychosocial support needs: Development and validation of the psychosocial risk questionnaire," *European Journal of Oncology Nursing*, vol. 63, p. 102240, 2023.

[61] M. Kusch, H. Labouvie, V. Schiewer, N. Talalaev, J. C. Cwik, S. Bussmann, L. Vaganian, A. L. Gerlach, A. Dresen, N. Cecon, S. Salm, T. Krieger, H. Pfaff, C. Lemmen,

L. Derendorf, S. Stock, C. Samel, **A. Hagemeier**, M. Hellmich, B. Leicher, G. Hultenschmidt, J. Swoboda, P. Haas, A. Arning, A. Gottel, K. Schwickerath, U. Graeven, S. Houwaart, H. Kerek-Bodden, S. Krebs, C. Muth, C. Hecker, M. Reiser, C. Mauch, J. Benner, G. Schmidt, C. Karlowsky, G. Vimalanandan, L. Matyschik, L. Galonska, A. Francke, K. Osborne, U. Nestle, M. Baumer, K. Schmitz, J. Wolf, and M. Hallek, "Integrated, cross-sectoral psycho-oncology (isPO): a new form of care for newly diagnosed cancer patients in germany," *BMC Health Serv Res*, vol. 22, no. 1, p. 543, 2022.

[62] T. J. Cleophas, A. H. Zwinderman, and B. van Ouwerkerk, "Clinical research: a novel approach to the analysis of repeated measures," *American journal of therapeutics*, vol. 19, no. 1, pp. e1–e7, 2012.

[63] M. A. Detry and Y. Ma, "Analyzing repeated measurements using mixed models," *JAMA*, vol. 315, no. 4, pp. 407–408, 2016.

[64] B. T. West, K. B. Welch, and A. T. Galecki, *Linear Mixed Models: A Practical Guide Using Statistical Software*. Boca Raton, FL: Chapman and Hall/CRC, 3 ed., 2022.

[65] V. A. Brown, "An introduction to linear mixed-effects modeling in R," *Advances in Methods and Practices in Psychological Science*, vol. 4, no. 1, 2021.

[66] R. A. Armstrong, "Recommendations for analysis of repeated-measures designs: testing and correcting for sphericity and use of manova and mixed model analysis," *Ophthalmic and Physiological Optics*, vol. 37, no. 5, pp. 585–593, 2017.

[67] L. N. Muhammad, "Guidelines for repeated measures statistical analysis approaches with basic science research considerations," *The Journal of Clinical Investigation*, vol. 133, no. 11, p. e171058, 2023.

[68] M. B. de Melo, D. Daldegan-Bueno, M. G. Menezes Oliveira, and A. L. de Souza, "Beyond anova and manova for repeated measures: Advantages of generalized estimated equations and generalized linear mixed models and its use in neuroscience research," *European Journal of Neuroscience*, vol. 56, no. 12, pp. 6089–6098, 2022.

[69] **A. Hagemeier**, C. Samel, and M. Hellmich, "The regression discontinuity design: methods and implementation with a worked example in health services research," *Zeitschrift für Evidenz, Fortbildung und Qualität im Gesundheitswesen*, vol. 172, pp. 71–77, 2022.

[70] **A. Hagemeier**, A. Adams, T. Krieger, S. Salm, N. Cecon-Stabel, A. Dresen, and M. Hellmich, "The impact of covid-19 on the interpretation of psycho-oncological support trial results: a quasi-experimental approach using the data from the new form of care "integrated cross-sectoral psycho-oncology (nFC-isPO)"," *BMC Health Services Research*, vol. 23, no. 1, p. 556, 2023.

[71] **A. Hagemeier**, K. D. Rosenberger, and M. Hellmich, "Extension of regression discontinuity designs (RDD) to evaluate time effects – a quasi-experimental approach using the data from the new form of care "integrated cross-sectoral psycho-oncology (nFC- isPO)"," *PLOS ONE*, vol. 21, no. 2, pp. 1–11, 2026.

[72] A. Schneider, G. Hommel, and M. Blettner, "Linear regression analysis: part 14 of a series on evaluation of scientific publications," *Deutsches Ärzteblatt International*, vol. 107, no. 44, p. 776, 2010.

[73] A. Gałecki and T. Burzykowski, *Linear Mixed-Effects Models Using R: A Step-by-Step Approach*. Springer Texts in Statistics, Springer New York, 1 ed., 2013.

[74] L. Fahrmeir, T. Kneib, and S. Lang, *Regression: Modelle, Methoden und Anwendungen*. Statistik und ihre Anwendungen, Springer, Berlin, Heidelberg, 2 ed., 2009.

[75] W. T. Hoyt, S. Leierer, and M. J. Millington, "Analysis and interpretation of findings using multiple regression techniques," *Rehabilitation Counseling Bulletin*, vol. 49, no. 4, pp. 223–233, 2006.

[76] R. Bender, A. Ziegler, and S. Lange, "Multiple regression," *DMW-Deutsche Medizinische Wochenschrift*, vol. 132, no. S 01, pp. e30–e32, 2007.

[77] M. J. Monsalves, A. S. Bangdiwala, A. Thabane, and S. I. Bangdiwala, "Level (logical explanations & visualizations of estimates in linear mixed models): recommendations for reporting multilevel data and analyses," *BMC Medical Research Methodology*, vol. 20, no. 1, p. 3, 2020.

[78] T. Haumann, R. Kassemeier, and J. Wieseke, "Multilevel modeling," in *Handbook of Market Research* (C. Homburg, M. Klarmann, and A. Vomberg, eds.), pp. 1–41, Cham, Switzerland: Springer, 2021.

[79] L. Ansmann, K. Kuhr, and C. Kowalski, "Mehrebenenanalysen in der organisationsbezogenen Versorgungsforschung–Nutzen, Voraussetzungen und Durchführung," *Das Gesundheitswesen*, vol. 79, no. 03, pp. 203–209, 2017.

