

Supplementary materials

Table of contents

1 Trial registry	1
2 The @myTabu framework	14
2.1 Structure of WBI	14
2.2 Theoretical foundation	17
2.2.1 Module 1: Motivation	19
2.2.2 Module 2: Supervision and social relationships	19
2.2.3 Module 3: Emotion management	19
2.2.4 Module 4: Problem solving	20
2.2.5 Module 5: Offense-supportive cognitions	20
2.2.6 Module 6: Sexuality	21
3 Outcomes and measurement timepoints	23
3.1 Description of trial endpoints	23
3.1.1 SSQ-P	23
3.1.2 SSQ-SO	23
3.1.3 SSQ-CF	23
3.1.4 PHQ-D	23
3.1.5 SSPI-2	23
3.1.6 ICD-11-Screener	24
3.1.7 CARES	24
3.1.8 Official re-offences	24
3.1.9 CVTRQ	24
3.1.10 RCQ	25
3.1.11 OQMPR	25
3.1.12 F-SozU	25
3.1.13 UCLA	25
3.1.14 DERS	25
3.1.15 NARQ	25
3.1.16 BIS-15	26
3.1.17 CUSI	26
3.1.18 SPSI-R	26
3.1.19 BMS	26
3.1.20 HBI-19	26
3.1.21 SSIC	27

3.1.22 SOI-R	27
3.1.23 EKK-R	27
3.1.24 ESIQ	27
3.1.25 WHO-5	27
3.1.26 QNP	27
3.1.27 MRQ	28
3.1.28 WAI-SR	28
3.1.29 EQ	28
4 Baseline characteristics	31
4.1 SSPI-2	31
5 Timing (protocol deviations)	32
5.1 Cumulative time from baseline to finish module	32
5.2 Time working on module	35
6 Primary Outcomes	39
6.1 CARES	39
6.2 Index of Desistance (IoD)	45
7 Secondary Outcomes	52
7.1 Differences from baseline	52
7.2 CVTRQ	54
7.3 RCQ	56
7.4 F-Soz-U	59
7.5 OQMPR	62
7.6 SOI-R	64
7.7 UCLA	66
7.8 BIS-15	69
7.9 CUSI	72
7.10 SPSI-R	75
7.11 EKK-R	77
7.12 DERS	80
7.13 NARQ	82
7.14 ESIQ	85
7.15 BMS	87
7.16 HBI-19	90
7.17 SSIC	92
8 Adverse events	95
8.1 Mood and Risk Questionnaire (MRQ)	95
8.2 Exclusion reasons	98
9 Ancillary Outcomes	99
9.1 WHO-5	99
9.2 WAI-SR: Online-Coach	107
9.3 WAI-SR: Community supervisor	110
10 CONSORT checklist	113

11 Statistical analysis plan	115
12 References	134

1 Trial registry

[Current Version](#)[Version History](#)[Download](#)

DRKS00021256

@myTabu - A placebo controlled randomized trial of a guided web-based intervention program for convicted child sexual abusers and child sexual exploitation material offenders

Organizational Data

DRKS-ID:

DRKS00021256

Recruitment Status:

Recruiting complete, study complete

Date of registration in DRKS:

2020-04-24

Last update in DRKS:

2024-10-29

Registration type:

Prospective

Acronym/abbreviation of the study

@myTabu

URL of the study

<http://www.kompetenz-gegen-missbrauch.de/mytabu> (Link: <http://www.kompetenz-gegen-missbrauch.de/mytabu>)

Brief summary in lay language

BACKGROUND: Sexual child abuse is one of the most harmful events for children. Treatment programs for child abusers are able to significantly reduce the risk to re-offend. In spite of some established treatment projects in primary prevention, treatment offers for child abusers conditionally released from prison are scarce – especially in rural areas. This is alarming, since most re-offences take place after release from prison. Community supervision is effective in reducing re-offenses but only with high financial effort. Online interventions provide a cost-efficient alternative and can be applied smooth in rural areas.

OBJECTIVES: The project aims to develop for the first time an online intervention for discharged child abusers during community supervision. It aims to establish an online risk assessment, which enables practitioners to assess the risk to re-offend repeatedly and efficient. It provides information on the effectiveness of the online

intervention.

METHODS: The effectiveness of the online intervention are evaluated in two federal states (Lower-Saxony and Baden-Wuerttemberg).

EXPECTED RESULTS: The project provides a cost-efficient online intervention, already integrated in community supervision of Lower-Saxony and Baden-Wuerttemberg. Legal guidelines are provided as well as an efficient way to assess the risk to re-offend online. By doing so, re-integration of child abusers into society is improved resulting in higher security for children.

Brief summary in scientific language

Background: There is a high demand for evidence-based and cost-effective treatment concepts for convicted child sexual abuser (CSA) and child sexual exploitation material offenders (CSEMOs) under community supervision (CS). The @myTabu-consortium developed a guided web-based intervention for convicted CSAs and CSEMOs under CS consisting of six online modules targeting psychological meaningful risk factors. The study aims to evaluate the effectiveness of this guided web-based intervention in reducing dynamic risk factors and the risk to re-offend compared to a placebo condition. Furthermore, these dynamic risk factors are measured before and after every module to evaluate their individual effectiveness to reduce the respective risk factor as well as risk to re-offend.

Methods: The methodological design is a placebo controlled randomized add-on trial ($N = 582$) with follow-ups at six time points. The placebo condition controls for attention and expectation effects and comprises the same amount of modules with a comparable time effort as the experimental intervention. The trial is conducted as an add-on to community supervision as usual. Primary outcomes are dynamic risk factors assessed by self-report risk assessment tools and officially recorded re-offenses.

Discussion: To the best of our knowledge, the study compares for the first time the (cost-) effectiveness of a guided web-based intervention for convicted CSAs and CSEMOs under community supervision against a placebo condition.

Health condition or problem studied

Free text:

Sexual offenders which are under community supervision (§§ 56, 57, 68 ff. StGB) because of child abuse (§§ 176, 176a, 176b StGB) or child sexual exploitation material use (§ 184b StGB).

Healthy volunteers:

No Entry

Interventions, Observational Groups

Arm 1:

The experimental condition is a web-based intervention conceptualized as a structured intervention which comprises six modules in a fixed order, each targeting different risk factors. The content is

provided online by psycho-educational blocks complemented by short videos and images. Eighteen different online exercises (e.g., fill-in-the-gap, multiple choice questions) are used to enable the user to train and to internalize the most important aspects of the psycho-educational blocks. The exercises are provided as self-guided tasks with automatic feedback and as coach-guided tasks, where the feedback is provided by online coaches. The feedback of the online coaches is manualized. Since guided web-based interventions are more effective than pure self-help programs, coach-guided exercises are one possibility to enhance the effectiveness of web-based interventions (Wild et al., 2019). Virtual participants accompany the participant, talk/narrate about their experiences and make suggestions from the perspective of someone who is also affected. Furthermore, an in-app messaging system allows the user to communicate in real-time with the online coach as well as the SO. In order to enhance the extrinsic motivation of participants, an automatic reward system is integrated based on a gamification approach. That is, participants receive coins for each finished block and virtual awards. The coins can be changed into Euros and when the participant has collected at least 20 Euros, he can request the payout online. The participant can receive a maximum compensation of 120 euros. Finally, an online first aid kit is provided for each user and filled with individual skills prepared by the user itself during the web-based intervention.

The provided content is divided into six online modules. The participant has one month to complete each module, thus the duration is 24 weeks. Each module is further divided into four sessions. Each session has a duration of about 60 minutes and the participant has one week to finish a session.

Arm 2:

The control condition is a placebo web-based intervention with the same mode, dose, and amount of support and attention by supporting online coaches as in the experimental condition - but the content of the modules is unrelated to the proposed dynamic risk factors. The content of the placebo condition comprises information on healthy living, e.g., healthy nutrition, sporting activities, or healthy sleep. The amount of exercises is the same as in the experimental condition, but exercises are specifically designed to not induce any sustainable effect.

Endpoints

Primary outcome:

Index of Desistance, IoD: The first primary endpoint is the difference in the pre-post changes of the Index of Desistance (IoD) in the experimental condition compared to the placebo condition. The IoD represents the individual risk to re-offend assessed through self-report. This composite endpoint reflects improvements with regard to psychological meaningful risk factors, offense behaviors as well as offence-related behaviors that have not been registered or convicted yet. The IoD comprises three self-report instruments, which were developed and are currently evaluated by members of the @myTabu-consortium: (1) The Acute-2007-SR is an online and self-report adaption of the Acute-2007 (hanson et al., 2013) risk-assessment tool in order to assess treatment induced short-term changes in recidivism risk factors. (2) The Checklist of Treatment Effectiveness (CTE) is developed to assess the differential treatment-induced changes in every individual treatment module. Thus, the CTE assesses the constructs that are targeted by the six modules of the invention. (3) The Checklist of Behavioral Misconduct (CMC) measures offense-related behavior as well as low- and high-level new offences. Offense-related behavior includes behaviors which are relevant for individuals previously convicted of sexual offenses prior to a new offense (e.g., recidivistic behavior like, e.g., if person convicted of sexual abuse of minors visits a playground or a school). The CMC could be seen as post-treatment deviant behavioral indicators. It is assessed pre and post each intervention module. The IoD will provide the individual risk level to re-offend by summing up the scores of the Acute-2007-SR, CTE, and CMC.

Officially recorded re-offenses in the experimental condition compared to the placebo condition are the second primary endpoint. Officially recorded re-offenses (based on the information of the Bundeszentralregisters, BZR) are obtained five years after last patient out.

Secondary outcome:

Pre-post difference scores of psychometric questionnaires in the experimental condition compared to the placebo condition are defined as the secondary endpoints. The questionnaires are assessed before and after each intervention module. This approach allows evaluating each web-based intervention module individually with regard to its targeted risk factors.

Corrections Victoria Treatment Readiness Questionnaire. The Corrections Victoria Treatment Readiness Questionnaire (CVTRQ Casey et al., 2007) is a self-report measure designed to assess treatment readiness in offenders who have been referred to a cognitive skills program. It comprises 20 items and consists of four scales: (1) Attitudes and motivation, (2) emotional reaction, (3) offending beliefs, and (4) efficacy. Each item has to be answered on a five-point Likert-scale. The CVTRQ shows an acceptable convergent validity, discriminant validity as well as predictive validity (Casey et al., 2007). To the best of our knowledge, there is no validated German version of the CVTRQ. @myTabu translated the CVTRQ. The translation was checked and translated back into English. Resulting differences due to adaptation were inspected and discussed in terms of item content.

Readiness to Change Questionnaire. The Readiness to Change Questionnaire (RCQ; Rollnick et al., 1992) is a 12-item questionnaire originally designed to identify the stage of change reached by an excessive drinker of alcohol. The German version (RCQ-D; Hannöver et al., 2002) was adapted within the project by changing the questions regarding drinking of alcohol in questions regarding the offences of CSAs and CSEMOs. Responses are made on a five-point Likert-scale. The RCQ-D as well as the RCQ shows good psychometric properties (Hannöver et al. 2002). Psychometric properties of the version adapted for CSAs and CSEMOs are not available yet.

Optimized Questionnaire for the Measurement of Psychological Reactance. The Questionnaire for the Measurement of Psychological Reactance (QMPR, Merz, 1983) is a questionnaire for the assessment of psychological reactance defined as the theory that people resist attempts to constrain either their thoughts or their behaviors. The Optimized Questionnaire for the Measurement of Psychological Reactance (OQMPR; Herzberg et al., 2002) is a German variant of the QMPR. The OQMPR consists of 12 statements on a five-point Likert-scale. It has a test-retest-reliability of $\alpha = .85$ (Herzberg et al., 2002).

Social support questionnaire, short version. The seven-item short version of the Social Support Questionnaire (F-SozU K-7) is an efficient questionnaire to assess perceived social support (Dunkel et al., 2005). Each item comprises a five-point Likert-scale. Internal consistency of the original long version showed Cronbach's alpha between .81 and .93 (Sommer et al., 1989).

University of California Los Angeles Loneliness Scale. The UCLA Loneliness Scale (Russel et al., 1996) consists of 20 items in order to assess subjective feelings of loneliness. The German short version (Bilsky et al, 1989) consists of 12 items on a four-point Likert-scale. A German study with CSOs and CSEMOs demonstrated a high reliability of $\alpha = .92$ (Neutze et al., 2012).

\paragraph{Bumby Molest Scale.} The Bumby Molest Scale (BMS; Bumby, et al., 1996) is a self-report questionnaire which assesses cognitive distortion in CSOs (German Version: Feelgood et al., 2009). It consists of 38 items on a four-point Likert-scale. The German version shows good construct validity,

internal consistency and test-retest reliability (Feelgood et al., 2009).

Social Problem-Solving Inventory Revised. The Social Problem-Solving Inventory Revised (SPSI-R; D'Zurilla et al., 2002) is a self-report questionnaire for the assessment of the five dimensions in the social problem-solving model: (1) positive problem orientation, (2) negative problem orientation, (3) rational problem solving, (4) impulsivity/carelessness style, and (5) avoidance style. The SPSI-R consists of 52 items on a five-point Likert-scale. It shows good psychometric properties for sexual offenders (Wakeling et al., 2007). The validated German version is a short form consisting of 25 items (Graf et al., 2003).

Difficulties in Emotion Regulation Scale - Sub-scale Impulsivity. The Difficulties in Emotion Regulation Scale (DERS; Gratz et al., 2004) is a self-report questionnaire to assess emotion dysregulation. The sub-scale 'impulse control difficulties' is used as outcome measure and comprises five items on a five-point Likert-scale. The German version is used, which shows a good internal consistency, construct and predictive validity (Ehring et al., 2008).

Negative Affect Repair Questionnaire. The Negative Affect Repair Questionnaire (NARQ; Scherer et al., 2013) is a self-report questionnaire to assess strategies to regulate negative affect in a systematic manner. It consists of 17 items on a five-point Likert-scale and provides a good construct validity. Reliability scores (Cronbach's alpha) for the three NARQ scales ranged between .71 and .80 (Scherer et al., 2013).

Barratt Impulsiveness Scale. The Barratt Impulsiveness Scale (BIS-11, Patton et al., 1995) is a questionnaire developed to assess the personality/behavioral construct of impulsiveness. It consists of 30 items on a four-point Likert-scale. The German short version (BIS-15, Meule et al., 2011) consisting of 15 items is used in the clinical trial. The BIS-15 is an efficient measure of impulsiveness with good internal consistency (alpha = .81; Meule et al., 2011).

Coping Using Sex Inventory. The Coping Using Sex Inventory (CUSI; Cortoni et al., 2001) is a questionnaire to assess the presence of and the degree to which sex was used to deal with problematic situations. It consists of 16 items on a five-point Likert-scale with a satisfying internal consistency (cortoni et al., 2001). The translation was checked and translated back into English by the authors. Resulting differences due to adaptation were inspected and discussed in terms of item content. For another German translation, it has been shown that the CUSI is able to assess therapy-induced changes of CSOs (Beier et al., 2015).

Hypersexual Behavior Inventory – 19. Sexual preoccupation is assessed by the Hypersexual Behavior Inventory – 19 (HBI-19; Klein et al., 2014). The HBI-19 is a three-factor measure (coping, control, and consequences) developed to assess hypersexual behaviour. The instrument consists of 19 items (e.g., "I use sex to forget sorrows of everyday life.") answered on a scale from 1 (never) to 5 (very often). The maximum score is 95, with higher scores indicating a higher level of sexual preoccupation. The questionnaire was shown to have good reliability (alpha = .90) and validity (Klein et al., 2014).

Emotional congruence with children. Emotional congruence with children is assessed by the Questionnaire on Emotional Congruence with Children – Revised (EKK-R; Mack et al., 2012) including three factors (special relationship to children, immaturity, and emotional closeness to children). Twenty items are answered on a four-point scale. The questionnaire demonstrated good reliability (alpha = .80) and validity (Mack et al., 2012).

Specific self-efficacy for modifying sexual interest in children. Specific self-efficacy for modifying

sexual interest in children (SSIC) is assessed by the Self-Efficacy for Modifying Sexual Interest in Children Scale (SSIC-Scale; Tozdan et al., 2015). Six items on the participant's conviction regarding being able to change their sexual interest in children (e.g., "I can succeed in reducing my sexual interest in children.") were answered on a scale from 1 (do not agree at all) to 5 (totally agree). The maximum score is 30, with higher scores indicating a higher level of self-efficacy. The instrument was shown to have good reliability ($\alpha = .87$) and validity (Tozdan et al., 2015). Four variables that were shown to be related to the SSIC (Tozdan et al., 2015, 2019) are also assessed: flexibility of sexual interest in children (three items on previous experiences concerning changes in the participants' sexual interest in children), exclusiveness of sexual interest in children (single-item question), motivation to change sexual interest in children (single-item question), and age of onset of sexual interest in children (single-item question).

Sexual interest in children. Sexual interest in children is assessed by two measurements. The Explicit Sexual Interest Questionnaire (ESIQ; Banse et al., 2010) directly assesses pedophilic interest. It consists of two scales measuring sexual behaviour (20 items, e.g., "I enjoyed orally stimulating a man.") and sexual fantasies (20 items, e.g., "I find it attractive to imagine a little boy sexually stimulating me."). All items are answered on a scale from 1 (totally disagree) to 5 (totally agree). The reliability of the instrument ranges between .86 and .97 (Banse et al., 2010). The second measurement is the Item 2a of the Sexual Outlet Inventory – Revised (SOI-R; Briken et al., 2010) which assesses the desire for sexual activity involving children on a visual analogue scale from 0 (desire is absent) to 100 (I have to act to satisfy the desire). Higher values on the scale indicate a stronger sexual interest in children (Briken et al., 2010).