[80] H. Singmann and D. Kellen, "An introduction to mixed models for experimental psychology," in *New Methods in Cognitive Psychology* (D. Spieler and E. Schumacher, eds.), Frontiers of Cognitive Psychology, pp. 4–31, New York, NY: Routledge, 2019.

[81] D. Bates, M. Mächler, B. Bolker, and S. Walker, "Fitting linear mixed-effects models using lme4," *Journal of Statistical Software*, vol. 67, no. 1, pp. 1–48, 2015.

[82] I. L. Maas, S. Nolte, O. B. Walter, T. Berger, M. Hautzinger, F. Hohagen, W. Lutz, B. Meyer, J. Schröder, C. Späth, J. P. Klein, S. Moritz, and M. Rose, "The regression discontinuity design showed to be a valid alternative to a randomized controlled trial for estimating treatment effects," *Journal of Clinical Epidemiology*, vol. 82, pp. 94–102, 2017.

[83] N. van Leeuwen, H. F. Lingsma, A. J. de Craen, D. Nieboer, S. P. Mooijaart, E. Richard, and E. W. Steyerberg, "Regression discontinuity design: simulation and application in two cardiovascular trials with continuous outcomes," *Epidemiology*, vol. 27, no. 4, pp. 503–511, 2016.

[84] A. G. O'Keeffe, S. Geneletti, G. Baio, L. D. Sharples, I. Nazareth, and I. Petersen, "Regression discontinuity designs: an approach to the evaluation of treatment efficacy in primary care using observational data," *BMJ*, vol. 349, 2014.

[85] D. Henry, P. Tolan, D. Gorman-Smith, and M. Schoeny, "Alternatives to randomized control trial designs for community-based prevention evaluation," *Prevention Science*, vol. 18, no. 6, pp. 671–680, 2017.

[86] C. Hausman and D. S. Rapson, "Regression discontinuity in time: Considerations for empirical applications," *Annual Review of Resource Economics*, vol. 10, no. 1, pp. 533–552, 2018.

[87] I. H. Zuckerman, E. Lee, A. K. Wutoh, Z. Xue, and B. Stuart, "Application of regression-discontinuity analysis in pharmaceutical health services research," *Health Services Research*, vol. 41, no. 2, pp. 550–563, 2006.

[88] T. Krieger, N. Cecon, S. Salm, P. Holger, D. Simic, C. Lemmen, A. Shukri, S. Stock, L. Derendorf, **A. Hagemeier**, A. Adams, M. Hellmich, A. Dresen, and for the isPO consortium, "Evaluation report on integrated cross-sector psycho-oncology (isPO)," evaluation report, Institute of Medical Sociology, Health Services Research and Rehabilitation Science, University of Cologne, 2022. Funded by the Innovation Committee of the Joint Federal Committee (GBA), Funding code: 01NVF17022, Language of the report: German.

[89] I. Thapa, P.-A. Laforcade, F. K. Bishop, J. Ferstad, M. Desai, D. M. Maahs, P. Prahalad, D. Zaharieva, D. Scheinker, and R. Johari, "Regression discontinuity in time: Evaluating the impact of evolving digital health interventions," *International Journal of Medical Informatics*, p. 106050, 2025.

[90] L. M. Smith, L. E. Lévesque, J. S. Kaufman, and E. C. Strumpf, "Strategies for evaluating the assumptions of the regression discontinuity design: a case study using a human papillomavirus vaccination programme," *International Journal of Epidemiology*, vol. 46, no. 3, pp. 939–949, 2017.

[91] D. D. Chaplin, T. D. Cook, J. Zurovac, J. S. Coopersmith, M. M. Finucane, L. N. Vollmer, and R. E. Morris, "The internal and external validity of the regression discontinuity design: A meta-analysis of 15 within-study comparisons," *Journal of Policy Analysis and Management*, vol. 37, no. 2, pp. 403–429, 2018.

[92] R. Jacob, P. Zhu, M.-A. Somers, and H. Bloom, "A practical guide to regression discontinuity," tech. rep., MDRC, New York, NY, 2012.

[93] M. D. Cattaneo, N. Idrobo, and R. Titiunik, *A Practical Introduction to Regression Discontinuity Designs: Foundations*. Elements in Quantitative and Computational Methods for the Social Sciences, Cambridge University Press, 2020.

[94] V. C. Wong, P. M. Steiner, and T. D. Cook, "Analyzing regression-discontinuity designs with multiple assignment variables: A comparative study of four estimation methods," *Journal of Educational and Behavioral Statistics*, vol. 38, no. 2, pp. 107–141, 2013.

[95] S. F. Reardon and J. P. Robinson, "Regression discontinuity designs with multiple rating-score variables," *Journal of research on Educational Effectiveness*, vol. 5, no. 1, pp. 83–104, 2012.

[96] I. K. Olomoda, M. Y. Olatunji, K. B. Ishola, A. K. Sanda, and A. O. Abubakar, "The effect of multivariate regression over multiple linear regression models using non-normal data," *Elicit Journal of Science Research*, vol. 1, no. 1, p. 28–39, 2025.

[97] J. F. Hair, W. C. Black, B. J. Babin, and R. E. Anderson, *Multivariate Data Analysis*. Cengage Learning EMEA, 8 ed., 2019.

Declaration of contribution

Anna Hagemeier wrote all publications independently and was significantly involved in the conception, realisation, data analysis and interpretation. Below you will find a brief description of which other co-authors were involved in the individual aspects of the publications.

1. Article:

The regression discontinuity design: Methods and implementation with a worked example in health services research

Authors: Anna Hagemeier; Christina Samel; Martin Hellmich

Affiliation of all authors: Institute of Medical Statistics and Computational Biology, Medical Faculty, University of Cologne, University Hospital Cologne, Cologne, Germany

Journal: Zeitschrift für Evidenz, Fortbildung und Qualität im Gesundheitswesen

Impact factor at time of publication: 1.1 (2022)

Citation:
Hagemeier, Anna, Christina Samel, and Martin Hellmich. The regression discontinuity design: methods and implementation with a worked example in health services research. Zeitschrift für Evidenz, Fortbildung und Qualität im Gesundheitswesen 172 (2022): 71-77. https://doi.org/10.1016/j.zefq.2022.04.014

Contributions:

Anna Hagemeier: Conceptualization,Writing- Original draft preparation, Visualization, Writing – review & editing, Methodology.
Christina Samel: Conceptualization, Writing – review & editing, Methodology.
Martin Hellmich: Supervision, Conceptualization, Writing – review & editing, Methodology.

2. Article:

The impact of COVID-19 on the interpretation of psycho-oncological support trial results: a quasi-experimental approach using the data from the new form of care "Integrated crosssectoral psycho-oncology (nFC-isPO)"

Authors: Anna Hagemeier[1], Anne Adams[1], Theresia Krieger[2], Sandra Salm[2,3], Natalia Cecon-Stabel[2], Antje Dresen[2] and Martin Hellmich[1]

Affiliations: [1] Institute of Medical Statistics and Computational Biology (IMSB), Medical Faculty, University of Cologne, University Hospital, Cologne, Germany
[2] Institute for Medical Sociology, Health Service Research and Rehabilitation Science (IMVR), Medical Faculty, Human Science Faculty, University of Cologne, University Hospital, Cologne, Germany
[3] Goethe University Frankfurt, Institute of General Practice, Frankfurt, Germany

Journal: BMC Health Services Research

Impact factor at time of publication: 2.8 (2022-2023)

Citation:
Hagemeier, A., Adams, A., Krieger, T., Salm, S., Cecon-Stabel, N., Dresen, A., & Hellmich, M. (2023). The impact of COVID-19 on the interpretation of psycho-oncological support trial results: a quasi-experimental approach using the data from the new form of care "Integrated cross-sectoral psycho-oncology (nFC-isPO)". BMC Health Services Research, 23(1), 1-8. https://doi.org/10.1186/s12913-023-09544-y

Contributions:

Conceptualization Anna Hagemeier and Martin Hellmich; **Data preparation and analysis** Anna Hagemeier; **Visualization** Anna Hagemeier; **Methodology** Anna Hagemeier, Anne Adams and Martin Hellmich; **Interpretation** Anna Hagemeier, Anne Adams; **Supervision** Martin Hellmich; **Writing - original draft** Anna Hagemeier; **Writing - review & editing** Anne Adams, Sandra Salm, Natalia Cecon-Stabel, Antje Dresen, Theresia Krieger and Martin Hellmich

3. Article:

Extension of Regression Discontinuity Designs (RDD) to evaluate time effects – a quasi-experimental approach using the data from the new form of care "Integrated cross-sectoral psycho-oncology (nFC- isPO)"

Authors: Anna Hagemeier[1], Kerstin Rosenberger[1] and Martin Hellmich[1,2]

Affiliations: [1] Institute of Medical Statistics and Computational Biology, Medical Faculty, University of Cologne, University Hospital Cologne, Cologne, Germany
[2] Department of Medical Statistics, University Medical Center Göttingen, Göttingen, Germany

Journal: PLOS ONE

Impact factor at time of publication: 2.6 (2024)

Citation:
Hagemeier A, Rosenberger KD, Hellmich M (2026) Extension of Regression Discontinuity Designs (RDD) to evaluate time effects – A quasi-experimental approach using the data from the new form of care "Integrated cross-sectoral psycho-oncology (nFC- isPO)". PLOS ONE 21(2): e0343414. https://doi.org/10.1371/journal.pone.0343414

Contributions:

Conceptualization Anna Hagemeier and Martin Hellmich; **Data preparation and analysis** Anna Hagemeier; **Visualization** Anna Hagemeier; **Methodology** Anna Hagemeier, Kerstin Rosenberger and Martin Hellmich; **Interpretation** Anna Hagemeier, Kerstin Rosenberger; **Supervision** Martin Hellmich; **Writing - original draft** Anna Hagemeier; **Writing - review & editing** Kerstin Rosenberger and Martin Hellmich

Acknowledgement

Thanks to my supervisor Prof. Dr. Martin Hellmich, who continually amazed me with his scientific expertise because he always had solutions to seemingly insoluble problems.

In addition to Martin Hellmich, I would also like to thank my two tutors, Prof. Dr. Nicole Skoetz and Prof. Dr. Elke Kalbe. They always encouraged me to be more courageous, to submit my publications to more prestigious journals and not to give up when faced with setbacks.

A very special thank you, if not the biggest, goes to Dr. Kerstin Rosenberger. Kerstin, you have always given me so much valuable advice in terms of content and personnel and have continually improved my texts. In addition, you have always cheered me up and encouraged me when the progress of my work or the next journal publication was once again delayed.

An equally big thank you goes to Wiebke Müller. Wiebke, just like Kerstin, you always motivated me with your positive attitude, not only in the office, but also during our writing sessions on various weekends. You also always took excellent care of our physical well-being!

Another big thanks goes to Julia Frank. You always read my texts with an open mind and a critical eye and gave me a different perspective in lots of parts, which has greatly benefited my work. Julia, thank you also for the Zoom calls at short notice, which sometimes involved supposedly small things like colour schemes, but where I would have gone crazy without your keen eye.

Of course, I mustn't forget my colleague Anne Adams, who has been in the same boat with me from the very beginning. Anne, we have not only shared an office for so long, but also supported each other emotionally. You always had an open ear for quick questions during the working day and often spent a long time thinking with me about one statistical problem or another. Thank you for that!