Study Design

Purpose:

Treatment

Allocation:

Randomized controlled study

Control:

Placebo

Phase:

III

Study type:

Interventional

Mechanism of allocation concealment:

No Entry

Blinding:

Yes

Assignment:

Parallel

Sequence generation:

No Entry

Who is blinded:

Assessor, Data analyst, Patient/subject

Recruitment

Recruitment Status:

Recruiting complete, study complete

Reason if recruiting stopped or withdrawn:

No Entry

Recruitment Locations

Recruitment countries:

Germany

Number of study centers:

Monocenter study

Recruitment location(s):

Other community supervision Lower-Saxony, Other community supervision Baden-Württemberg, Other community supervision Brandenburg, Other community supervision Rheinland-Pfalz, Other community supervision Nordrhein-Westfalen, Other community supervision Sachsen, Other community supervision Sachsen-Anhalt, Other community supervision Thüringen, Other community supervision Bayern, Other community supervision Hessen, Other community supervision Mecklenburg-Vorpommern, Other community supervision Berlin, Other community supervision Schleswig-Holstein

Recruitment period and number of participants

Planned study start date:

2021-01-01

Actual study start date:

2021-03-01

Planned study completion date:

2024-09-30

Actual Study Completion Date:

2024-09-30

Target Sample Size:

582

Final Sample Size:

369

Inclusion Criteria

Sex:

All

Minimum Age:

18 Years

Maximum Age:

no maximum age

Additional Inclusion Criteria:

(1) being under community supervision, (2) being convicted for at least one child abuse (§§ 176, 176a, 176b StGB) or child sexual exploitation material use (§ 184b StGB), and (3) being at least 18 years of age.

Exclusion Criteria

(1) a probation period shorter than six months at enrolment, (2) no access to a PC, tablet or smart phone, (3) a severe acute psychiatric disorder (e.g., acute psychosis), (4) a severe cerebro-organic disorder, (5) a severe cognitive impairment, (6) no banking account for receiving the allowance, (7) withdrawal of the informed consent, and (8) no written informed consent

Addresses**Primary Sponsor**

UNIVERSITÄTSMEDIZIN GÖTTINGEN GEORG-AUGUST-UNIVERSITÄT Ressort Forschung und Lehre Studienzentrum UMG

Jutta Heinrich

Von-Bar-Str. 2/4

37075 Göttingen

Germany

Telephone:

+49 551 / 39-60829

Fax:

+49 551 / 39-60846

 Contact per E-Mail

URL:

No Entry

Investigator Sponsored/Initiated Trial (IST/IIT):

Yes

Contact for Scientific Queries

Universitätsmedizin Göttingen Klinik für Psychiatrie und Psychotherapie - Forensische Psychiatrie

Dr. Peter Fromberger

Rosdorfer Weg 70

37081 Göttingen

Germany

Telephone:

0551-4022108

Fax:

0551-4022110

 [Contact per E-Mail](#)

URL:

<http://www.psychiatrie.med.uni-goettingen.de/de/content/forschung/257.html> (Link: <http://www.psychiatrie.med.uni-goettingen.de/de/content/forschung/257.html>)

Contact for Public Queries

Universitätsmedizin Göttingen Klinik für Psychiatrie und Psychotherapie - Forensische Psychiatrie
Dr. Peter Fromberger
Rosdorfer Weg 70
37081 Göttingen
Germany

Telephone:

0551-4022108

Fax:

0551-4022110

 [Contact per E-Mail](#)

URL:

<http://www.psychiatrie.med.uni-goettingen.de/de/content/forschung/257.html> (Link: <http://www.psychiatrie.med.uni-goettingen.de/de/content/forschung/257.html>)

Principal Investigator

Universitätsmedizin Göttingen
Dr. Peter Fromberger
Rosdorfer Weg 70
37081 Göttingen
Germany

Telephone:

+495514022108

Fax:

No Entry

 [Contact per E-Mail](#)

URL:

No Entry

Sources of Monetary or Material Support

Government or public funding body, financed by tax revenue (e.g. the German DFG, BMFTR)

Bundesministerium für Bildung und Forschung Dienstsitz Berlin
Friedrichstraße 130 B
10117 Berlin
Germany

Telephone:

No Entry

Fax:

No Entry

 Contact per E-Mail

URL:

<http://www.bmbf.de> (Link: <http://www.bmbf.de>)

Ethics Committee

Address Ethics Committee

Ethikkommission der Universitätsmedizin Göttingen
Von-Siebold-Straße 3
37075 Göttingen
Germany

Telephone:

+49-551-3961261

Fax:

+49-551-3969536

 Contact per E-Mail

URL:

No Entry

Vote of leading Ethics Committee

Date of ethics committee application:

2020-01-27

Ethics committee number:

16/2/20

Vote of the Ethics Committee:

Approved

Date of the vote:

2020-02-21

Further identification numbers

Other WHO Primary Registry or Data Provider ID:

No Entry

EudraCT Number:

No Entry

UTN (Universal Trial Number):

No Entry

EUDAMED Number:

No Entry

IPD - Individual Participant Data

Do you plan to make participant-related data (IPD) available to other researchers in an anonymized form?:

No

IPD Sharing Plan:

No Entry

Study protocol and other study documents

Study protocols:

- [!\[\]\(cfd7b04e499fc632d34654cfc84d7bee_img.jpg\) Studienprotokoll \(Version 1.0.0\)](#)

Study abstract:

No Entry

Other study documents:

No Entry

Background literature:

No Entry

Related DRKS studies:

No Entry

Publication of study results

Planned publication:

No Entry

Publications/study results:

No Entry

Date of the first journal publication of results:

No Entry

DRKS entry published for the first time with results:

No Entry

Basic reporting

Basic Reporting / Results tables:

No Entry

Brief summary of results:

No Entry

2 The @myTabu framework

Please note, that the following descriptions are partly adapted from Fromberger et al. (2021).¹

2.1 Structure of WBI

The WBI @myTabu consists of three isolated sections, one only for participants, one only for the online coaches, and one section for community supervisors (see Figure 1). The access to these sections was secured by username and password full-filling the criteria of the DSGVO. On the landing page of the coach section, a dashboard informs the coach on tasks he has to perform (e.g. unanswered messages by a client, open guided tasks). Each coach had only access to the participants he was responsible for. The landing page links to the in-app messenger and a page for working on guided tasks together with the participant. The supervisor section, only accessible by participating community supervisors, provides a landing page which summarized important key aspects (treatment progress, critical events) of all participants, which have been supervised by this supervisor. Due to security reasons, the supervisor was able to see the raw values of the primary endpoint. If a item of the primary endpoint has been graded as a potential hint to recidivism by lawyers within the research team during development of the app, additional judicial information are provided specifically for the federal state the supervision officer is working at.

The participant section, only accessible by participants, provides on the landing page key aspects of their progress, e.g. current saldo, gudied exercises the have to work on, a help page, a first aid kit as well as a section for the WBI. The participants are remembered on open guided tasks or being late with the intervention by in-app alerts. The WBI starts with two introductory sessions, in order to explain the participants how the app works. After finishing the instruction section, the participants is linked to the first lection of the WBI. The structure and organization of the content was identical across both trial arms. The WBI was structured into six modules, each module into four sessions and each session into six to eight lections. A lection was divided into different content blocks: guided exercises, psychoeducation, examples (virtual participants), and automated exercises. Virtual participants accompanied the participant, narrate about their experiences, and make suggestions from the perspective of someone who is also affected (scored by professional actor)At the beginning of each session, key elements from the previous session or module were recapped. At the end of each session, the most important points of that session or module were summarized. Content created during exercises was partly dynamic integrated in consecutive sessions. For example, at the start of the intervention participants created a list of positive activities, which was later used to suggest ways in which they could reward themselves for completing a session or module. In both trial arms, the program included digital psycho-educational blocks incorporating multimedia elements such as videos and graphics, as well as guided and unguided exercises. A total of eighteen different online exercises (e.g., fill-in-the-gap tasks, multiple-choice questions, four-field-tasks) were integrated into the WBI. These exercises were either offered as self-guided tasks with automated feedback or as coach-guided tasks in which feedback was provided by online coaches. In the intervention arm, the most relevant aspects of each session were delivered as guided tasks. In the placebo arm, the allocation of tasks as guided or unguided was determined based on a predefined matching procedure to replicate the distribution of guided and unguided tasks in the intervention arm. The feedback of the online coaches was manualized: Within a special section of the WBI, the coach was able to choose pre-build templates for each guided task (see Figure 2). Furthermore, an in-app messaging system allowed the user to communicate with the online coach as well as the supervision officer in a secured environment. It was ensured that messages and feedback for guided tasks has been provided within 36 hours by the coach. In order to enhance the

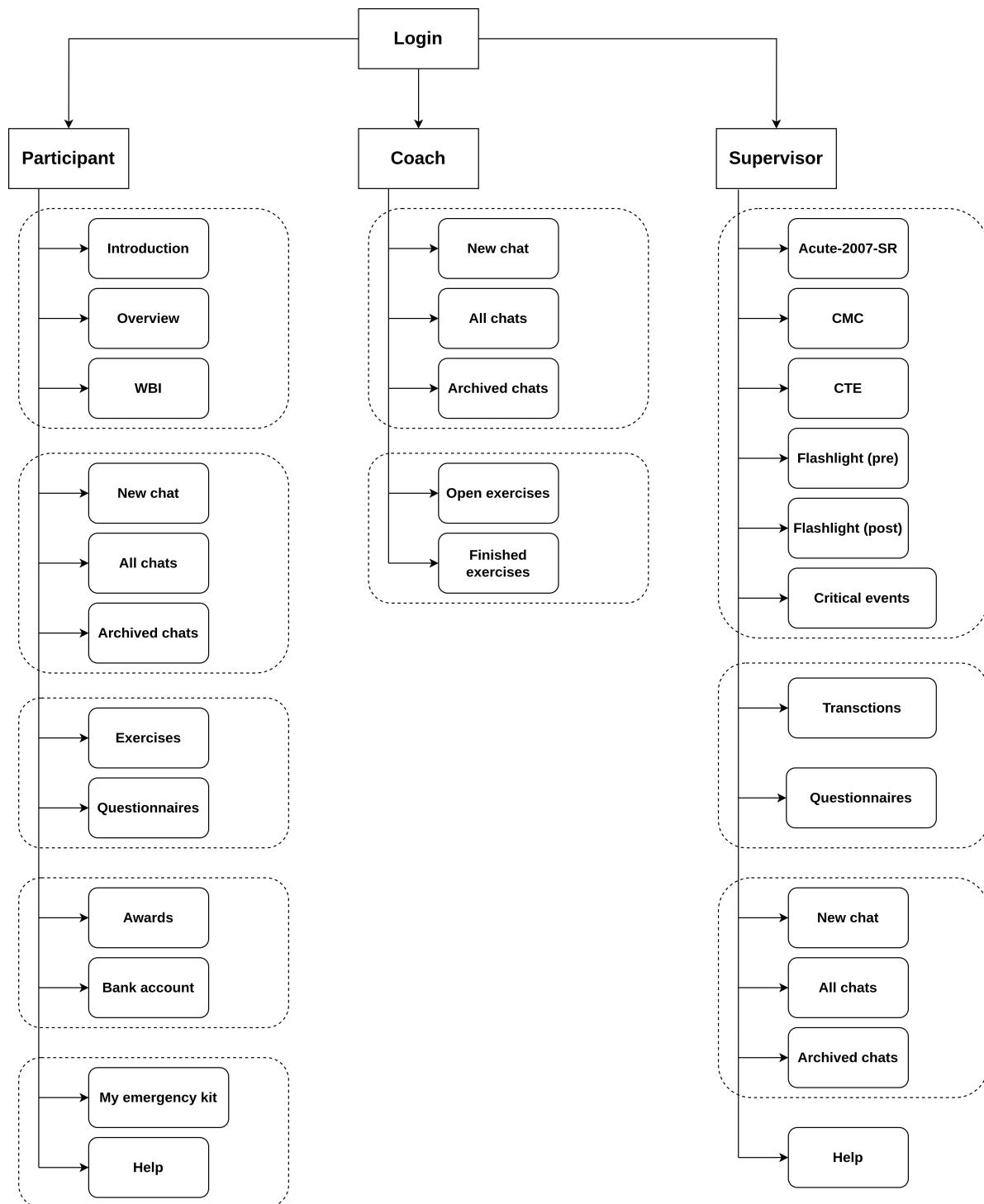


Figure 1: Sitemap of the WBI @myTabu. WBI = Web-Based Intervention.

extrinsic motivation of participants, an automatic reward system is integrated based on a gamification approach. That is, participants receive coins and virtual awards for each finished content block. The coins can be changed into Euros, if the participant has collected at least 15 Euros, he can request the payout online. The participant can receive a maximum compensation of 120 Euro. Finally, an online first aid kit was provided for each user and filled with individual skills prepared by the user itself during guided tasks in the intervention arm. In the placebo arm, the first add kit was not individualized but pre-filled with basic hints. Throughout all modules, the speech was friendly, empathetic, simple, and easy to understand (e.g., no technical terms, examples of everyday life) in order to account for the different cognitive functioning level of participants.

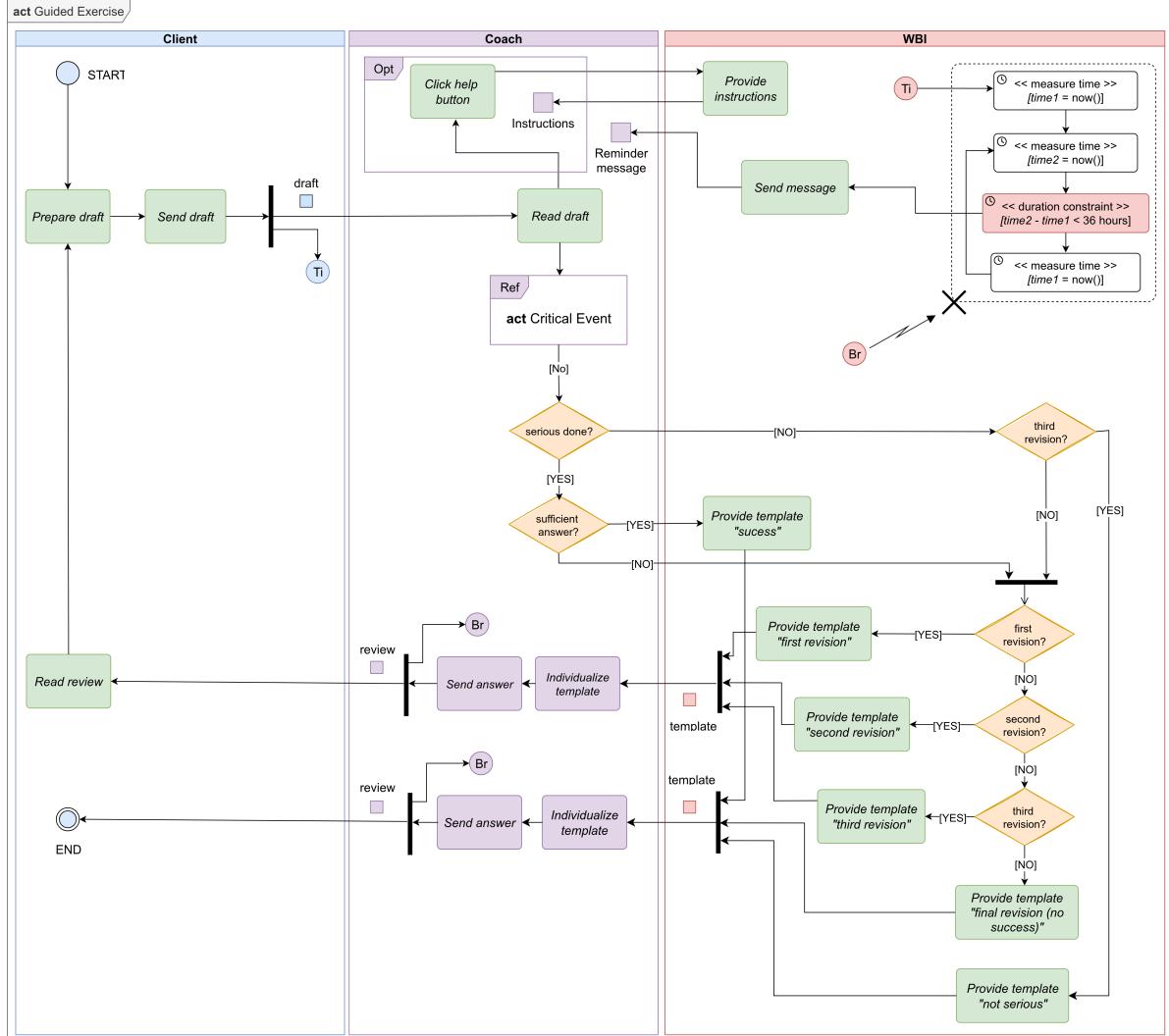


Figure 2: Process flow of guided exercises. 29 guided exercises have been developed for each trial arm. The figure shows the process of reviewing a guided task by study therapists.