I would also like to thank Stefan Paulus-Kötter for his linguistic corrections, because even the best content is useless if it doesn't sound good linguistically.

Finally, I would like to express my personal gratitude to my family and, above all, to my partner Stephan Peiffer. Steph, you have been with me through all the ups and downs, always supporting me, always telling me not to give up, and showing me how far I have come!

List of further Publications

[Pub1] **A. Hagemeier**, M. Oberste, K. D. Rosenberger, R. Roth, M. Hellmich, V. Fluegel, K. Ruettger, K. Dadaczynski, C. Joisten, L. Mause, N. Scholten, J. Glaubach, M. Hehn, I. Bernhard, I. Aydemir, M. Redaelli, D. Simic, A. Alayli, C. Lemmen, S. Stock, and on beahlf of the fruehstArt Study Group, "Evaluation of an innovative family-centred care and prevention intervention for children with overweight and obesity: a mixed-methods study protocol of the randomised controlled fruehstart study in germany," *BMJ Open*, vol. 16, no. 1, 2026.

[Pub2] K. Ruettger, V. Fluegel, **A. Hagemeier**, K. D. Rosenberger, M. Hellmich, C. Joisten, L. Mause, N. Scholten, J. Glaubach, M. Hehn, I. Bernhard, M. Redaèlli, D. Simic, A. Alayli, S. Stock, and K. Dadaczynski, "Formative evaluation of an early family-centred prevention programme for childhood overweight and obesity (fruehstart): A study protocol," *Children*, vol. 12, no. 12, p. 1613, 2025.

[Pub3] S. Ferdinandus, A. Rühle, A. Lamrani, C. Frei, J. Kaufmann, M. Mäurer, G. Wurschi, P. Jiang, F. Ehret, A. Baehr, A. Hardt, R. Bodensohn, L. Käsmann, M. Waltenberger, D. Scafa, J. Layer, E. Troost, S. Elkhamisy, D. Jazmati, C. Franklin, S. Neppl, **A. Hagemeier**, and M. Trommer, "Abscopal effects in patients with malignant melanoma treated with radiotherapy and immune checkpoint inhibition: analysis of a large observational multicenter study," *Journal for Immunotherapy of Cancer*, vol. 13, no. 10, p. e012717, 2025.

[Pub4] L. Mause, J. Hoffmann, T. Dresbach, **A. Hagemeier**, M. Hellmich, K. L. Kastor, B. Kuehne, A. Reimer, C. Willerscheidt, N. Scholten, and the Neo-CamCare study group, "Effects of webcam access to preterm infants in nicus on parental psychological strain," *Acta Paediatrica*, vol. 114, no. 6, pp. 1421–1431, 2025.

[Pub5] L. C. Mack, **A. Hagemeier**, and D. M. Forner, "Influence of stage and age on survival of patients with vulvar cancer in germany: a retrospective study," *BMJ Open*, vol. 14, no. 8, p. e077960, 2024.

[Pub6] I. Suarez, D. Rauschning, C. Schuller, **A. Hagemeier**, M. Stecher, C. Lehmann, P. Schommers, S. Schlabe, J. J. Vehreschild, C. Koll, C. Schwarze-Zander, J. C. Wasmuth, A. Klingmuller, J. K. Rockstroh, G. Fatkenheuer, C. Boesecke, and J. Rybniker, "Incidence and risk factors for hiv-tuberculosis coinfection in the cologne-bonn region: a retrospective cohort study," *Infection*, 2024.

[Pub7] K. Schmidt-Hellerau, N. Baade, M. Gunther, N. Scholten, C. H. Lindemann, C. Leisse, C. Oberrohrmann, S. Peter, N. Jung, I. Suarez, C. Horn, P. Ihle,

J. Kupper-Nybelen, **A. Hagemeier**, M. Hellmich, and C. Lehmann, "Outpatient parenteral antimicrobial therapy (opat) in germany: insights and clinical outcomes from the k-apat cohort study," *Infection*, vol. 52, pp. 1407–1414, 2024.

[Pub8] N. Cecon-Stabel, M. Omairat, S. Salm, **A. Hagemeier**, A. Dresen, and T. Krieger, "Does the mode of care delivery affect therapeutic alliance, patient care satisfaction or patient reported outcomes? psycho-oncological care evaluation data collected during the covid-19 pandemic," *Journal of Psychosocial Oncology*, pp. 1–21, 2024.

[Pub9] S. Peter, C. Oberrohrmann, H. Pfaff, C. Lehmann, K. Schmidt-Hellerau, V. Brandes, C. Leisse, C. H. Lindemann, P. Ihle, J. Kupper-Nybelen, **A. Hagemeier**, and N. Scholten, "Exploring patients' perspectives: a mixed methods study on outpatient parenteral antimicrobial therapy (opat) experiences," vol. 24, no. 1, p. 544, 2024.

[Pub10] R. Al-Monajjed, P. Albers, J. Droop, D. Fugmann, J. Noldus, R.-J. Palisaar, M. Ritter, J. Ellinger, P. Krausewitz, M. Truß, B. Hadaschik, V. Grünwald, A.-J. Schrader, P. Papavassilis, N. Ernstmann, B. Schellenberger, A. Moritz, C. Kowalski, M. Hellmich, P. Heiden, **A. Hagemeier**, D. Horenkamp-Sonntag, M. Giessing, L. Pauler, S. Dieng, M. Peters, G. Feick, A. Karger, and PRO-P study group, "PRO-P: evaluating the effect of electronic patient-reported outcome measures monitoring compared with standard care in prostate cancer patients undergoing surgery—study protocol for a randomized controlled trial," *Trials*, vol. 25, no. 1, p. 754, 2024.

[Pub11] A. Reimer, L. Mause, J. Hoffmann, **A. Hagemeier**, T. Dresbach, N. Scholten, and C. Neo, "How does stress affect maternal and paternal perceptions of relationship strain after a preterm birth? results of a retrospective survey study," *Acta Paediatr*, vol. 112, no. 4, pp. 762–769, 2023.

[Pub12] V. Bolshinsky, H. Ismail, M. Li, J. Basto, R. Schier, **A. Hagemeier**, K. M. Ho, A. Heriot, and B. Riedel, "Clinical covariates that improve surgical risk prediction and guide targeted prehabilitation: an exploratory, retrospective cohort study of major colorectal cancer surgery patients evaluated with preoperative cardiopulmonary exercise testing," *Perioperative Medicine*, vol. 11, no. 1, p. 20, 2022.

[Pub13] H. Ecker, S. Wingen, **A. Hagemeier**, C. Plata, B. W. Bottiger, and W. A. Wetsch, "Dispatcher self-assessment and attitude toward video assistance as a new tool in simulated cardiopulmonary resuscitation," *Western Journal of Emergency Medicine*, vol. 23, no. 2, pp. 229–234, 2022.

[Pub14] C. Kowalski, N. T. Sibert, C. Breidenbach, **A. Hagemeier**, R. Roth, T. Seufferlein, S. Benz, S. Post, R. Siegel, A. Wiegering, R. Winkels, S. Bieck-Messemer, J. Fahlke, C. Reissfelder, M. Fuchs, T. Herzog, R. Weihrauch, J. Faber-Mertens, H. Rudolph, L. Puskas, K. Kohlhaw, M. Szczerbinska, H. Scheuerlein, P. A. Neumann, S. Hollerbach, M. Riechmann, E. W. Kolbe, N. Weigert, J. Koninger, C. Klink, S. Mussa, A. K. Horn, L. Staib, J. Werner, J. Jahne, M. Aly, H. Mork, R. Grutzmann, P. Piso, S. Dieng, and S. Wesselmann, "Outcome quality after colorectal cancer resection

in certified colorectal cancer centers-patient-reported and short-term clinical outcomes," *Deutsches Ärzteblatt International*, vol. 119, no. 47, pp. 821–828, 2022.

[Pub15] M. Kusch, H. Labouvie, V. Schiewer, N. Talalaev, J. C. Cwik, S. Bussmann, L. Vaganian, A. L. Gerlach, A. Dresen, N. Cecon, S. Salm, T. Krieger, H. Pfaff, C. Lemmen, L. Derendorf, S. Stock, C. Samel, **A. Hagemeier**, M. Hellmich, B. Leicher, G. Hultenschmidt, J. Swoboda, P. Haas, A. Arning, A. Gottel, K. Schwickerath, U. Graeven, S. Houwaart, H. Kerek-Bodden, S. Krebs, C. Muth, C. Hecker, M. Reiser, C. Mauch, J. Benner, G. Schmidt, C. Karlowsky, G. Vimalanandan, L. Matyschik, L. Galonska, A. Francke, K. Osborne, U. Nestle, M. Baumer, K. Schmitz, J. Wolf, and M. Hallek, "Integrated, cross-sectoral psycho-oncology (isPO): a new form of care for newly diagnosed cancer patients in germany," *BMC Health Services Research*, vol. 22, no. 1, p. 543, 2022.

[Pub16] N. Scholten, C. Leisse, V. Brandes, C. Oberrohrmann, P. Ihle, S. Peter, **A. Hagemeier**, M. Hellmich, C. H. Lindemann, C. Samel, H. Pfaff, and C. Lehmann, "Outpatient parenteral antimicrobial therapy in germany: a prospective cohort study protocol," *BMJ Open*, vol. 12, no. 11, p. e061417, 2022.

[Pub17] C. M. Schoppmeier, J. Helpap, **A. Hagemeier**, M. J. Wicht, and A. G. Barbe, "Using the modified schirmer test for dry mouth assessment: A cross-sectional study," *European Journal of Oral Sciences*, vol. 130, no. 4, p. e12880, 2022.

[Pub18] A. G. Barbe, S. Spiritus, **A. Hagemeier**, M. J. Noack, and G. Rohrig, "Oral health assessment of seniors under outpatient care by family doctors : Development and validation of the geriatric outpatient oral health screening," *Zeitschrift für Gerontologie und Geriatrie*, vol. 54, no. 3, pp. 255–263, 2021.

[Pub19] S. A. Elsner, S. S. Salek, A. Y. Finlay, **A. Hagemeier**, C. J. Bottomley, A. Katalinic, and A. Waldmann, "Validation of the german version of the family reported outcome measure (from-16) to assess the impact of disease on the partner or family member," *Health Qual Life Outcomes*, vol. 19, no. 1, p. 106, 2021.

[Pub20] M. Z. Hummelsheim, S. Hamacher, **A. Hagemeier**, M. J. Noack, and A. G. Barbe, "Care need and dry mouth as risk indicators for impaired taste and smell," *Scientific Reports*, vol. 11, no. 1, p. 20419, 2021.

[Pub21] M. Isa, J. Holzki, **A. Hagemeier**, M. A. Rothschild, and C. J. Cote, "Anatomical in vitro investigations of the pediatric larynx: A call for manufacturer redesign of tracheal tube cuff location and perhaps a call to reconsider the use of uncuffed tracheal tubes," *Anesthesia and Analgesia*, vol. 133, no. 4, pp. 894–902, 2021.

[Pub22] L. Karapanos, I. Akbarov, V. Zugor, R. Kokx, **A. Hagemeier**, and A. Heidenreich, "Safety and mid-term surgical results of anterior urethroplasty with the tissue-engineered oral mucosa graft MukoCell((R)): A single-center experience," *International Journal of Urology*, vol. 28, no. 9, pp. 936–942, 2021.