2.2 Theoretical foundation

Already established therapeutic concepts from in-face community-based treatment programs for individuals who committed sexual offenses against children^{2,3} as well as cognitive-behavioral treatment concepts build the theoretical ground of the WBI; see¹ for more details. The in-person programs have been adapted for online use. Both programs have been designed for CSA as well as CSEM offenders. Parts relevant to the meaningful risk factors of both in-face programs have been digitalized and parts on risk factors not supported by empirical evidence have been skipped in order to shorten the WBI. The content of the intervention arm focused exclusively on psychological meaningful risk factors^{4,5}, see Figure 3 for an overview of the theoretical concept of the WBI. To achieve a reduction in re-offenses, it seems to be most promising to target dynamic factors, which are changeable and strongly associated with recidivism⁵. Mann et al. (2010) introduced the term meaningful psychological risk factors for sexual offending by identifying those risk factors in the literature, which are dynamic and therefore changeable by treatment and are empirically proved in enhancing the risk for recidivism. Recently, Seto et al. (2023) replicated the findings and proofed the meaningful risk factors with a broader and more recent literature basis. The empirical basis for risk factors for CSAM-exclusive offences is not as broad as for sexual offending in the sense of in-person offences.⁶ elaborated in their review risk factors for CSEM-exclusive offenders. Atypical sexual interests, sexual pre-occupation, sexual arousal, access to the internet and children, and distorted cognitions are assumed to play a key role in recidivism. Thus, there is an overlap between risk factors for CSA offenses and risk factors for CSAM offenses⁷. But it is assumed, that e.g. the content of offense-supportive attitudes differs from the content of offense-supportive attitudes for mixed offenders or CSA-exclusive offenders. Access to children or the internet is a so-called facilitating factor in the sense of the Motivation-Facilitation Model of sexual offending and therefore only indirect changeable by treatment⁸. In consequence, addressed risk factors were the same for participants with CSA offences and participants with CSEM offences. Virtual participants (four male and one female offender with different offense histories) accompanied the participant, narrate about their experiences, and make suggestions from the perspective of someone who is also affected (scored by professional actor). Guided exercises are used at the core of cognitive-behavioral treatment within the WBI. Guided exercises have been developed to train important new concepts online. They ensure, that the participants have to think about the previous contents to answer them correctly, since the online coach reviews the answers. Additionally, guided exercises are the most probable content block in which the online coach gets in touch with the participant (see Figure 2). All modules within the intervention arm comprised techniques of motivational interviewing⁹ to enhance or maintain the motivation of the participants. The WBI is designed for female as well as male participants, but due to a lack of knowledge with regard to specific risk factors for female sex offenders at development of the WBI, the content remains identical for women and men. In order to make the WBI suitable for female participants, the used language was adapted and one female virtual participant has been designed and integrated. The placebo arm was developed to ensure the same mode, dose, and amount of support and attention by supporting online coaches as in the intervention arm - but the content of the modules was unrelated to meaningful dynamic risk factors. The content of the placebo condition comprised information on healthy living, e.g., healthy nutrition, physical activities, or healthy sleep. The amount of guided exercises was the same as in the intervention arm ($n = 29$ guided exercises), but the exercises have been specifically designed to not induce any sustainable effect. In the following paragraphs, the theoretical background of each module is described in detail.

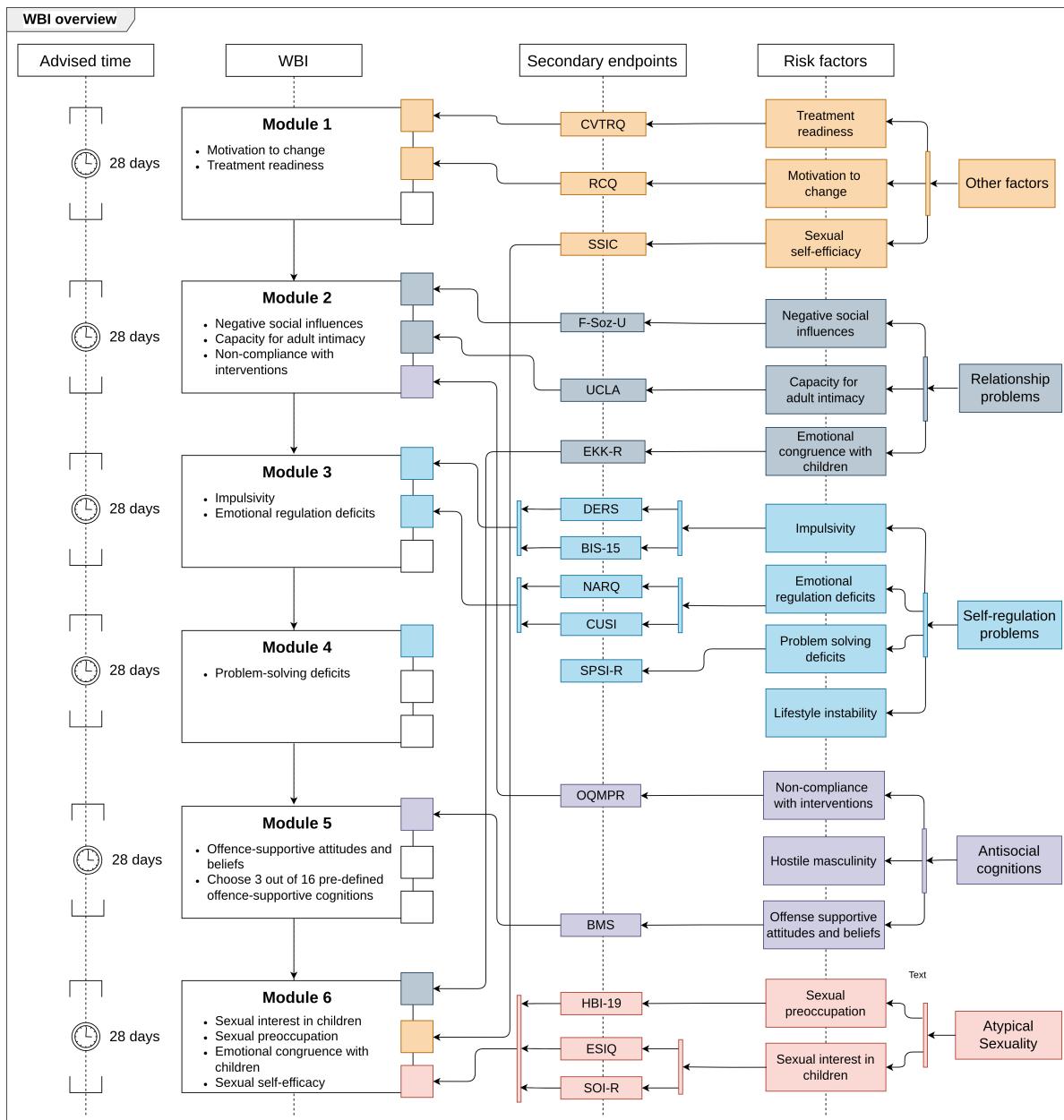


Figure 3: Overview of the WBI. The figure shows the correspondence between meaningful risk-factors, secondary outcomes of the clinical trial, and the content of the trial arm. Abbreviations: BIS-15 = Barratt Impulsiveness Scale-15, BMS = Bumby Molest Scale, CUSI = Coping Using Sex Inventory, CVTRQ = Corrections Victoria Treatment Readiness Questionnaire, DERS = Difficulties in Emotion Regulation Scale, EKK-R = Questionnaire on Emotional Congruence with Children-Revised, ESIQ = Explicit Sexual Interest Questionnaire, F-Soz-U = Seven-item short version of the Social Support Questionnaire, HBI-19 = Hypersexual Behavior Inventory-19, NARQ = Negative Affect Repair Questionnaire, OQMPR = Questionnaire for the Measurement of Psychological Reactance, RCQ = Readiness to Change Questionnaire - German version, SOI-R = Sexual Outlet Inventory revised, subscale desire for sexual activity with children, SPSI-R = Social Problem-Solving Inventory Revised, SSIC = Specific self-efficacy for modifying Sexual Interest in Children, UCLA = UCLA Loneliness Scale - German short version.

2.2.1 Module 1: Motivation

Ward et al. identify motivation to change and subsequent treatment engagement as intermediate targets that must be addressed before criminogenic needs can be modified.¹⁰ Empirical evidence indicates that higher treatment motivation is associated with greater therapeutic success among individuals who committed CSA offences and with programme dropout in web-based interventions for alcohol use disorder.^{11,12} Reviews of web-based interventions suggest that most treatment discontinuation occurs before the active treatment phase begins and that the timing of dropout is comparable between in-person and therapist-guided web-based treatments.¹³ As motivation to change is likely to influence treatment engagement, outcomes, and dropout among ICCO, these findings highlight the importance of addressing motivation at the outset of the study.^{11,12} The motivation module is therefore grounded in motivational interviewing principles, which conceptualise behavioural change as a process marked by ambivalence.⁹ Therapists support participants in exploring reasons for and against change, with a focus on personal values and goal setting to foster intrinsic motivation. The approach is guided by four core principles—expressing empathy, developing discrepancy, avoiding argumentation, and supporting self-efficacy—implemented through open-ended questioning, reflective listening, affirmations, change-talk strategies, resistance management, confidence-building techniques, and summaries.

2.2.2 Module 2: Supervision and social relationships

Theoretical and empirical models of criminal behaviour underscore the relevance of social influences for recidivism.¹⁴ Negative social environments and resistance to rules and supervision have been identified as meaningful risk factors for sexual reoffending.^{5,15} Accordingly, interventions for ICCO under CS should promote the development of protective social networks. Such networks extend beyond family and peers to include therapists, supervision officers, and social workers. The supervision and social relationships module of @myTabu therefore aims to strengthen motivation to establish positive social connections and to reduce resistance to supervision through psychoeducational and cognitive-behavioural techniques, including the illustration of associations between negative social networks and reoffending and the provision of strategies to disengage from criminogenic influences. A further established risk factor for sexual reoffending is the absence of emotionally intimate adult relationships.⁵ Individuals who have committed sexual offences and maintain stable adult partnerships show lower recidivism rates than those who do not, and deficits in emotionally intimate adult relationships are linearly associated with sexual reoffending.^{15,16} Consequently, this module provides information on these associations and offers guidance to support the development of emotionally intimate adult relationships. To date, the effectiveness of the specific supervision-related content described here has not been empirically evaluated, either in in-person or in web-based interventions.

2.2.3 Module 3: Emotion management

Lifestyle impulsiveness as a pattern characterised by low self-control, instability in employment and housing, absence of meaningful daily routines, irresponsible decision-making, and limited or unrealistic long-term goals is associated with recidivism. Lifestyle impulsiveness is a well-established risk factor for sexual reoffending with a broad empirical base.⁵ For dysfunctional coping, Knight and Thornton demonstrated that external coping—defined as a tendency to respond to stressful and negatively valenced emotional events in a reckless and impulsive manner—predicts sexual recidivism.¹⁷ To address these risk factors, @myTabu integrates skills training derived from Dialectical Behaviour Therapy (DBT¹⁸) and Acceptance and Commitment Therapy (ACT¹⁹) with the aim of reducing impulsivity and strengthening

coping capacities in response to stressful or negatively experienced life events. DBT skills training targets emotion regulation and distress tolerance by teaching alternative behavioural strategies to replace external or sexualised coping. DBT has previously been shown to reduce impulsive behaviour in forensic populations, and web-based interventions based on DBT principles have demonstrated effectiveness, particularly in reducing stress, across several psychiatric conditions.^{20–22} ACT focuses on fostering a meaningful life by promoting acceptance of negative internal experiences and providing strategies for adaptive responses to emotional distress.¹⁹ ACT has been successfully applied in offender treatment settings and has shown promising results when delivered via web-based interventions for various psychiatric disorders.^{21–23}

2.2.4 Module 4: Problem solving

Meta-analyses have demonstrated a significant association between poor social problem-solving abilities and sexual recidivism.^{4,5} Social problem solving is defined as a self-directed cognitive-behavioural process through which individuals attempt to identify and implement effective solutions to problems encountered in everyday life.²⁴ Within the widely accepted cognitive model proposed by D'Zurilla and Goldfried, problem solving comprises two dimensions: problem orientation and problem-solving style. Problem orientation may be positive (PPO) or negative (NPO). PPO is associated with adaptive coping and reduced emotional distress, whereas NPO is linked to avoidance and ineffective coping strategies.²⁵ According to this model, three problem-solving styles can be distinguished: rational problem solving (RPS), impulsive coping style (ICS), and avoidance style (AS). RPS involves a systematic and planned application of skills to achieve adaptive solutions. In contrast, ICS is characterised by unsystematic responses, often involving the implementation of the first available option, while AS is marked by avoidance of problems and reliance on external solutions. Empirical evidence indicates that ICCO exhibit higher levels of NPO and lower levels of PPO than the general population and show a greater tendency towards ICS and AS compared with RPS.²⁶ Moreover, poor problem-solving skills in this population have been shown to be amenable to change through psychological interventions.^{27,28} The problem-solving module therefore aims to improve deficient problem-solving skills by drawing on the well-established principles of Problem-Solving Therapy (PST²⁶). The module focuses on fostering a positive problem orientation and strengthening rational problem-solving skills.

2.2.5 Module 5: Offense-supportive cognitions

Offense-supportive attitudes among ICCO can be summarised as beliefs that children are capable of sexually mature relationships, are not harmed by sexual contact with adults, or actively provoke sexual behaviour in adults.^{5,29} These attitudes must be clearly distinguished from attempts to excuse or justify one's own specific offences, which should not be the focus of cognitive treatment.⁴ A meta-analysis by Helmus et al. demonstrated that only offense-supportive attitudes, as defined above, constitute a psychologically meaningful risk factor in terms of predicting sexual recidivism among ICCO.³⁰ In contrast, offence-specific justifications have not been shown to be predictive and are therefore not considered primary targets of intervention. The offense-supportive attitudes module applies cognitive restructuring to modify offense-supportive beliefs. Cognitive restructuring is the most widely used and best-evaluated technique for changing maladaptive cognitions in this population. Modifying implicit theories, attitudes, and beliefs is challenging, particularly among ICCO, as such beliefs may not be experienced as harmful by the individual and may instead serve to minimise perceived harm resulting from their own offences.²⁹ Nevertheless, several studies have demonstrated that cognitive restructuring can effectively reduce implicit offense-supportive beliefs and attitudes.³¹

2.2.6 Module 6: Sexuality

Individuals who have committed sexual offences have been shown to be more sexually active and more strongly interested in sexual topics than non-offending individuals.³² Sexual preoccupation is characterised by intrusive sexual thoughts and behaviours that are difficult or impossible to control (84) and has been shown to be strongly associated with sexual reoffending.¹⁵ Accordingly, a primary aim of the sexuality module is to provide basic knowledge about sexuality and to support the reduction of excessive sexual thoughts and behaviours. The module further aims to increase awareness of high-risk situations that may trigger sexual thoughts and behaviours. Participants are provided with strategies to manage and, where possible, reduce sexual preoccupation by limiting sexual thoughts and activities. Emotional congruence with children represents another central risk factor addressed in the sexuality module. ICCO under CS may experience a perceived emotional closeness to children and, when identifying themselves with children, may find it easier to initiate contact with them.³³ Emotional congruence with children is therefore considered a risk factor strongly associated with sexual reoffending.³⁴ The sexuality module seeks to reduce emotional congruence with children through psychoeducational approaches, emphasising that emotionally equal relationships between adults and children are not possible and providing strategies to establish greater emotional distance from children. Sexual interest in children, including sexual fantasies involving children, constitutes one of the most important risk factors for sexual reoffending among ICCO⁵ and has been described as a potential motivational factor underlying sexual abuse of children.³⁵ Reducing sexual interest in children while strengthening non-deviant sexual interests may therefore contribute to risk reduction. The extent to which sexual interests can change across the lifespan remains debated, and empirical evidence is currently insufficient to draw definitive conclusions.³⁶ Clinically, both changes in sexual interests and highly persistent pedophilic interests have been observed, supporting the conceptualisation of sexual interest in children as a dynamic risk factor that may change over time.³⁶ Within the sexuality module, cognitive-behavioural techniques are applied to influence sexual interest in and, where possible, sexual fantasies about children, while promoting sexual interest in adults. The techniques employed in this module have demonstrated effectiveness in in-person treatments and, at least among younger individuals, have also been applied in web-based interventions.³⁷⁻⁴⁰

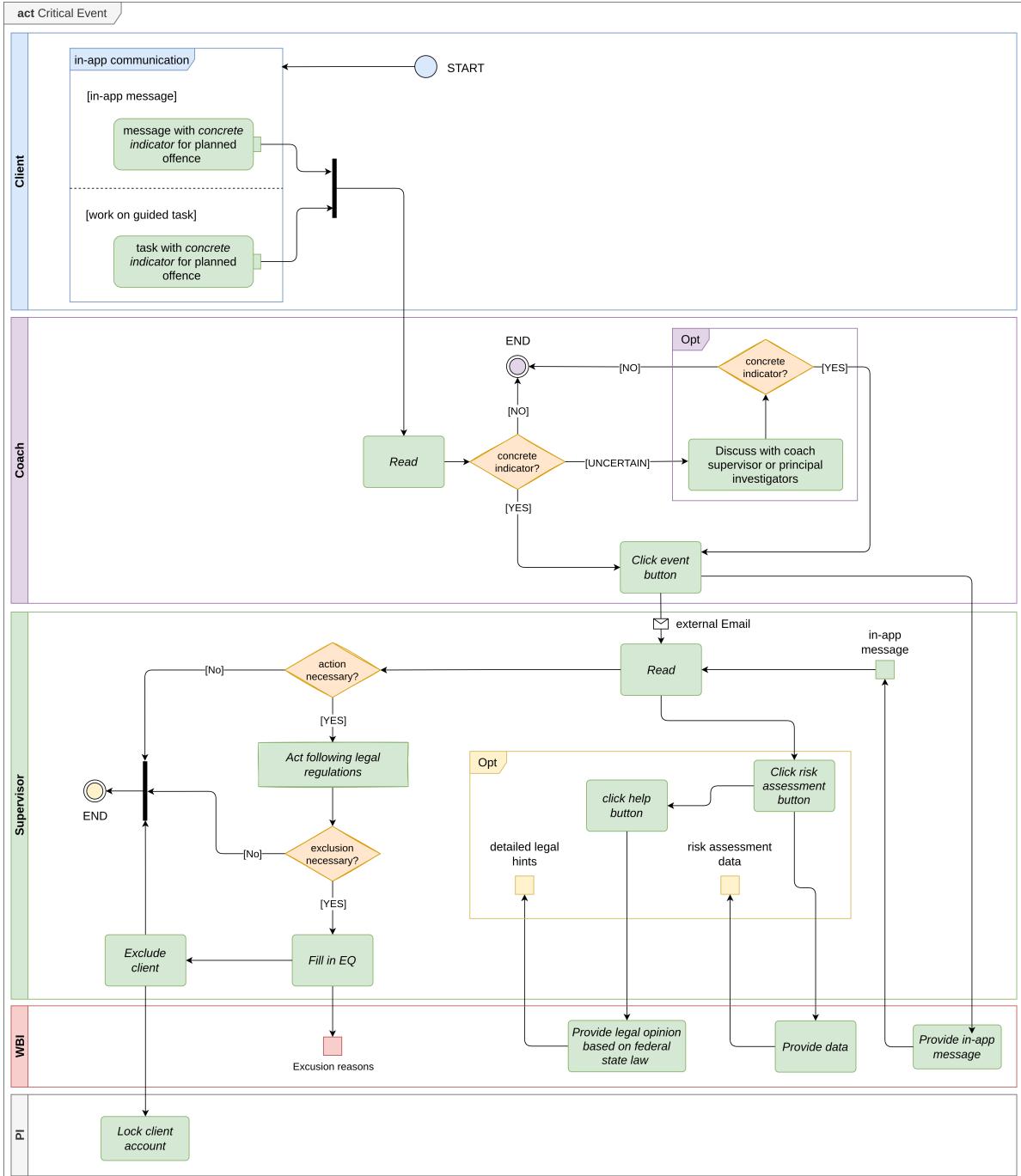


Figure 4: Action diagram for critical events. Predefined standard operating procedures have been used in order to ensure the security of society during the clinical trial. The figure shows exemplary the process for the occurrence of a critical event (hints on concrete preparations of a re-offence). Please note, no critical events occurred during the clinical trial.