[Pub23] G. Rohrig, A. Spiritus, **A. Hagemeier**, S. H. M. Derman, M. J. Noack, and A. G. Barbe, "Subjective oral health problems of geriatric outpatients and general medical geriatric findings-a pilot study," *Zeitschrift für Gerontologie und Geriatrie*, vol. 54, no. 1, pp. 47–53, 2021.

[Pub24] N. Scholten, S. Bretthauer, K. Eilermann, **A. Hagemeier**, M. Hellmich, J. Hoffmann, D. Horenkamp-Sonntag, C. Jannes, L. Kuntz, P. Mantell, L. Mause, A. Muller, A. Reimer, C. Samel, I. Spiecker Genannt Dohmann, S. Wobbe-Ribinski, C. Woopen, and T. Dresbach, "The effects of webcams on german neonatal intensive care units - study protocol of a randomised crossover trial (neo-camcare)," *BMC Health Services Research*, vol. 21, no. 1, p. 456, 2021.

[Pub25] G. Röhrig, A. Spiritus, **A. Hagemeier**, S. H. M. Derman, M. J. Noack, and A. G. Barbe, "Subjektive mundgesundheitsprobleme ambulant betreuter geriatrischer patienten und allgemeinmedizinisch-geriatrische befunde–eine pilotstudie.," *Zeitschrift für Gerontologie und Geriatrie*, vol. 54, no. 1, pp. 47–53, 2021.

[Pub26] I. Jenniches, C. Lemmen, J. C. Cwik, M. Kusch, H. Labouvie, N. Scholten, A. Gerlach, S. Stock, C. Samel, **A. Hagemeier**, M. Hellmich, P. Haas, M. Hallek, H. Pfaff, and A. Dresen, "Evaluation of a complex integrated, cross-sectoral psycho-oncological care program (isPO): A mixed-methods study protocol," *BMJ Open*, vol. 10, no. 3, p. e034141, 2020.

[Pub27] A. G. Barbe, S. Spiritus, **A. Hagemeier**, M. J. Noack, and G. Röhrig, "Erfassung der mundgesundheit von ambulant betreuten senioren durch hausärzte: Entwicklung und validierung des geriatrischen ambulanten mundgesundheits-screenings," *Zeitschrift für Gerontologie und Geriatrie*, vol. 54, no. 3, p. 255, 2020.

[Pub28] G. Röhrig, I. Becker, **A. Hagemeier**, K. Gutensohn, T. Nebe, working group laboratory diagnostics of the German Society of Hematology, O. (DGHO), and the working group anemia in the aged of the German Geriatric Society (DGG), "White blood cell counts, crp, ggt and ldh in the elderly german population," *Journal of Laboratory Medicine*, vol. 43, no. 2, pp. 57–65, 2019.

Appendix

Appendix of 1. Article

Appendix A

Program code 1: Generate data in SPSS

```
************************************************************.
* Title: Worked Example of the RDD with SPSS                *.
* Author: Anna Hagemeier (anna.hagemeier@uni-koeln.de)      *.
* last Update: 01.12.21                                     *.
************************************************************.

* Simulate test data.

SET RNG=MT MTINDEX=20210527.

INPUT PROGRAM.
LOOP #I=1 TO 500. /*Number of cases*/.
COMPUTE HADS_Diff_Treat=RV.UNIFORM(2.0,7.00) + RV.NORMAL(2,2.5).
COMPUTE HADS_Diff_Cont=RV.UNIFORM(0.00,1.00) + RV.NORMAL(0,1.5).
COMPUTE HADS_Pre=RV.UNIFORM(0.00,42.00).
IF HADS_Pre >14.5 HADS_Post=HADS_Pre-HADS_Diff_Treat.
IF HADS_Pre<= 14.5 HADS_Post=HADS_Pre-HADS_Diff_Cont.
IF HADS_Pre <= 14.5 HADS_group = 0.
IF HADS_Pre > 14.5 HADS_group = 1.
IF HADS_Post < 0 HADS_Post =0.
IF HADS_Post > 42 HADS_Post =42.
END CASE.
END LOOP.
END FILE.
END INPUT PROGRAM.
EXECUTE.

FORMATS  HADS_Diff_Cont HADS_Diff_Treat HADS_Pre HADS_Post (F6.2).
FORMATS  HADS_group (F2.0).
VALUE LABELS HADS_group
0 'HADS<=14.5'
1 'HADS>14.5'.

* Create an interaction term to test whether or not the slopes differ
significantly.
COMPUTE Interact=HADS_group * HADS_Pre.

* Create shift to measure the treatment effect at the pre-defined cut-off.
COMPUTE Shift=(HADS_Pre - 14.5)*(HADS_Pre>14.5).
EXECUTE.

DATASET NAME HADS_RDD WINDOW=FRONT.
OUTPUT NAME=Output1.
EXECUTE.
```

Program code 2: SPSS syntax to conduct a regression discontinuity analysis with the R extension

```
STATS RDD
DEPENDENT=HADS_Post
RUNNING= HADS_Pre
THRESHOLD=14.5
/OPTIONS KERNEL=Rectangular BANDWIDTH = 1.5 SETYPE=HC1 MCCRARY=NO PLOT=YES.
```

1

Program code 3: SPSS syntax to run the regression discontinuity analysis via simple regression approach

```
* Filter Data to conduct Regression within pre-defined bandwidth (b=(13,16)).

DATASET ACTIVATE HADS_RDD.
USE ALL.
COMPUTE filter_$=(HADS_Pre >= 13 & HADS_Pre <= 16).
VARIABLE LABELS filter_$ 'HADS_Pre >= 13 & HADS_Pre <= 16 (FILTER)'.
VALUE LABELS filter_$ 0 'Not Selected' 1 'Selected'.
FORMATS filter_$ (f1.0).
FILTER BY filter_$.
EXECUTE.

FREQUENCIES HADS_group.
EXECUTE.

REGRESSION
  /MISSING LISTWISE
  /STATISTICS COEFF OUTS CI(95) R ANOVA CHANGE
  /CRITERIA=PIN(.05) POUT(.10)
  /NOORIGIN
  /DEPENDENT HADS_Post
  /METHOD=ENTER HADS_Pre
  /METHOD=ENTER  HADS_group Interact Shift.
```

2

A worked example with R

Program code: RDD with the 'rdrobust' package in R (version 4.0.0). The decision for rdrobust is based on the package search and the corresponding number of downloads of the packages (see table 'Results Package Search).

```
################################################################################
# Title: Worked Example of the RDD with R
# Author: Anna Hagemeier (anna.hagemeier@uni-koeln.de)
# last Update-Date: 01.12.21
################################################################################

# read necessary packages
library(haven)
library(rdrobust)
library(pkgsearch)

# In R one can conduct a package search with the package "pkgsearch" to find the
# most frequently used RDD packages
packagesearch <- ps("regression discontinuity", size = 10)

# read and view data (path need t be set individually to where the data was saved)
path <- "Promotion/data/SPSS_Data_Worked_Example.sav"

Data_Worked_Example <- read_sav(path)
View(Data_Worked_Example)

# prepare parameters for prediction
rdd_data <- Data_Worked_Example
threshold <- 14.5
bandwidth <- 1.5 # b=(13,16) <- 14.5 ± 1.5
x <- rdd_data$HADS_Pre
y <- rdd_data$HADS_Post

# plot of the parameters
plot(x,y, xlim = c(0,45), ylim = c(0,45), xlab = "assignment variable",
ylab = "outcome", main = "Scatterplot of Pre-HADS vs. Post-HADS")

# predict regression discontinuity
rdd_prediction <- rdrobust(y, x, c=threshold, h=bandwidth, kernel = 'uniform')
summary(rdd_prediction)

# plot of the regression discontinuity
rdplot(y,x,c=threshold, h=bandwidth, p=1)
```

3

Results Package Search: Extract of the results of a package search on the RDD in R limited to 10 packages sorted by downloads of the last month (search from: 2021-12-01)

	Package-name	Title	Last Release (Version)	Downloads*
1	rdrobust	Robust Data-Driven Statistical Inference in Regression-Discontinuity Designs	2021-09-02 (1.0.6)	2653
2	rdd	Regression Discontinuity Estimation	2016-03-14 (0.57)	1465
3	rddensity	Manipulation Testing Based on Density Discontinuity	2021-03-04 (2.2)	1282
4	rddtools	Toolbox for Regression Discontinuity Design ('RDD')	2021-05-16 (1.5.0)	892
5	DRIP	Discontinuous Regression and Image Processing	2020-09-22 (1.6)	742
6	rdpower	Power Calculations for RD-Designs	2020-12-18 (2.0)	645
7	rdmulti	Analysis of RD-Designs with Multiple Cut-offs or Scores	2021-05-19 (0.8)	641
8	cosa	Bound Constrained Optimal Sample Allocation	2019-08-24 (2.0.0)	617
9	rddapp	Regression Discontinuity Design Application	2020-02-20 (1.2.1)	107
10	optrdd	Optimized Regression Discontinuity Designs	2018-08-17 (1.0.2)	25

* Raw number of package downloads last month, from the RStudio CRAN mirror (search from: 2021-12-01)

Results of rdrobus: Result of the RDD estimation with 'rdrobust'.

```
Call: rdrobust

Number of Obs.                  500
BW type                      Manual
Kernel                      Uniform
VCE method                       NN

Number of Obs.                 170         330
Eff. Number of Obs.             23          12
Order est. (p)                   1           1
Order bias  (q)                  2           2
BW est. (h)                  1.500       1.500
BW bias (b)                  1.500       1.500
rho (h/b)                    1.000       1.000
Unique Obs.                    170         330

================================================================================
        Method     Coef. Std. Err.         z     P>|z|      [ 95% C.I. ]
================================================================================
  Conventional    -4.704      1.432    -3.284     0.001     [-7.511 , -1.896]
        Robust         -          -    -4.749     0.000     [-7.817 , -3.250]
================================================================================
```

4

A worked example with Stata

Program code: STATA do-file to conduct the rdd analysis with rdrobust

```
/*
# Title: Worked Example of the RDD with STATA
# Author: Anna Hagemeier (anna.hagemeier@uni-koeln.de)
# last Update-Date: 01.12.21
*/
clear all

* install the rdrobust fucntion
net install rdrobust,
from(https://raw.githubusercontent.com/rdpackages/rdrobust/master/stata)

*read the data created in SPSS (SPSS_Data_Worked_Example.sav)
import spss using "<your path>", case(lower)

* plot the parameters

scatter hads_post hads_pre

* predict rdd with rdrobust (https://rdpackages.github.io/rdrobust/)

rdrobust hads_post hads_pre, c(14.5) h(1.5 1.5) kernel(uniform)

* plot rdd and use linear model fit

rdplot hads_post hads_pre, c(14.5) h(1.5 1.5) p(1) kernel(uniform)
```

Results of rdrobust within STATA

Sharp RD estimates using local polynomial regression.

Cutoff c = 14.5	Left of c	Right of c
Number of obs	170	330
Eff. Number of obs	23	12
Order est. (p)	1	1
Order bias (q)	2	2
BW est. (h)	1.500	1.500
BW bias (b)	1.500	1.500
rho (h/b)	1.000	1.000

Number of obs =	500
BW type =	Manual
Kernel =	Uniform
VCE method =	NN

Outcome: hads_post. Running variable: hads_pre.