3 Outcomes and measurement timepoints

3.1 Description of trial endpoints

Note that the following description of the primary and secondary outcome measures is adapted from¹.

3.1.1 SSQ-P

The Sample Specifications Questionnaire for the Participant (SSQ-P) is a self developed online questionnaire asking the participants about their demographics, lifestyle, current and former therapeutic treatments and offending history.

3.1.2 SSQ-SO

The Sample Specifications Questionnaire for the SO (SSQ-SO) is a self-developed online questionnaire for the supervision officer concerning the current living condition, offending and clinical history, and index offense of the participant.

3.1.3 SSQ-CF

The Sample Specifications Questionnaire based on the Court File (SSQ-CF) is a self-developed checklist for coding court files. Information from the court files are coded by the study investigators to determine whether the participant meets the study inclusion criteria. This includes information e.g on the index offense, and former offenses. In case that items of the SSQ-CF cannot be filled out adequately by utilizing court files, the participants was interviewed directly.

3.1.4 PHQ-D

The Patient Health Questionnaire (PHQ-D⁴¹) is a screening tool for mental disorders. Classification of results generally takes place on a syndrome level. Syndromes examined are: Somatoform disorders, major depressive syndrome, other depressive syndrome, panic syndrome, other anxiety syndromes, eating disorder syndrome, alcohol abuse. Tests are scored using a stencil. Utilizing a sum score for each syndrome, conclusions pertaining to the severity of each syndrome can be drawn. The measure shows good to very good diagnostic validity, especially regarding panic disorder and major depression⁴².

3.1.5 SSPI-2

The Revised Scale for Pedophilic Interests (SSPI-2⁴³) is a structured rating scale of assessing pedophilic interests based on the offending behavior. It comprises five items (number, age, gender, relationship of victims, and child pornography) and is significantly associated with phallometrically assessed sexual arousal to children. The SSPI-2 is derived from information of the participants as well as the supervision officer and is only defined for hands-on offenders⁴³.

3.1.6 ICD-11-Screener

The ICD-11-Screener⁴⁴ is a self-assessment procedure to ascertain indications for the presence of sexual dysfunction, gender dysphoria, or paraphilic disorder. However, the instrument is not suitable for diagnosis. It is assumed that the ICD-11-Screener can identify individuals who have the disorders described above. The ICD-11-Screener is not validated yet. Due to a mistake by the PI, this questionnaire was not mentioned in the SAP.

3.1.7 CARES

The Child Sexual Abuse Risk Evaluation Self-Report (CARES) is a composite instrument integrating the CARES-A, which assesses acute dynamic risk factors, and the CARES-S, which measures stable dynamic risk factors.⁴⁵ The CARES total score is calculated by summing the CARES-A and CARES-S total scores, producing a range from 0 to 102, with higher scores indicating greater severity of both acute and stable dynamic risk factors. The CARES-A scale evaluates seven acute dynamic risk factors associated with sexual reoffending in individuals convicted of at least one (contact or non-contact) sexual offense against children: (1) Preoccupation with Potential Victims, (2) Hostility, (3) Hypersexuality, (4) Rejection of Control Measures, (5) Emotional Crisis, (6) Loss of Social Support, and (7) Substance Use. Comprising 21 items, the CARES-A yields scores from 0 to 42; higher scores denote more pronounced acute dynamic risk. Reliability coefficients for the CARES-A range from $\alpha = .79$ to $\alpha = .83$. Analyses of social desirability revealed no significant correlations between CARES-A scores and either the exaggeration of positive qualities or negative qualities.⁴⁵ The CARES-S scale assesses six stable dynamic predictors of sexual recidivism. It consists of 30 items covering: (1) Problematic Sexual Interests, (2) Problematic Emotion-Regulation, (3) Lack of Problem-Solving Skills, (4) Offense-Supportive Cognitions, and (5) Lack of Social Support. Five of these subscales correspond to empirically supported psychological risk factors for sexual recidivism^{4,5}. The CARES-S total score ranges from 0 to 60; higher scores indicate stronger stable dynamic risk factors. The reduced CARES-S achieved an internal consistency of $\alpha = .53$. As with the CARES-A, no significant correlations with social desirability measures were found in the evaluation study.⁴⁵

3.1.8 Official re-offences

Officially registered re-offenses will be assessed five years after last-patient-out (September 2030).

3.1.9 CVTRQ

The Corrections Victoria Treatment Readiness Questionnaire (CVTRQ⁴⁶) is a self-report measure designed to assess treatment readiness in offenders who have been referred to a cognitive skills program. It comprises 20 items and consists of four scales: (1) Attitudes and motivation, (2) emotional reaction, (3) offending beliefs, and (4) efficacy. Each item has to be answered on a five-point Likert-scale. The CVTRQ shows an acceptable convergent validity, discriminant validity as well as predictive validity⁴⁶. To the best of our knowledge, there is no validated German version of the CVTRQ. @myTabu translated the CVTRQ. The translation was checked and translated back into English. Resulting differences due to adaptation were inspected and discussed in terms of item content.

3.1.10 RCQ

The Readiness to Change Questionnaire (RCQ⁴⁷) is a 12 item questionnaire originally designed to identify the stage of change reached by individuals who excessively drink alcohol. The German version (RCQ-D)⁴⁸ was adapted within the project by changing the questions regarding drinking of alcohol in questions regarding CSA and CSEM offenses. Responses are made on a five-point Likert-scale. The RCQ-D as well as the RCQ shows good psychometric properties.⁴⁸ Psychometric properties of the version adapted for ISAC and ICCSEM are not available yet.

3.1.11 OQMPR

The Questionnaire for the Measurement of Psychological Reactance (QMPR⁴⁹) is a questionnaire for the assessment of psychological reactance defined as the theory that people resist attempts to constrain either their thoughts or their behaviors. The Optimized Questionnaire for the Measurement of Psychological Reactance (OQMPR)⁵⁰ is a German variant of the QMPR. The OQMPR consists of 12 statements on a five-point Likert-scale. It has a test-retest-reliability of $\alpha = 0.85$.⁵⁰

3.1.12 F-SozU

Seven-item short version of the Social Support Questionnaire (F-SozU⁵¹) is an efficient questionnaire to assess perceived social support. Each item comprises a five-point Likertscale. Internal consistency of the original long version showed Cronbach's α between 0.81 and 0.93.⁵²

3.1.13 UCLA

The UCLA Loneliness Scale (UCLA⁵³) consists of 20 items in order to assess subjective feelings of loneliness. The German short version⁵⁴ consists of 12 items on a four-point Likert-scale. A German study with CSA and CSEM offenders demonstrated a high reliability of $\alpha = 0.92$.⁵⁵

3.1.14 DERS

The Difficulties in Emotion Regulation Scale (DERS⁵⁶) is a self-report questionnaire to assess emotion dysregulation. The sub-scale “impulse control difficulties” is used as outcome measure and comprises five items on a five-point Likert-scale. The German version is used, which shows a good internal consistency, construct and predictive validity.⁵⁷

3.1.15 NARQ

The Negative Affect Repair Questionnaire (NARQ⁵⁸) is a self-report questionnaire to assess strategies to regulate negative affect in a systematic manner. It consists of 17 items on a five-point Likert-scale and provides a good construct validity. Reliability scores (Cronbach's α) for the three NARQ scales ranged between 0.71 and 0.80.⁵⁸

3.1.16 BIS-15

The Barratt Impulsiveness Scale-15 (BIS-15⁵⁹) is a questionnaire developed to assess the personality/behavioral construct of impulsiveness. It consists of 30 items on a four-point Likert-scale. The German short version BIS-15⁶⁰ consisting of 15 items is used in the clinical trial. The BIS-15 is an efficient measure of impulsiveness with good internal consistency of $\alpha = 0.81$.⁶⁰

3.1.17 CUSI

The Coping Using Sex Inventory (CUSI⁶¹) is a questionnaire to assess the presence of and the degree to which sex was used to deal with problematic situations. It consists of 16 items on a five-point Likert-scale with a satisfying internal consistency.⁶¹ The translation was checked and translated back into English by the authors. Resulting differences due to adaptation were inspected and discussed in terms of item content.

3.1.18 SPSI-R

The Social Problem-Solving Inventory Revised (SPSI-R²⁴) is a self-report questionnaire for the assessment of the five dimensions in the social problem-solving model: (1) positive problem orientation, (2) negative problem orientation, (3) rational problem solving, (4) impulsivity/carelessness style, and (5) avoidance style. The SPSI-R consists of 52 items on a five-point Likert-scale. It shows good psychometric properties for individuals who have sexually offended²⁷. The validated German version is a short form consisting of 25 items⁶².

3.1.19 BMS

The Bumby Molest Scale (BMS⁶³) is a self-report questionnaire which assesses offense supportive cognitions of CSA or CSEM offenders (German Version:⁶⁴). It consists of 38 items on a four-point Likert-scale. The German version shows good construct validity, internal consistency, and test-retest reliability⁶⁴.

3.1.20 HBI-19

Sexual preoccupation is assessed by the Hypersexual Behavior Inventory-19 (HBI-19⁶⁵). The HBI-19 is a three-factor measure (coping, control, and consequences) developed to assess hypersexual behavior. The instrument consists of 19 items (e.g., 'I use sex to forget sorrows of everyday life.') answered on a scale from 1 (never) to 5 (very often). The maximum score is 95, with higher scores indicating a higher level of sexual preoccupation. The questionnaire was shown to have good reliability ($\alpha = 0.90$) and validity.⁶⁵

3.1.21 SSIC

The Specific self-efficacy for modifying Sexual Interest in Children (SSIC⁶⁶) comprises six items on the participant's conviction regarding the ability to change their sexual interest in children (e.g., 'I can succeed in reducing my sexual interest in children') were answered on a scale from 1 (do not agree at all) to 5 (totally agree). The maximum score is 30, with higher scores indicating a higher level of self-efficacy. The instrument was shown to have good reliability ($\alpha = 0.87$) and validity⁶⁶.

3.1.22 SOI-R

The Item 2a of the Sexual Outlet Inventory revised (SOI-R⁶⁷) assesses the desire for sexual activity involving children on a visual analog scale from 0 (desire is absent) to 100 ("I have to act to satisfy the desire")⁶⁷. Higher values on the scale indicate a stronger sexual interest in children.⁶⁷

3.1.23 EKK-R

Emotional congruence with children is assessed by the Questionnaire on Emotional Congruence with Children-Revised (EKK-R⁶⁸) including three factors (special relationship to children, immaturity, and emotional closeness to children). Twenty items are answered on a four-point scale. The questionnaire demonstrated good reliability ($\alpha = 0.80$) and validity.⁶⁸

3.1.24 ESIQ

The Explicit Sexual Interest Questionnaire (ESIQ⁶⁹) directly assesses pedophilic interest. It consists of two scales measuring sexual behavior (20 items, e.g., 'I enjoyed orally stimulating a man.') and sexual fantasies (20 items, e.g., 'I find it attractive to imagine a little boy sexually stimulating me.'). All items are answered on a scale from 1 (totally disagree) to 5 (totally agree). The reliability of the instrument ranges between 0.86 and 0.97⁶⁹.

3.1.25 WHO-5

The WHO-5 Well-Being Index (WHO-5⁷⁰) is a questionnaire that measures current mental well-being with five items. The items are rated on a six-point Likert-scale ranging from 0 (at no time) to 5 (all of the time). The WHO-5 has shown good validity in measuring subjective well-being in clinical studies⁷¹. The norm in Germany is mean 65.7. The threshold for a clinically relevant change is defined as a change of at least 10 points on the total score.⁷¹ Due to a mistake by the PI, this questionnaire was not mentioned in the SAP.

3.1.26 QNP

The Questionnaire of Non-Participation (QNP) was developed to assess reasons for not taking part in the study. It comprises eleven items on a six-point Likert-scale. Items are e.g., 'I will not take part in the study because I do not need any further help.' or 'I will not take part in the study because I am afraid of negative consequences if my SO gets additional information about me.' The QNP will only be administered if a potential participant will not take part in the study but has given informed consent to

fill out this questionnaire. Due to a mistake by the PI, this questionnaire was not mentioned in the SAP.

3.1.27 MRQ

The Mood and Risk Questionnaire (MRQ) consists of ten self-developed items, of which six are asking about thoughts and behaviors associated with potential re-offenses and four asking about the psychological and emotional state of the participant. The MRQ was developed by the @myTabu team in order to assess adverse events in a structured manner and with adapted adverse events for individuals who have a history of CSA or CSEM. The MRQ has been assessed before and after each session of the WBI and differs between the items in the pre and post measurements: all items of the pre measurement refer explicitly to the time between finishing last post measurement and before starting the new session (time not working on the WBI). The post measurement explicitly asks for the time during working on the session. The MRQ has the following one-item subscales: (1) Contact planning (“...have you thought about how to best make contact with a child?”); (2) Contact preparation (“...have you made preparations to be able to make contact with a child?”); (3) Urge for CSA (“...have you felt that you must commit a sexual act with a child?”); (4) Urge for CSEM (“...have you felt that you might soon (again) watch child sexual abuse material?”); (5) Sexual tension (“...have you felt that sexual tension has built up within you?”); (6) Control of sexual thoughts and activity (“...have you felt that it is difficult for you to control your sexual thoughts and activities?”); (7) Unable to cope with the mental burden (“...have you felt that you can no longer endure burdens?”); (8) Suicidal ideations (“...have you thought about taking your own life?”); (9) Mental crisis (“At present, I feel that my behavior and experience are impaired due to mental problems.”); (10) Very bad negative mood (“How do you feel at the moment?”). An adverse event was defined as the highest possible score on each one-item self-report subscale ('Often' for subscales 1–8, 'Strongly agree' for subscale 9, and 'Very bad' for subscale 10). The MRQ post questionnaire explicitly asks for the time during working on the WBI, whereas the MRQ pre questionnaire exclusively asks for the last week before measurement timepoint. Please note, that the occurrence of a report of the highest possible score was defined as adverse event post-hoc. Due to a mistake by the PI, this questionnaire was not mentioned in the SAP.

3.1.28 WAI-SR

The Working Alliance Inventory-Short Revised (WAI-SR^{72,73}) measures three dimensions of therapeutic alliance: (a) bond, (b) task, and (c) goal. For the purpose of this study, the original items of the German version⁷² were adapted to be suitable for a web-based intervention. The items have to be rated on a five-point Likertscale ranging from 1 (never) to 5 (always). It was adapted to suit to the online-coach (WAI-SR COACH) as well as the supervision officer (WAI-SR SUPERVISOR) resulting in 16 items. Internal consistency (Cronbach's α) of the original questionnaire ranges from 0.82 to 0.90⁷². Due to a mistake by the PI, this questionnaire was not mentioned in the SAP.

3.1.29 EQ

The Exclusion Questionnaire (EQ) is a questionnaire to assess the exclusion reasons of the supervision officer. The EQ was specifically developed for this clinical trial by the project team. Only supervision officers were able to exclude a participant from the trial by clicking a button in the WBI. After clicking the button, the supervision officer had to fill in the EQ before the PI was informed to exclude the

participant by Email. The EQ asks about eight possible exclusion reasons (e.g. re-offences, withdraw of informed consent) and one additional free text from. Multiple answers on the eight exclusion reasons are possible. Due to a mistake by the PI, this questionnaire was not mentioned in the SAP.

Table 1: Measurement timepoints for all trial endpoints. Please note, that the Mood and Risk Questionnaire (MRQ) is not listed in the table. The MRQ Pre was assessed before each session, the MRQ Post after each session. The measurement time-points of the WAI-SR COACH and WAI-SR SUPERVISOR have been administered before session two in each module. Please see Fromberger et al.¹ for table of measurement timepoints with a higher time resolution.