Method	Coef.	Std. Err.	z	P>|z|	[95% Conf.	Interval]
Conventional	-4.7036	1.4322	-3.2841	0.001	-7.51073	-1.89645
Robust	-	-	-4.7491	0.000	-7.81717	-3.24983

5

Confirmation of ethics approval

Final ethical approval for the isPO study was granted on 18 December 2018 under reference number 18-092. The official letter from the Ethics Committee of the Faculty of Medicine at the University of Cologne was removed prior to the publication of this thesis in order to protect data and ensure that no signatures were made publicly available.

Attachment

Curriculum vitae

Work Experience

January 2018 – October 2026	Statistican at Institute of Medical Statistics and Computational Biology at University of Cologne
June 2016 – June 2017	Research assistant at Institute of Social Medicine and Epidemiology at University of Luebeck
April 2016 – June 2016	Internship at Institute of Social Medicine and Epidemiology at University of Luebeck
May 2013 – December 2016	Assistant at Health Information Forum of UKSH Luebeck

Professional development (extract)

May 19 to 23, 2025	Causal Inference for Assessing Effectiveness in Real World Data and Clinical Trials: A Practical Hands-on Workshop, UMIT Tirol
March 27, 2025	Leading projects successfully/project management University Hospital Cologne
March 3, 2024	Refresher course for members of a clinical trial investigation team in accordance with European Regulation (EU) No 536/2014
November 14 to 15, 2022	Current Developments in Cluster Randomised Trials and Stepped Wedge Designs (Webinar)
January 21, 2021	rstudio::global (2021), Online Conference
January 29 to 31, 2020	SQL for epidemiologists, AGENS School Cologne

March 19 to 22, 2018	Introduction to health services research and methodological principles DNVF Spring School, Bonn

Education

July 2017	Degree: Master of Science Master thesis at University of Luebeck at the Institute of Social medicine and Epidemiology: "*Epidemiology of the vulva carcinoma in Schleswig-Holstein*"
April 2015 – July 2017	Studies: Master of Science in Mathematics in Medicine and Life Sciences with specialization in Biometrics and Statistics
May 2015	Degree: Bachelor of Science in Mathematics in Medicine and Life Sciences Bachelor thesis at University of Luebeck at Institute of Medical Biometrics and Statistics: "*K-nearest-neighbours for survival analysis*"
October 2011 – May 2015	Studies: Bachelor of Science in Mathematics in Medicine and Life Sciences at University of Luebeck
June 2011	Abitur at Gymnasium Horn-Bad-Meinberg

Affidavit (Eidesstattliche Erklärung)

Hiermit versichere ich an Eides statt, dass ich die vorliegende Dissertation selbstständig und ohne die Benutzung anderer als der angegebenen Hilfsmittel und Literatur angefertigt habe. Alle Stellen, - einschließlich Tabellen, Karten und Abbildungen -, die wörtlich oder sinngemäß aus veröffentlichten und nicht veröffentlichten anderen Werken im Wortlaut oder dem Sinn nach entnommen sind, sind in jedem Einzelfall als Entlehnung kenntlich gemacht. Ich versichere an Eides statt, dass diese Dissertation noch keiner anderen Fakultät oder Universität zur Prüfung vorgelegen hat; dass sie – abgesehen von unten angegebenen Teilpublikationen und eingebundenen Artikeln und Manuskripten – noch nicht veröffentlicht worden ist sowie, dass ich eine Veröffentlichung der Dissertation vor Abschluss der Promotion nicht ohne Genehmigung der / des Vorsitzenden des IPHS-Promotionsausschusses vornehmen werde. Die Bestimmungen dieser Ordnung sind mir bekannt. Die von mir vorgelegte Dissertation ist von Prof. Dr. Martin Hellmich betreut worden.
Darüber hinaus erkläre ich hiermit, dass ich die Ordnung zur Sicherung guter wissenschaftlicher Praxis und zum Umgang mit wissenschaftlichem Fehlverhalten der Universität zu Köln gelesen und sie bei der Durchführung der Dissertation beachtet habe und verpflichte mich hiermit, die dort genannten Vorgaben bei allen wissenschaftlichen Tätigkeiten zu beachten und umzusetzen.

Übersicht der Publikationen:

- **Hagemeier, Anna**, Christina Samel, and Martin Hellmich. (2022). The regression discontinuity design: methods and implementation with a worked example in health services research. Zeitschrift für Evidenz, Fortbildung und Qualität im Gesundheitswesen 172: 71-77. `https://doi.org/10.1016/j.zefq.2022.04.014`
- **Hagemeier, Anna**, Adams, Anne, Krieger, Theresia, Salm, Sandra, Cecon-Stabel, Natalia, Dresen, Antje, and Hellmich, Martin. (2023). The impact of COVID-19 on the interpretation of psycho-oncological support trial results: a quasi-experimental approach using the data from the new form of care "Integrated cross-sectoral psycho-oncology (nFC-isPO)". BMC Health Services Research, 23(1), 1-8. `https://doi.org/10.1186/s12913-023-09544-y`
- **Hagemeier, Anna**, Rosenberger, Kerstin Daniela and Hellmich, Martin. (2026). Extension of Regression Discontinuity Designs to evaluate time effects - a quasi-experimental approach using the data from the new form of care "Integrated cross-sectoral psyco-oncology (nFC- isPO)". PLOS ONE 21(2): e0343414. `https://doi.org/10.1371/journal.pone.0343414`

Ich versichere, dass ich alle Angaben wahrheitsgemäß nach bestem Wissen und Gewissen gemacht habe und verpflichte mich, jedmögliche, die obigen Angaben betreffenden Veränderungen, dem IPHS-Promotionsausschuss unverzüglich mitzuteilen.

Hürth, 03.07.2026 ....

Ort, Datum Anna Hagemeier