	Allocation	Introduction		Module 1		Module 2		Module 3		Module 4		Module 5		Module 6		Follow-Up
		pre	post	pre	post	pre	post	pre	post	pre	post	pre	post	pre	post	
BASELINE CHARACTERISTICS																
SSQ-P		X														
SSQ-SO																
SSQ-CF	X															
PHQ-D		X														
SSPI-2	X															
ICD-11-Screener		X														
CARES				X		X		X		X		X		X		X
PRIMARY OUTCOMES																
Official re-offences																X
CVTRQ	X			X	X											
SECONDARY OUTCOMES																
RCQ			X	X												
OQMPR					X	X										
F-SozU					X	X										
UCLA					X	X										
DERS							X	X								
NARQ							X	X								
BIS-15							X	X								
CUSI							X	X								
SPSI-R									X	X						
BMS									X	X						
HBI-19											X	X				
SSIC											X	X				
SOI-R													X	X		
EKK-R													X	X		
ESIQ																
SBV-R		X														
ANCILLARY OUTCOMES																
ATQ	X															
WHO-5			X	X	X	X	X	X	X	X	X	X	X	X	X	
QNP	X															
WAI-SR COACH			X		X		X		X		X		X		X	
WAI-SR SUPERVISOR			X		X		X		X		X		X		X	

Abbreviations: BIS-15 = Barratt Impulsiveness Scale-15, BMS = Bumby Molest Scale, CUSI = Coping Using Sex Inventory, CVTRQ = Corrections Victoria Treatment Readiness Questionnaire, DERS = Difficulties in Emotion Regulation Scale, EKK-R = Questionnaire on Emotional Congruence with Children-Revised, ESIQ = Explicit Sexual Interest Questionnaire, F-Soz-U = Seven-item short version of the Social Support Questionnaire, HBI-19 = Hypersexual Behavior Inventory-19, NARQ = Negative Affect Repair Questionnaire, OQMPR = Questionnaire for the Measurement of Psychological Reactance, RCQ = Readiness to Change Questionnaire - German version, SOI-R = Sexual Outlet Inventory revised, subscale desire for sexual activity with children, SPSI-R = Social Problem-Solving Inventory Revised, SSIC = Specific self-efficacy for modifying Sexual Interest in Children, UCLA = UCLA Loneliness Scale - German short version.

Table 2: Baseline values of the SSPI-2. Note, that the SSPI-2 is only defined for participants who have at least one hands-on offences.⁴³ Abbreviations: Q1 = 25th percentile, Q3 = 75th percentile.

Characteristic	Overall N = 221 ¹	Intervention N = 108 ¹	Placebo N = 113 ¹
Any hands-off offences			
Yes	186 (84%)	90 (83%)	96 (85%)
No	35 (16%)	18 (17%)	17 (15%)
More than one hands-on victims			
Yes	43 (45%)	20 (43%)	23 (46%)
No	53 (55%)	26 (57%)	27 (54%)
Missing	125	62	63
Any hands-on victims under 15 years old			
Yes	43 (44%)	21 (44%)	22 (45%)
No	54 (56%)	27 (56%)	27 (55%)
Missing	124	60	64
Any hands-on victims under 12 years old			
Yes	68 (72%)	31 (66%)	37 (77%)
No	27 (28%)	16 (34%)	11 (23%)
Missing	126	61	65
Any extra-familiar hands-on victims			
Yes	75 (77%)	35 (74%)	40 (80%)
No	22 (23%)	12 (26%)	10 (20%)
Missing	124	61	63
SSPI-2 score			
Median (Q1 – Q3)	3 · 00 (2 · 00 – 4 · 00)	3 · 00 (2 · 00 – 4 · 00)	3 · 00 (2 · 00 – 4 · 00)
Missing	136	66	70

¹ n (%)

4 Baseline characteristics

4.1 SSPI-2

5 Timing (protocol deviations)

5.1 Cumulative time from baseline to finish module

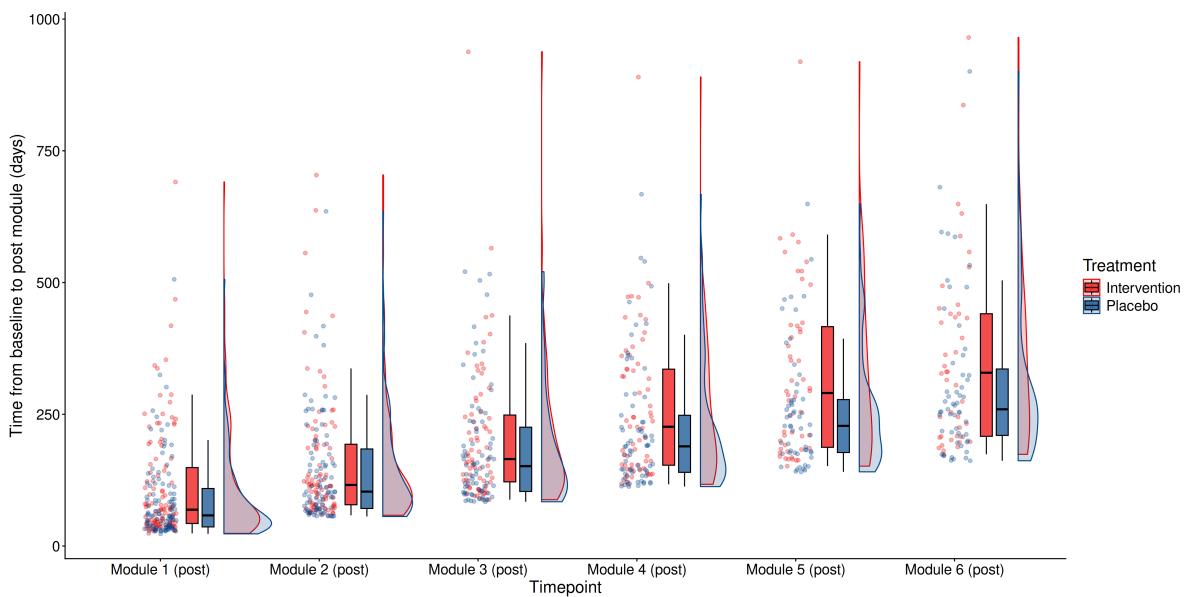


Figure 5: Cumulative time from baseline to finish module. The cumulative time from baseline to finish module is defined as time from beginning of module one to the end of the respective module. Note, that the cumulative time from baseline to finish module consist of the time of working on the content and time for filling in questionnaires and time needed to revise guided tasks (and an undefined error term for doing things not related to the WBI).

Table 3: Mixed model for repeated measures (MMRM) for the cumulative time from baseline to finish module. The cumulative time from baseline to finish module is defined as time from beginning of module one to the end of the respective module. Note, that the cumulative time from baseline to finish module consist of the time of working on the content and time for filling in questionnaires and time needed to revise guided tasks (and an undefined error term for doing things not related to the WBI).

Variable	Beta	95% CI	p-value
(Intercept)	87	69 to 105	<0 · 0001
Timepoint * treatment			
Module 1 (post) * Intervention	27	1 · 9 to 53	0 · 035
Module 2 (post) * Intervention	44	3 · 3 to 85	0 · 034
Module 3 (post) * Intervention	77	13 to 141	0 · 019
Module 4 (post) * Intervention	113	30 to 195	0 · 0078
Module 5 (post) * Intervention	126	30 to 223	0 · 011
Module 6 (post) * Intervention	132	29 to 236	0 · 013
Timepoint			
Module 1 (post)	0 · 00	Ref.	
Module 2 (post)	69	55 to 83	<0 · 0001
Module 3 (post)	149	118 to 180	<0 · 0001
Module 4 (post)	217	172 to 262	<0 · 0001
Module 5 (post)	289	235 to 344	<0 · 0001
Module 6 (post)	333	273 to 393	<0 · 0001

Abbreviations: CI = Confidence Interval, StGB = German penalty law

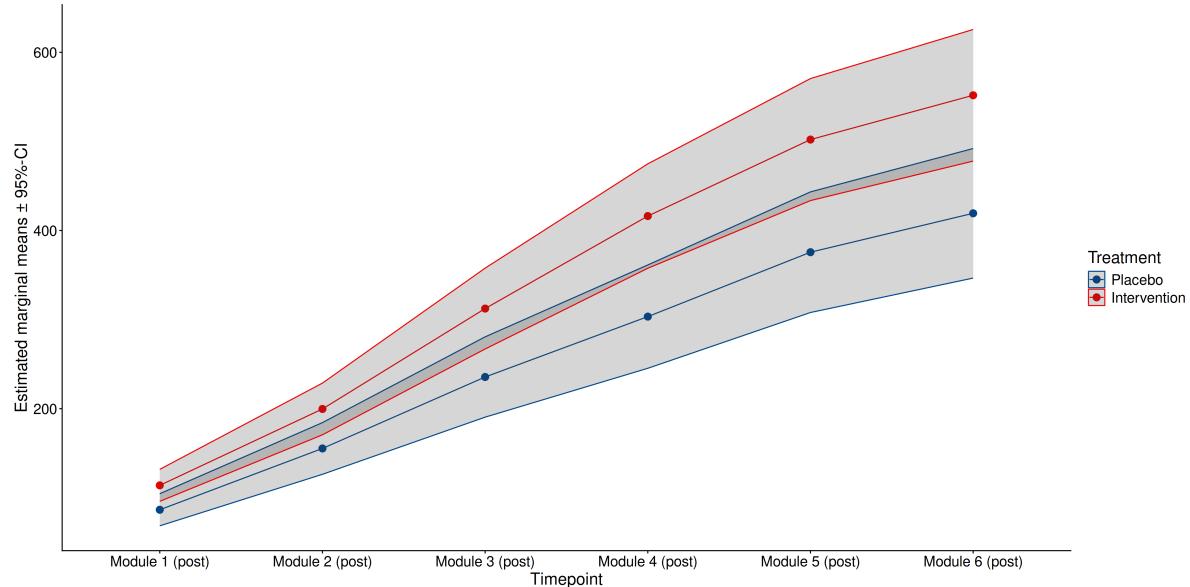


Figure 6: Estimated marginal means of the MMRM for the cumulative time from baseline to finish module. The cumulative time from baseline to finish module is defined as time from beginning of module one to the end of the respective module. Note, that the cumulative time from baseline to finish module consist of the time of working on the content and time for filling in questionnaires and time needed to revise guided tasks (and an undefined error term for doing things not related to the WBI).

Table 4: Pairwise comparisons for the cumulative time from baseline to finish module is defined as time from beginning of module one to the end of the respective module. Note, that the cumulative time from baseline to finish module consist of the time of working on the content and time for filling in questionnaires and time needed to revise guided tasks (and an undefined error term for doing things not related to the WBI).

Characteristic	Intervention N = 108	Placebo N = 113	p	q-value
Module 1 (advised: 28 days)			0 · 053	0 · 24
Mean (SD)	114 (110)	87 (76)		
Median (Q1, Q3)	69 (43, 151)	58 (36, 112)		
Min, Max	24, 691	23, 506		
Missing	2	7		
Module 2 (advised: 56 days)			0 · 25	0 · 30
Mean (SD)	162 (126)	145 (104)		
Median (Q1, Q3)	116 (78, 193)	103 (71, 192)		
Min, Max	58, 704	56, 635		
Missing	23	22		
Module 3 (advised: 84 days)			0 · 10	0 · 30
Mean (SD)	210 (135)	185 (111)		
Median (Q1, Q3)	165 (121, 250)	152 (103, 229)		
Min, Max	88, 938	84, 521		
Missing	38	37		
Module 4 (advised: 112 days)			0 · 048	0 · 24
Mean (SD)	260 (138)	223 (122)		
Median (Q1, Q3)	226 (153, 336)	189 (140, 248)		
Min, Max	117, 890	113, 667		
Missing	46	44		
Module 5 (advised: 140 days)			0 · 028	0 · 17
Mean (SD)	325 (161)	258 (113)		
Median (Q1, Q3)	290 (186, 418)	228 (177, 279)		
Min, Max	152, 919	141, 649		
Missing	58	58		
Module 6 (advised: 168 days)			0 · 13	0 · 30
Mean (SD)	363 (181)	309 (152)		
Median (Q1, Q3)	329 (208, 441)	259 (208, 347)		
Min, Max	174, 965	162, 901		
Missing	66	58		

Abbreviation: Q1 = 25th percentile, Q3 = 75th percentile, p_{adj} = Holm-Bonferroni adjusted p.

5.2 Time working on module

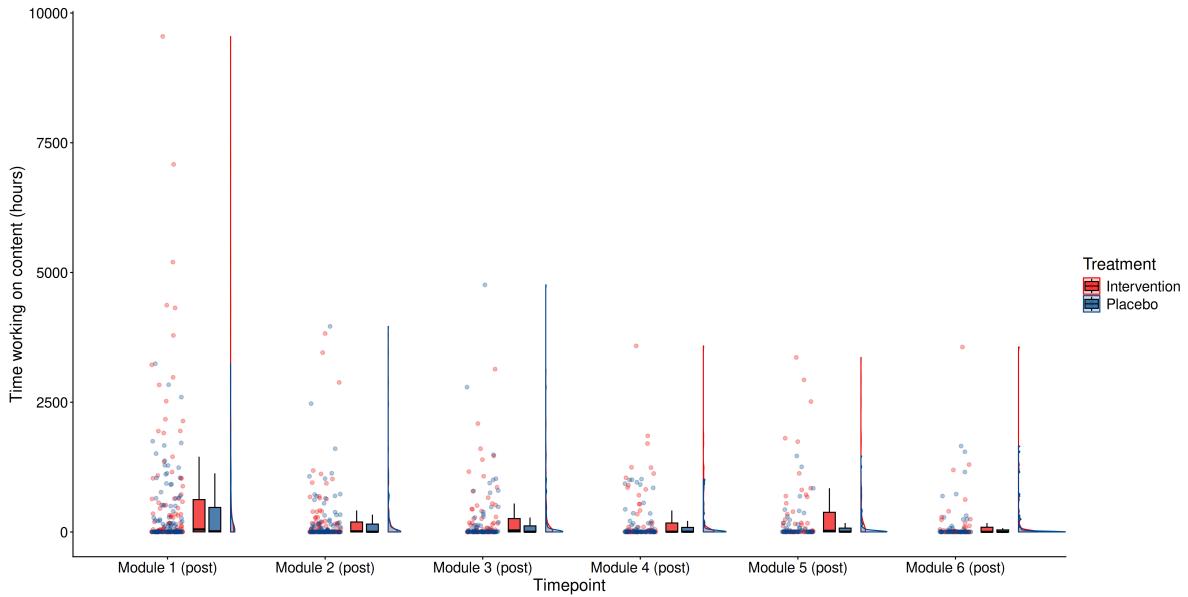


Figure 7: Rainplot for time working on content by module. The time working on content is defined as the sum of the time from creating a data base entry for lection data (at first visit of the lection page) to the time last change of database entry for this lection data occurred for each lection the participant worked on in one module. This time difference can be interpreted as a narrow estimator for the time needed to finish the content of a WBI without the time needed for filling in questionnaires, reviews of guided tasks, or waiting times. But it includes the time needed for first answer of a guided exercise and times, in which the participant not logged in or logged in but not worked on the lection.

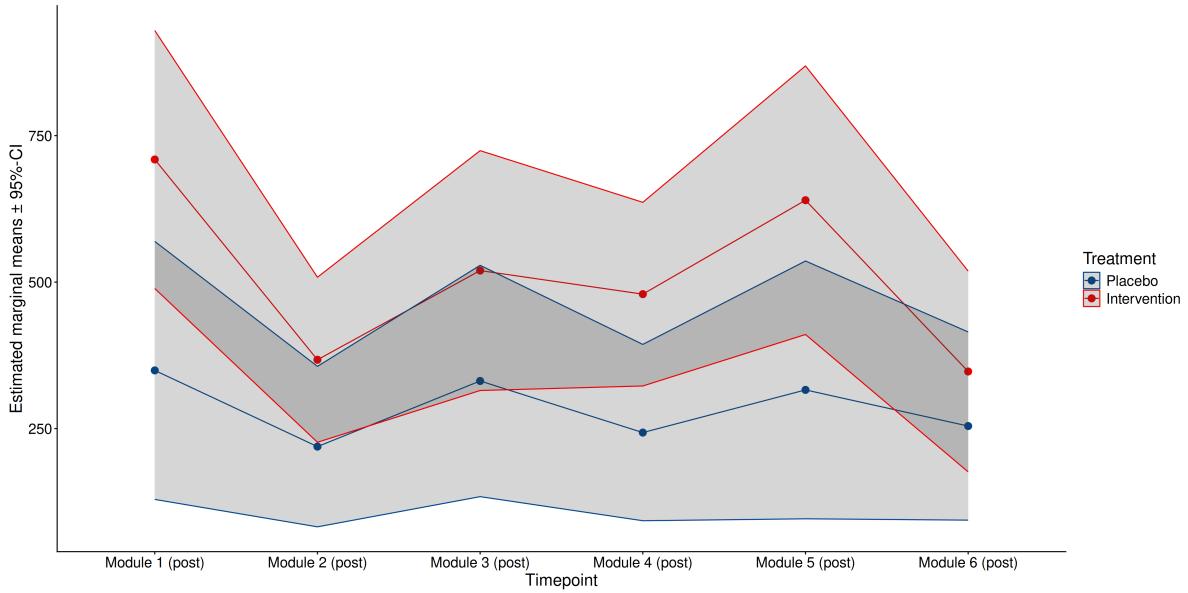


Figure 8: Estimated marginal means for time working on content of one module of the WBI. The time working on content is defined as the sum of the time from creating a data base entry for lection data (at first visit of the lection page) to the time last change of database entry for this lection data occurred for each lection the participant worked on in one module. This time difference can be interpreted as a narrow estimator for the time needed to finish the content of a WBI without the time needed for filling in questionnaires, reviews of guided tasks, or waiting times. But it includes the time needed for first answer of a guided exercise and times, in which the participant not logged in or logged in but not worked on the lection.

Table 5: Mixed model for repeated measures (MMRM) for time working on content of one module of the WBI. The time working on content is defined as the sum of the time from creating a data base entry for lection data (at first visit of the lection page) to the time last change of database entry for this lection data occurred for each lection the participant worked on in one module. This time difference can be interpreted as a narrow estimator for the time needed to finish the content of a WBI without the time needed for filling in questionnaires, reviews of guided tasks, or waiting times. But it includes the time needed for first answer of a guided exercise and times, in which the participant not logged in or logged in but not worked on the lection.

Variable	Beta	95% CI	p-value
(Intercept)	349	129 to 570	0 · 0020
Timepoint * treatment			
Module 1 (post) * Intervention	360	49 to 671	0 · 024
Module 2 (post) * Intervention	148	-48 to 345	0 · 14
Module 3 (post) * Intervention	188	-96 to 473	0 · 19
Module 4 (post) * Intervention	236	19 to 453	0 · 034
Module 5 (post) * Intervention	324	6 · 1 to 641	0 · 046
Module 6 (post) * Intervention	93	-142 to 328	0 · 42
Timepoint			
Module 1 (post)	0 · 00	Ref.	
Module 2 (post)	-130	-317 to 57	0 · 17
Module 3 (post)	-18	-255 to 219	0 · 88
Module 4 (post)	-106	-327 to 115	0 · 34
Module 5 (post)	-33	-304 to 237	0 · 81
Module 6 (post)	-95	-327 to 137	0 · 42

Abbreviation: CI = Confidence Interval, StGB = German penalty law.

Table 6: Pairwise comparisons for time working on content by module. The time working on content is defined as the sum of the time from creating a data base entry for lection data (at first visit of a lection page) to the time last change of database entry for this lection data occurred for each lection the participant worked on in one module. This time difference can be interpreted as a narrow estimator for the time needed to finish the content of a WBI without the time needed for filling in questionnaires, reviews of guided tasks, or waiting times. But it includes the time needed for first answer of a guided exercise and times, in which the participant not logged in or logged in but not worked on the lection.

Characteristic	Intervention N = 108	Placebo N = 113	p	q-value
Module 1 (pre-tests: about 4 hours)			0 · 18	>0 · 99
Mean (SD)	709 (1 500)	349 (629)		
Median (Q1, Q3)	49 (2, 637)	12 (2, 498)		
Min, Max	1, 9 549	0, 3 242		
Missing	2	7		
Module 2 (pre-tests: about 4 hours)			0 · 77	>0 · 99
Mean (SD)	273 (668)	218 (553)		
Median (Q1, Q3)	11 (1, 193)	4 (1, 157)		
Min, Max	0, 3 825	0, 3 962		
Missing	24	22		
Module 3 (pre-tests: about 4 hours)			0 · 46	>0 · 99
Mean (SD)	283 (561)	237 (680)		
Median (Q1, Q3)	29 (2, 258)	4 (2, 120)		
Min, Max	1, 3 137	0, 4 762		
Missing	39	38		
Module 4 (pre-tests: about 4 hours)			0 · 89	>0 · 99
Mean (SD)	293 (620)	145 (297)		
Median (Q1, Q3)	4 (1, 173)	4 (2, 89)		
Min, Max	1, 3 586	0, 1 019		
Missing	47	45		
Module 5 (pre-tests: about 4 hours)			0 · 25	>0 · 99
Mean (SD)	398 (775)	126 (297)		
Median (Q1, Q3)	22 (2, 391)	5 (3, 76)		
Min, Max	1, 3 364	1, 1 464		
Missing	58	58		
Module 6 (pre-tests: about 4 hours)			0 · 95	>0 · 99
Mean (SD)	202 (603)	139 (361)		
Median (Q1, Q3)	2 (1, 93)	2 (1, 40)		
Min, Max	1, 3 564	0, 1 655		
Missing	66	60		

Abbreviation: Q1 = 25th percentile, Q3 = 75th percentile, p_{adj} = Holm-Bonferroni adjusted p.

6 Primary Outcomes

6.1 CARES

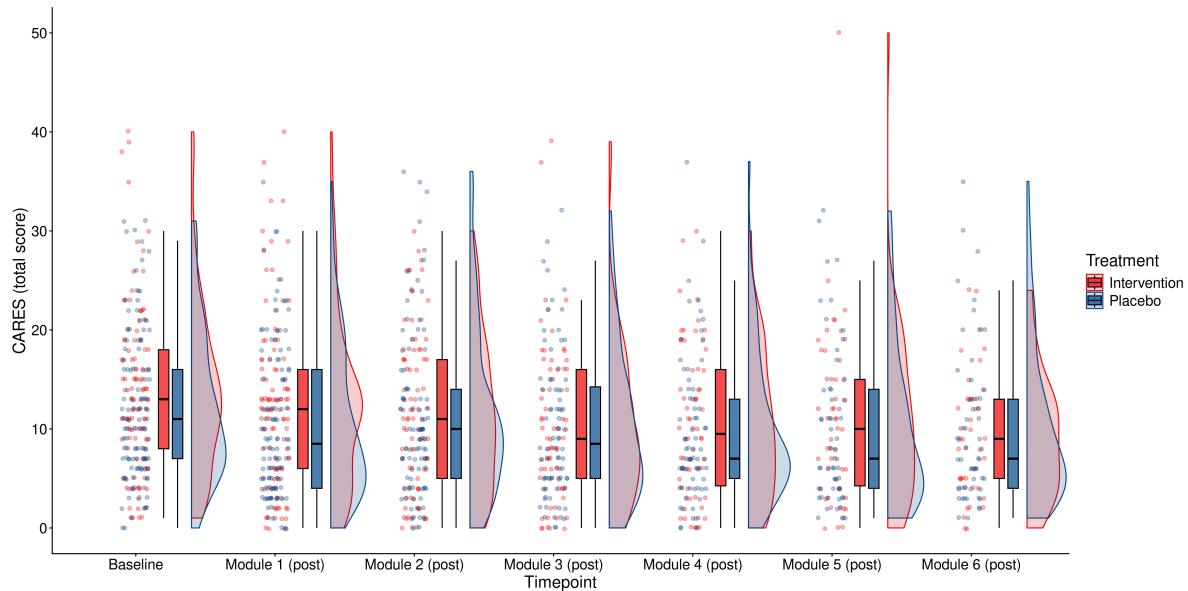


Figure 9: Rainplot of the primary endpoint Child Sexual Abuse Risk Evaluation Self-Report (CARES total score).

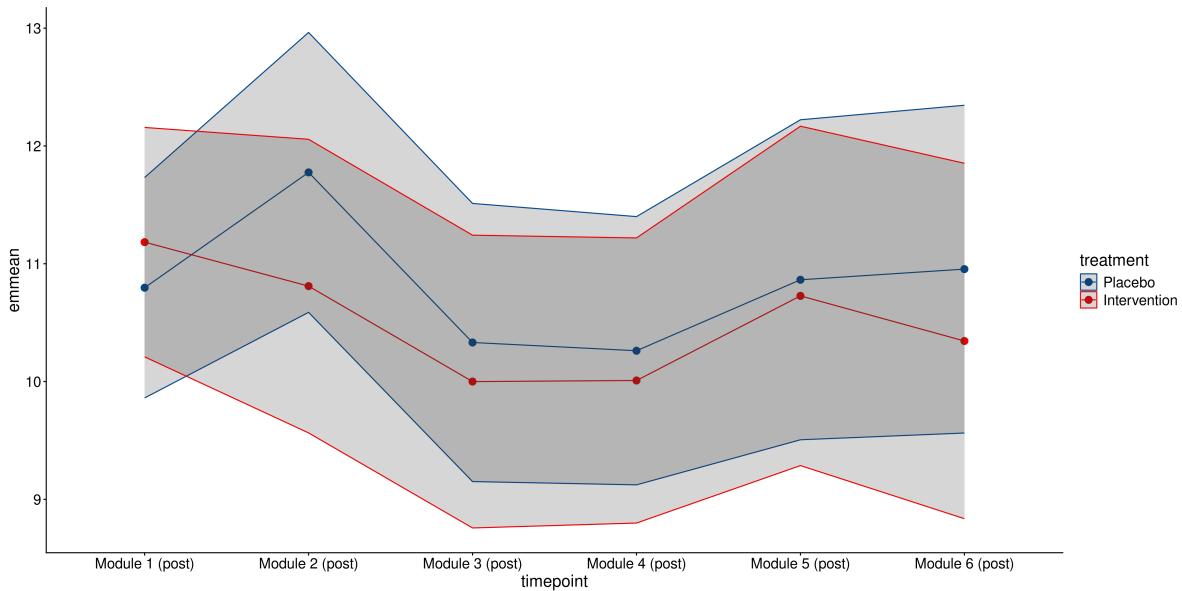


Figure 10: Estimated marginal means of the MMRM for the primary endpoint Child Sexual Abuse Risk Evaluation Self-Report (CARES total score).

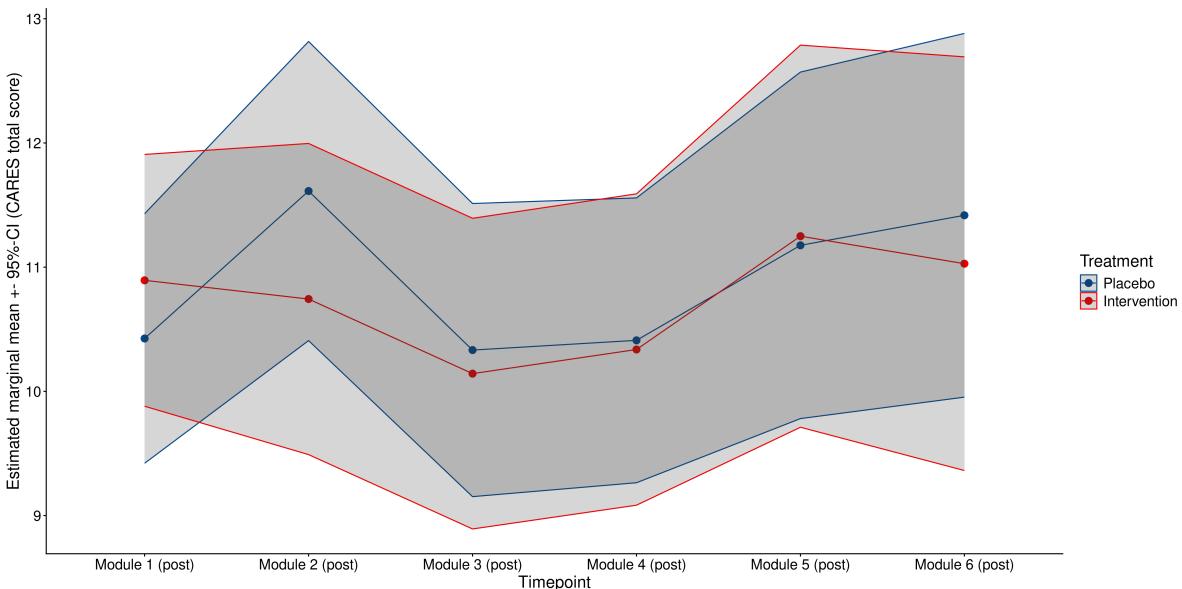


Figure 11: Estimated marginal means of the MMRM for the primary endpoint Child Sexual Abuse Risk Evaluation Self-Report (CARES total score) with time since baseline as additional factor.

Table 7: Explorative MMRM for the primary endpoint Child Sexual Abuse Risk Evaluation Self-Report (CARES total score) with time since baseline as additional factor. Time since baseline is defined as the time between the baseline measurement and the respective consecutive measurement time-point.

Variable	Beta	95% CI	p-value
(Intercept)	-0 · 09	-1 · 7 to 1 · 5	0 · 91
Type of supervision			
community supervision	0 · 00	Ref.	
post-release supervision	0 · 31	-1 · 2 to 1 · 8	0 · 68
Additional treatment			
No	0 · 00	Ref.	
Yes	-1 · 0	-2 · 2 to 0 · 19	0 · 10
Offense type			
Hands-off (only §184b StGB)	0 · 00	Ref.	
Hands-on (at least one conviction §176 ff StGB)	0 · 99	-0 · 24 to 2 · 2	0 · 11
Baseline value	0 · 84	0 · 77 to 0 · 91	<0 · 0001
Static recidivism risk	0 · 40	-0 · 11 to 0 · 90	0 · 12
Timepoint * treatment			
Module 1 (post) * Intervention	0 · 47	-0 · 75 to 1 · 7	0 · 45
Module 2 (post) * Intervention	-0 · 87	-2 · 5 to 0 · 76	0 · 29
Module 3 (post) * Intervention	-0 · 19	-1 · 8 to 1 · 4	0 · 82
Module 4 (post) * Intervention	-0 · 07	-1 · 6 to 1 · 5	0 · 93
Module 5 (post) * Intervention	0 · 07	-1 · 8 to 2 · 0	0 · 94
Module 6 (post) * Intervention	-0 · 39	-2 · 4 to 1 · 6	0 · 70
Timepoint			
Module 1 (post)	0 · 00	Ref.	
Module 2 (post)	1 · 2	0 · 17 to 2 · 2	0 · 022
Module 3 (post)	-0 · 09	-1 · 1 to 0 · 94	0 · 86
Module 4 (post)	-0 · 01	-1 · 2 to 1 · 2	0 · 98
Module 5 (post)	0 · 75	-0 · 65 to 2 · 1	0 · 29
Module 6 (post)	0 · 99	-0 · 41 to 2 · 4	0 · 17
Time since baseline (years)	-1 · 2	-2 · 4 to 0 · 01	0 · 051

Abbreviations: CI = Confidence Interval, StGB = German penalty law

Table 8: Sensitivity analysis for primary endpoint Child Sexual Abuse Risk Evaluation Self-Report (CARES total score). The table summarizes the results of a reference-based multiple imputation followed by ANCOVA models for the difference in the CARES compared to baseline.

Parameter	Estimate	95% CI	p
Group differences Module 1 (post)	-0.90	(-2.21, 1.06)	0.385
Least square mean [intervention] Module 1 (post)	24.20	(22.07, 25.21)	*< 0.001*
Least square mean [placebo] Module 1 (post)	23.30	(21.28, 24.98)	*< 0.001*
Group differences Module 2 (post)	0.46	(-1.35, 2.91)	0.994
Least square mean [intervention] Module 2 (post)	22.89	(20.56, 24.74)	*< 0.001*
Least square mean [placebo] Module 2 (post)	23.35	(20.23, 25.45)	*< 0.001*
Group differences Module 3 (post)	0.88	(-1.08, 3.53)	0.169
Least square mean [intervention] Module 3 (post)	20.30	(17.36, 22.49)	*< 0.001*
Least square mean [placebo] Module 3 (post)	21.18	(19.1, 23.89)	*< 0.001*
Group differences Module 4 (post)	0.41	(-1.6, 3.04)	0.964
Least square mean [intervention] Module 4 (post)	20.09	(17.56, 21.86)	*< 0.001*
Least square mean [placebo] Module 4 (post)	20.50	(18, 22.75)	*< 0.001*
Group differences Module 5 (post)	0.36	(-2.65, 3.61)	0.832
Least square mean [intervention] Module 5 (post)	20.53	(15.88, 22.31)	*< 0.001*
Least square mean [placebo] Module 5 (post)	20.89	(17.85, 23.46)	*< 0.001*
Group differences Module 6 (post)	2.27	(-0.81, 4.84)	0.172
Least square mean [intervention] Module 6 (post)	18.09	(15.81, 19.75)	*< 0.001*
Least square mean [placebo] Module 6 (post)	20.35	(18.22, 22.75)	*< 0.001*

Table 9: Pairwise comparisons of the primary endpoint Child Sexual Abuse Risk Evaluation Self-Report (CARES total score) for the intervention arm.

Module	pre N = 106	post N = 106	p	q-value
Module 1 (post)			0 · 0050	0 · 030
Mean (SD)	14 (8)	12 (9)		
Median (Q1, Q3)	13 (8, 18)	12 (6, 16)		
Min, Max	1, 40	0, 40		
Module 2 (post)			0 · 58	>0 · 99
Mean (SD)	12 (8)	12 (8)		
Median (Q1, Q3)	12 (6, 15)	11 (5, 17)		
Min, Max	0, 40	0, 30		
Missing	22	22		
Module 3 (post)			0 · 010	0 · 050
Mean (SD)	11 (8)	10 (8)		
Median (Q1, Q3)	11 (4, 17)	9 (5, 16)		
Min, Max	0, 28	0, 39		
Missing	38	38		
Module 4 (post)			0 · 89	>0 · 99
Mean (SD)	11 (8)	11 (8)		
Median (Q1, Q3)	10 (6, 16)	10 (5, 16)		
Min, Max	0, 39	0, 30		
Missing	46	46		
Module 5 (post)			0 · 44	>0 · 99
Mean (SD)	10 (7)	11 (9)		
Median (Q1, Q3)	9 (4, 14)	10 (4, 15)		
Min, Max	0, 30	0, 50		
Missing	56	56		
Module 6 (post)			>0 · 99	>0 · 99
Mean (SD)	10 (7)	9 (6)		
Median (Q1, Q3)	8 (4, 15)	9 (5, 13)		
Min, Max	0, 25	0, 24		
Missing	64	64		

Abbreviation: Q1 = 25th percentile, Q3 = 75th percentile, p_{adj} = Holm-Bonferroni adjusted p.

Table 10: Pairwise comparisons of the primary endpoint Child Sexual Abuse Risk Evaluation Self-Report (CARES total score) for the placebo arm.

Module	pre N = 106	post N = 106	p	q-value
Module 1 (post)			<0 · 0001	<0 · 0001
Mean (SD)	12 (7)	11 (8)		
Median (Q1, Q3)	11 (7, 16)	9 (4, 16)		
Min, Max	0, 31	0, 35		
Module 2 (post)			0 · 050	0 · 21
Mean (SD)	10 (8)	11 (8)		
Median (Q1, Q3)	8 (4, 16)	10 (5, 14)		
Min, Max	0, 35	0, 36		
Missing	15	15		
Module 3 (post)			0 · 041	0 · 21
Mean (SD)	12 (9)	10 (8)		
Median (Q1, Q3)	10 (5, 18)	8 (5, 14)		
Min, Max	0, 36	0, 32		
Missing	31	31		
Module 4 (post)			0 · 98	>0 · 99
Mean (SD)	10 (7)	10 (7)		
Median (Q1, Q3)	8 (5, 14)	8 (5, 14)		
Min, Max	0, 29	0, 37		
Missing	38	38		
Module 5 (post)			0 · 78	>0 · 99
Mean (SD)	9 (8)	10 (8)		
Median (Q1, Q3)	7 (4, 13)	7 (4, 14)		
Min, Max	0, 37	1, 32		
Missing	51	51		
Module 6 (post)			0 · 62	>0 · 99
Mean (SD)	10 (8)	10 (8)		
Median (Q1, Q3)	7 (4, 14)	7 (4, 13)		
Min, Max	1, 32	1, 35		
Missing	53	53		

Abbreviation: Q1 = 25th percentile, Q3 = 75th percentile, p_{adj} = Holm-Bonferroni adjusted p.

Table 11: Pairwise comparisons of difference from baseline of the primary endpoint Child Sexual Abuse Risk Evaluation Self-Report (CARES total score).

Characteristic	Placebo N = 106	Intervention N = 106	p	q-value
Module 1 (post)				>0 · 99
Mean (SD)	-1 · 5 (4 · 2)	-1 · 3 (5 · 0)	0 · 68	
Median (Q1, Q3)	-1 · 0 (-4 · 0, 1 · 0)	-1 · 0 (-4 · 0, 2 · 0)		
Min, Max	-13 · 0, 12 · 0	-21 · 0, 15 · 0		
Module 2 (post)			0 · 052	0 · 31
Mean (SD)	1 · 0 (5 · 5)	-0 · 3 (3 · 9)		
Median (Q1, Q3)	1 · 0 (-2 · 0, 4 · 0)	-1 · 0 (-2 · 0, 2 · 0)		
Min, Max	-17 · 0, 32 · 0	-14 · 0, 8 · 0		
Missing	15	22		
Module 3 (post)			0 · 51	>0 · 99
Mean (SD)	-1 · 5 (5 · 5)	-0 · 8 (4 · 1)		
Median (Q1, Q3)	-1 · 0 (-3 · 0, 1 · 0)	-1 · 0 (-3 · 0, 1 · 0)		
Min, Max	-34 · 0, 11 · 0	-13 · 0, 12 · 0		
Missing	31	38		
Module 4 (post)			0 · 98	>0 · 99
Mean (SD)	0 · 0 (4 · 3)	-0 · 2 (3 · 5)		
Median (Q1, Q3)	0 · 0 (-2 · 0, 2 · 0)	0 · 0 (-2 · 0, 2 · 0)		
Min, Max	-11 · 0, 13 · 0	-9 · 0, 9 · 0		
Missing	38	46		
Module 5 (post)			0 · 86	>0 · 99
Mean (SD)	0 · 6 (4 · 9)	0 · 7 (4 · 8)		
Median (Q1, Q3)	1 · 0 (-3 · 0, 3 · 0)	0 · 0 (-2 · 0, 2 · 0)		
Min, Max	-14 · 0, 16 · 0	-10 · 0, 21 · 0		
Missing	51	56		
Module 6 (post)			0 · 85	>0 · 99
Mean (SD)	0 · 21 (3 · 01)	-0 · 17 (3 · 88)		
Median (Q1, Q3)	0 · 00 (-1 · 00, 2 · 00)	0 · 00 (-2 · 00, 2 · 00)		
Min, Max	-9 · 00, 6 · 00	-16 · 00, 9 · 00		
Missing	53	64		

Abbreviation: Q1 = 25th percentile, Q3 = 75th percentile, p_{adj} = Holm-Bonferroni adjusted p.

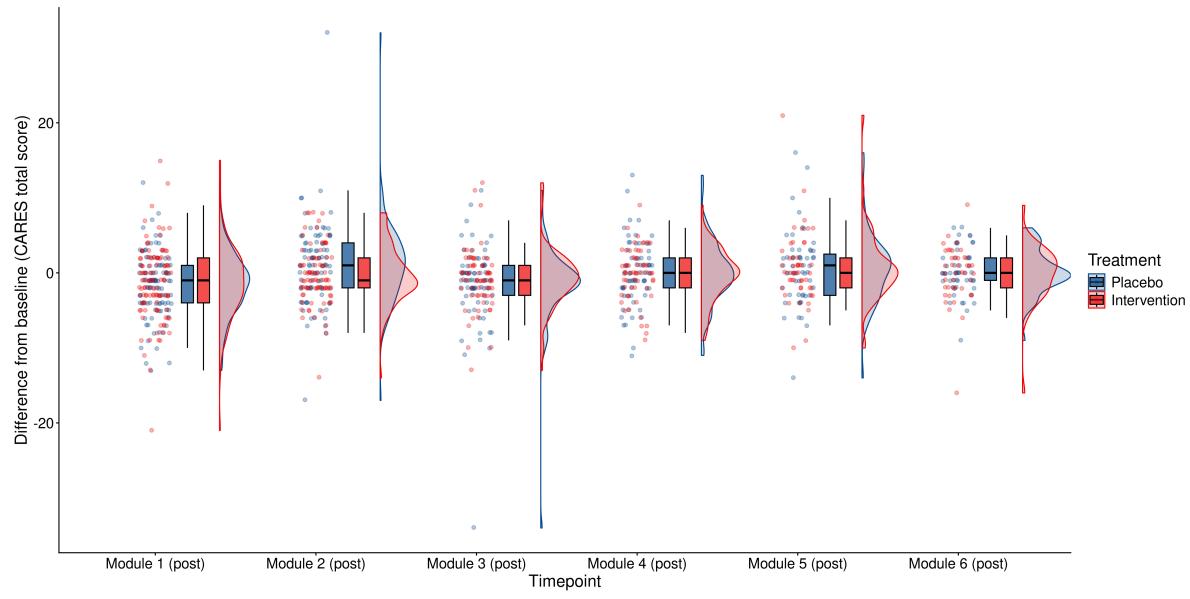


Figure 12: Rainplot of primary endpoint Child Sexual Abuse Risk Evaluation Self-Report (CARES difference from baseline scores).

6.2 Index of Desistance (IoD)

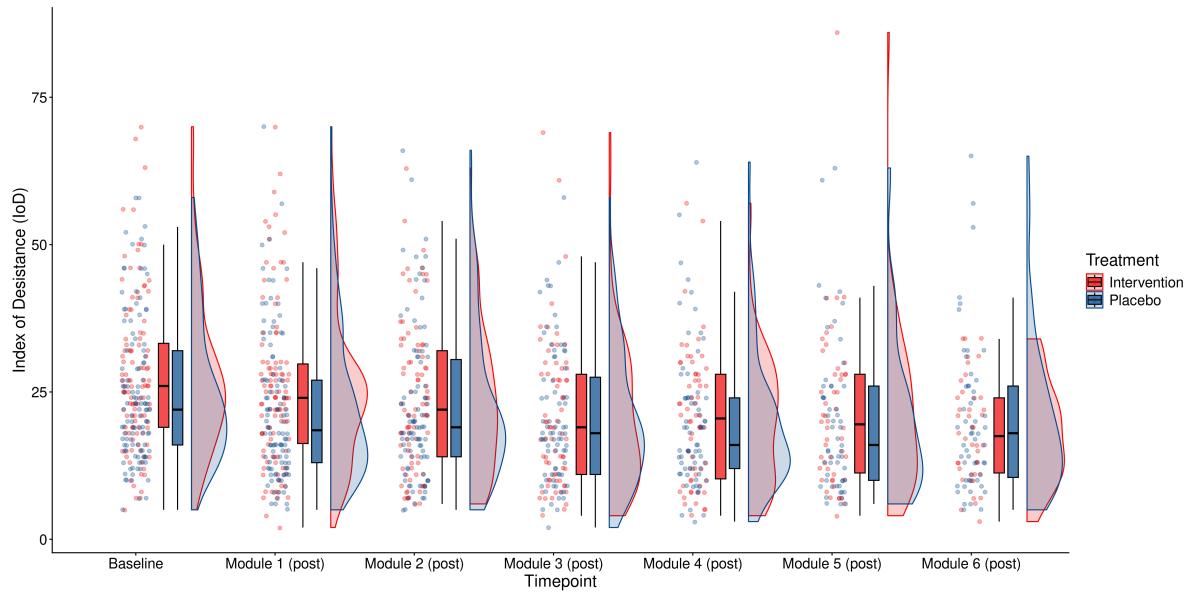


Figure 13: Rainplot of the Index of Desistance (IoD total score). Note, that the IoD is the preliminary version of the CARES and mentioned in the Statistical analysis protocol as primary outcome.

Table 12: Mixed model for repeated measures (MMRM) for the Index of Desistance (IoD total score).

Variable	Beta	95% CI	p-value
(Intercept)	3 · 1	1 · 2 to 5 · 0	0 · 0017
Type of supervision			
community supervision	0 · 00	Ref.	
post-release supervision	1 · 2	-0 · 29 to 2 · 8	0 · 11
Additional treatment			
No	0 · 00	Ref.	
Yes	-1 · 3	-2 · 5 to -0 · 10	0 · 034
Offense type			
Hands-off (only §184b StGB)	0 · 00	Ref.	
Hands-on (at least one conviction §176 ff StGB)	1 · 2	-0 · 08 to 2 · 4	0 · 065
IoD_baseline	0 · 88	0 · 83 to 0 · 92	<0 · 0001
Timepoint			
Baseline	0 · 00	Ref.	
Module 1 (post)	-2 · 2	-4 · 2 to -0 · 27	0 · 027
Module 2 (post)	-2 · 9	-4 · 9 to -0 · 95	0 · 0040
Module 3 (post)	-5 · 0	-7 · 0 to -2 · 9	<0 · 0001
Module 4 (post)	-5 · 4	-7 · 5 to -3 · 2	<0 · 0001
Module 5 (post)	-3 · 9	-6 · 2 to -1 · 6	0 · 0009
Module 6 (post)	-4 · 7	-7 · 1 to -2 · 3	0 · 0001
Treatment			
Intervention	0 · 00	Ref.	
Placebo	-0 · 30	-2 · 2 to 1 · 5	0 · 75
Treatment * Timepoint			
Placebo * Module 1 (post)	-0 · 50	-3 · 3 to 2 · 3	0 · 71
Placebo * Module 2 (post)	0 · 71	-2 · 1 to 3 · 5	0 · 62
Placebo * Module 3 (post)	0 · 41	-2 · 5 to 3 · 3	0 · 78
Placebo * Module 4 (post)	1 · 1	-1 · 9 to 4 · 1	0 · 47
Placebo * Module 5 (post)	0 · 84	-2 · 4 to 4 · 0	0 · 61
Placebo * Module 6 (post)	1 · 9	-1 · 4 to 5 · 1	0 · 26

Abbreviation: CI = Confidence Interval, StGB = German penalty law.

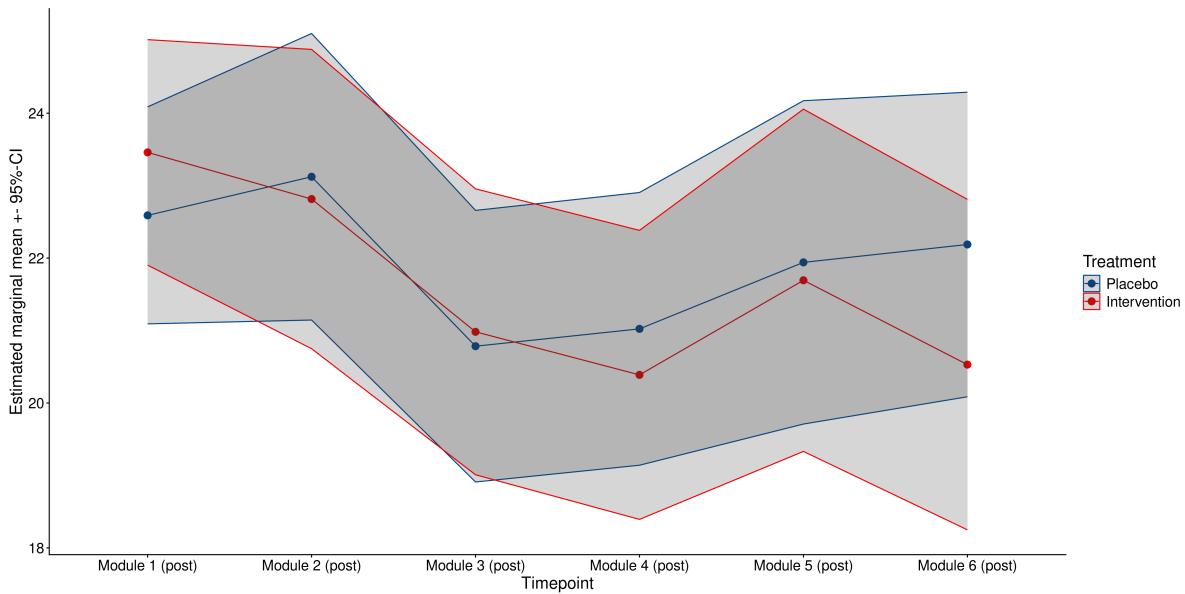


Figure 14: Estimated marginal means of the MMRM for the IoD total score.

Table 13: Explorative MMRM for Index of Desistance (IoD total score) with the additional factor time since beginning of the first module. StGB = German penalty law.

Variable	Beta	95% CI	p-value
(Intercept)	0 · 35	-2 · 5 to 3 · 2	0 · 81
Type of supervision			
community supervision	0 · 00	Ref.	
post-release supervision	1 · 2	-1 · 3 to 3 · 6	0 · 34
Additional treatment			
No	0 · 00	Ref.	
Yes	-1 · 7	-3 · 6 to 0 · 27	0 · 092
Offense type			
Hands-off (only §184b StGB)	0 · 00	Ref.	
Hands-on (at least one conviction §176 ff StGB)	1 · 8	-0 · 17 to 3 · 8	0 · 073
IoD_baseline	0 · 83	0 · 75 to 0 · 90	<0 · 0001
Static recidivism risk	0 · 57	-0 · 25 to 1 · 4	0 · 17
Timepoint * treatment			
Module 1 (post) * Intervention	0 · 91	-1 · 0 to 2 · 9	0 · 36
Module 2 (post) * Intervention	-0 · 26	-3 · 0 to 2 · 4	0 · 85
Module 3 (post) * Intervention	0 · 26	-2 · 3 to 2 · 8	0 · 84
Module 4 (post) * Intervention	-0 · 55	-3 · 1 to 2 · 0	0 · 67
Module 5 (post) * Intervention	-0 · 16	-3 · 3 to 3 · 0	0 · 92
Module 6 (post) * Intervention	-1 · 6	-4 · 6 to 1 · 4	0 · 30
Timepoint			
Module 1 (post)	0 · 00	Ref.	
Module 2 (post)	0 · 63	-0 · 99 to 2 · 3	0 · 44
Module 3 (post)	-1 · 6	-3 · 3 to 0 · 08	0 · 061
Module 4 (post)	-1 · 3	-3 · 3 to 0 · 63	0 · 18
Module 5 (post)	-0 · 33	-2 · 6 to 2 · 0	0 · 78
Module 6 (post)	-0 · 01	-2 · 3 to 2 · 3	>0 · 99
Time since baseline (years)	-0 · 56	-2 · 5 to 1 · 4	0 · 57

Abbreviation: CI = Confidence Interval, StGB = German penalty law.

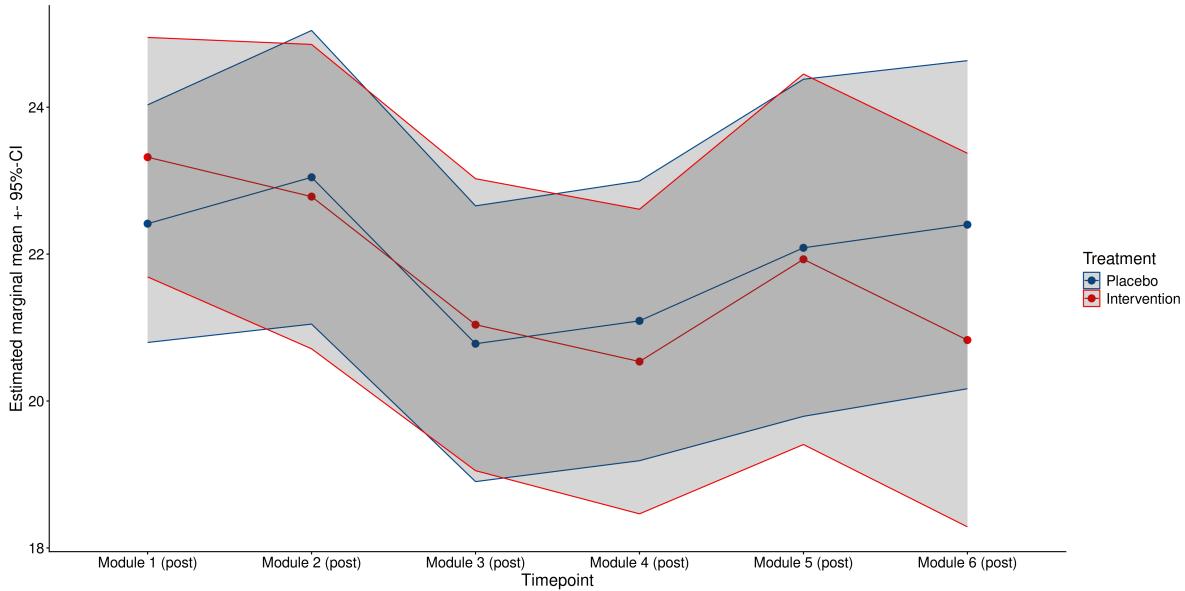


Figure 15: Estimated marginal means of the MMRM for the Index of Desistance (IoD total score) with time since baseline as additional factor.

Table 14: Sensitivity analysis for the Index of Desistance (IoD total score). Summarized are the results of the reference-based multiple imputation followed by ANCOVA models for the difference in the IoD (total score) compared to baseline.

Parameter	Estimate	95% CI	p
Group differences Module 1 (post)	-0.78	(-2.52, 0.91)	0.372
Least square mean [intervention] Module 1 (post)	-2.09	(-3.69, -1.49)	*< 0.001*
Least square mean [placebo] Module 1 (post)	-2.87	(-4.4, -1.83)	*< 0.001*
Group differences Module 2 (post)	0.08	(-2.53, 2.96)	0.964
Least square mean [intervention] Module 2 (post)	-2.56	(-4.53, -1.45)	*< 0.001*
Least square mean [placebo] Module 2 (post)	-2.47	(-5.28, -0.12)	*< 0.001*
Group differences Module 3 (post)	-0.38	(-2.5, 2.41)	0.982
Least square mean [intervention] Module 3 (post)	-4.40	(-6.67, -2.51)	*< 0.001*
Least square mean [placebo] Module 3 (post)	-4.78	(-6.45, -2.8)	*< 0.001*
Group differences Module 4 (post)	0.46	(-1.95, 3.43)	0.836
Least square mean [intervention] Module 4 (post)	-4.99	(-6.59, -2.53)	*< 0.001*
Least square mean [placebo] Module 4 (post)	-4.54	(-5.88, -2.89)	*< 0.001*
Group differences Module 5 (post)	0.14	(-2.34, 3.34)	0.878
Least square mean [intervention] Module 5 (post)	-3.75	(-6.48, -2.2)	*< 0.001*
Least square mean [placebo] Module 5 (post)	-3.61	(-6.38, -2.19)	*< 0.001*
Group differences Module 6 (post)	1.52	(-1.87, 3.49)	0.186
Least square mean [intervention] Module 6 (post)	-4.90	(-6.01, -2.8)	*< 0.001*
Least square mean [placebo] Module 6 (post)	-3.38	(-4.82, -1.99)	*< 0.001*

Table 15: Explorative pairwise comparison of the Index of Desistance (IoD total score) for the intervention arm.

Module	pre N = 106	post N = 106	p	q-value
Module 1 (post)			0 · 0050	0 · 025
Mean (SD)	28 (13)	25 (13)		
Median (Q1, Q3)	26 (19, 33)	24 (16, 30)		
Min, Max	5, 70	2, 70		
Module 2 (post)			0 · 70	>0 · 99
Mean (SD)	24 (14)	24 (13)		
Median (Q1, Q3)	23 (15, 28)	22 (14, 33)		
Min, Max	2, 70	6, 63		
Missing	22	22		
Module 3 (post)			0 · 0010	0 · 0060
Mean (SD)	23 (12)	21 (13)		
Median (Q1, Q3)	22 (13, 31)	19 (11, 29)		
Min, Max	6, 49	4, 69		
Missing	38	38		
Module 4 (post)			0 · 51	>0 · 99
Mean (SD)	21 (13)	21 (11)		
Median (Q1, Q3)	20 (12, 29)	21 (11, 28)		
Min, Max	4, 69	4, 57		
Missing	46	46		
Module 5 (post)			0 · 56	>0 · 99
Mean (SD)	20 (12)	21 (14)		
Median (Q1, Q3)	21 (10, 28)	20 (11, 28)		
Min, Max	5, 57	4, 86		
Missing	56	56		
Module 6 (post)			0 · 75	>0 · 99
Mean (SD)	19 (10)	18 (8)		
Median (Q1, Q3)	17 (10, 28)	18 (11, 24)		
Min, Max	4, 41	3, 34		
Missing	64	64		

Abbreviation: Q1 = 25th percentile, Q3 = 75th percentile, p_{adj} = Holm-Bonferroni adjusted p.

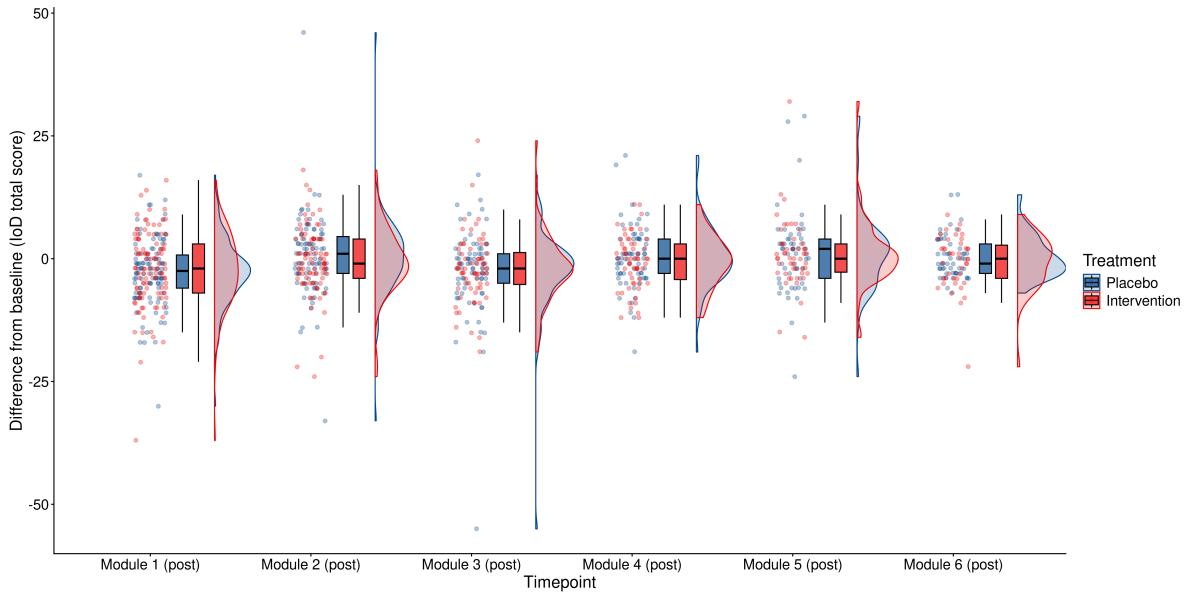


Figure 16: Rainplot of the differences from baseline for the Index of Desistance (IoD total score).

Table 16: Explorative pairwise comparison of the Index of Desistance (IoD total score) for the placebo arm.

Module	pre N = 106	post N = 106	p	q-value
Module 1 (post)			<0 · 0001	<0 · 0001
Mean (SD)	25 (12)	22 (12)		
Median (Q1, Q3)	22 (16, 32)	19 (13, 27)		
Min, Max	5, 58	5, 70		
Module 2 (post)			0 · 68	>0 · 99
Mean (SD)	22 (13)	22 (13)		
Median (Q1, Q3)	19 (12, 27)	19 (14, 31)		
Min, Max	5, 70	5, 66		
Missing	15	15		
Module 3 (post)			0 · 040	0 · 20
Mean (SD)	23 (14)	20 (12)		
Median (Q1, Q3)	19 (14, 31)	18 (11, 27)		
Min, Max	5, 66	2, 58		
Missing	31	31		
Module 4 (post)			>0 · 99	>0 · 99
Mean (SD)	20 (11)	20 (12)		
Median (Q1, Q3)	18 (11, 26)	17 (12, 24)		
Min, Max	2, 47	3, 64		
Missing	38	38		
Module 5 (post)			0 · 67	>0 · 99
Mean (SD)	19 (12)	20 (13)		
Median (Q1, Q3)	15 (11, 24)	16 (10, 26)		
Min, Max	3, 64	6, 63		
Missing	51	51		
Module 6 (post)			0 · 80	>0 · 99
Mean (SD)	20 (13)	20 (13)		
Median (Q1, Q3)	16 (10, 26)	17 (10, 26)		
Min, Max	6, 63	5, 65		
Missing	53	53		

Abbreviation: Q1 = 25th percentile, Q3 = 75th percentile, p_{adj} = Holm-Bonferroni adjusted p.

Table 17: Difference of the Index of Desistance (IoD total score) between before (pre) and after (post) each module.

Characteristic	Placebo N = 106	Intervention N = 106	p	q-value
Module 1 (post)				
Mean (SD)	-3 (7)	-2 (8)	0 · 45	>0 · 99
Median (Q1, Q3)	-3 (-6, 1)	-2 (-7, 3)		
Min, Max	-30, 17	-37, 16		
Module 2 (post)			0 · 33	>0 · 99
Mean (SD)	1 (8)	0 (7)		
Median (Q1, Q3)	1 (-3, 5)	-1 (-4, 4)		
Min, Max	-33, 46	-24, 18		
Missing	15	22		
Module 3 (post)			0 · 85	>0 · 99
Mean (SD)	-3 (9)	-2 (7)		
Median (Q1, Q3)	-2 (-5, 1)	-2 (-6, 2)		
Min, Max	-55, 17	-19, 24		
Missing	31	38		
Module 4 (post)			0 · 44	>0 · 99
Mean (SD)	0 · 3 (6 · 3)	-0 · 7 (5 · 4)		
Median (Q1, Q3)	0 · 0 (-3 · 0, 4 · 0)	0 · 0 (-4 · 5, 3 · 0)		
Min, Max	-19 · 0, 21 · 0	-12 · 0, 11 · 0		
Missing	38	46		
Module 5 (post)			0 · 88	>0 · 99
Mean (SD)	1 · 0 (8 · 4)	0 · 9 (7 · 1)		
Median (Q1, Q3)	2 · 0 (-4 · 0, 4 · 0)	0 · 0 (-3 · 0, 3 · 0)		
Min, Max	-24 · 0, 29 · 0	-16 · 0, 32 · 0		
Missing	51	56		
Module 6 (post)			0 · 76	>0 · 99
Mean (SD)	0 · 3 (4 · 3)	-0 · 6 (5 · 5)		
Median (Q1, Q3)	-1 · 0 (-3 · 0, 3 · 0)	0 · 0 (-4 · 0, 3 · 0)		
Min, Max	-7 · 0, 13 · 0	-22 · 0, 9 · 0		
Missing	53	64		

Abbreviation: Q1 = 25th percentile, Q3 = 75th percentile, p_{adj} = Holm-Bonferroni adjusted p.

7 Secondary Outcomes

7.1 Differences from baseline

Table 18: Difference between before (pre) and after (post) for each secondary endpoint. Differences (from pre to post) of secondary endpoints have been compared between the treatment groups using the Brunner-Munzel test for unpaired data.

Characteristic	Placebo N = 113	Intervention N = 108	p	q-value
CVTRQ (total score)			0 · 19	>0 · 99
Mean (SD)	0 · 5 (4 · 9)	1 · 5 (5 · 7)		
Median (Q1, Q3)	0 · 0 (-3 · 0, 3 · 0)	1 · 0 (-2 · 0, 5 · 0)		
Min, Max	-11 · 0, 15 · 0	-12 · 0, 28 · 0		
RCQ (total score)			0 · 21	>0 · 99
Mean (SD)	0 · 3 (4 · 5)	-0 · 5 (4 · 6)		
Median (Q1, Q3)	0 · 0 (-3 · 0, 3 · 0)	-1 · 0 (-4 · 0, 3 · 0)		
Min, Max	-14 · 0, 14 · 0	-12 · 0, 13 · 0		
F-Soz-U (total score)			0 · 22	>0 · 99
Mean (SD)	0 · 14 (0 · 46)	0 · 01 (0 · 56)		
Median (Q1, Q3)	0 · 14 (-0 · 14, 0 · 43)	0 · 00 (-0 · 29, 0 · 43)		
Min, Max	-0 · 86, 1 · 14	-2 · 29, 1 · 29		
Missing	19	17		
OQMPR (total score)			0 · 94	>0 · 99
Mean (SD)	-0 · 9 (5 · 4)	-1 · 3 (5 · 9)		
Median (Q1, Q3)	-1 · 0 (-5 · 0, 2 · 0)	-1 · 0 (-4 · 0, 3 · 0)		
Min, Max	-15 · 0, 18 · 0	-21 · 0, 20 · 0		
Missing	19	17		
UCLA (total score)			0 · 28	>0 · 99
Mean (SD)	-0 · 11 (3 · 23)	-0 · 54 (2 · 98)		
Median (Q1, Q3)	0 · 00 (-2 · 00, 2 · 00)	0 · 00 (-2 · 00, 1 · 00)		
Min, Max	-13 · 00, 10 · 00	-12 · 00, 11 · 00		
Missing	19	17		
BIS-15 (total score)			0 · 92	>0 · 99
Mean (SD)	0 · 0 (3 · 9)	0 · 0 (3 · 4)		
Median (Q1, Q3)	-0 · 5 (-2 · 0, 2 · 0)	0 · 0 (-2 · 0, 2 · 0)		
Min, Max	-19 · 0, 11 · 0	-10 · 0, 10 · 0		
Missing	35	36		
CUSI (total score)			0 · 55	>0 · 99
Mean (SD)	0 · 8 (5 · 3)	0 · 4 (6 · 5)		
Median (Q1, Q3)	0 · 0 (-2 · 0, 2 · 0)	0 · 0 (-3 · 0, 2 · 5)		
Min, Max	-8 · 0, 23 · 0	-11 · 0, 27 · 0		
Missing	35	36		
DERS (subscale impulsivity)			0 · 024	0 · 36
Mean (SD)	-0 · 15 (1 · 24)	0 · 49 (2 · 10)		
Median (Q1, Q3)	0 · 00 (0 · 00, 0 · 00)	0 · 00 (0 · 00, 1 · 00)		
Min, Max	-4 · 00, 5 · 00	-5 · 00, 7 · 00		
Missing	35	36		
NARQ (subscale externalizing strategies)			0 · 30	>0 · 99
Mean (SD)	-0 · 29 (2 · 53)	0 · 17 (2 · 22)		
Median (Q1, Q3)	0 · 00 (-1 · 00, 1 · 00)	0 · 00 (0 · 00, 1 · 00)		
Min, Max	-16 · 00, 4 · 00	-12 · 00, 5 · 00		
Missing	35	36		
SPSI-R (total score)			0 · 97	>0 · 99
Mean (SD)	0 · 02 (1 · 46)	0 · 08 (1 · 66)		
Median (Q1, Q3)	0 · 00 (-0 · 90, 0 · 80)	0 · 00 (-1 · 00, 1 · 20)		
Min, Max	-4 · 20, 4 · 60	-4 · 00, 4 · 40		
Missing	41	46		
BMS (total score)			0 · 14	>0 · 99
Mean (SD)	-1 (7)	2 (10)		
Median (Q1, Q3)	0 (-4, 1)	0 (-3, 5)		
Min, Max	-20, 23	-12, 39		
Missing	54	55		
EKK-R (total score)			0 · 49	>0 · 99
Mean (SD)	0 · 3 (3 · 8)	-0 · 6 (4 · 8)		
Median (Q1, Q3)	0 · 0 (-2 · 0, 3 · 0)	0 · 0 (-2 · 0, 2 · 0)		
Min, Max	-9 · 0, 9 · 0	-17 · 0, 9 · 0		
Missing	60	66		
ESIQ (total score)			0 · 0020	0 · 032
Mean (SD)	-0 · 1 (7 · 9)	-3 · 6 (5 · 4)		
Median (Q1, Q3)	0 · 0 (-1 · 0, 2 · 0)	-3 · 0 (-8 · 0, 0 · 0)		
Min, Max	-22 · 0, 31 · 0	-14 · 0, 11 · 0		
Missing	560	66		
HBI-19 (total score)			0 · 92	>0 · 99
Mean (SD)	0 · 0 (5 · 5)	-1 · 0 (6 · 9)		
Median (Q1, Q3)	0 · 0 (-2 · 0, 2 · 0)	0 · 0 (-4 · 0, 2 · 0)		
Min, Max	-11 · 0, 18 · 0	-23 · 0, 18 · 0		
Missing	60	66		
SSIC (total score)			0 · 077	>0 · 99
Mean (SD)	0 · 1 (4 · 5)	0 · 6 (3 · 6)		
Median (Q1, Q3)	0 · 0 (-1 · 0, 0 · 0)	0 · 0 (0 · 0, 2 · 0)		
Min, Max	-18 · 0, 20 · 0	-12 · 0, 9 · 0		
Missing	60	66		
SQL (Kappa score)			0 · 961	0 · 95

Table 19: Linear regression model for the Corrections Victoria Treatment Readiness Questionnaire CVTRQ (total score); Number of observations: 221.

Variable	N	95% CI	p-value	adjusted p value
(Intercept)	221	18	9.1, 26	<0.001
Baseline value (pre)	221	0.81	0.71, 0.91	<0.001
Treatment				
Placebo	113	0.00	Ref.	
Intervention	108	0.89	-0.49, 2.3	0.2
Offense type				
Hands-on (at least one conviction §176 ff StGB)	89	0.00	Ref.	
Hands-off (only §184b StGB)	132	-0.22	-1.8, 1.3	0.8
Type of supervision				
community supervision	176	0.00	Ref.	
post-release supervision	45	-0.51	-2.4, 1.4	0.6
Additional treatment				
No	148	0.00	Ref.	
Yes	73	0.24	-1.3, 1.7	0.7
Static recidivism risk (baseline)	221	-0.47	-1.1, 0.17	0.2

Abbreviation: CI = Confidence Interval

7.2 CVTRQ

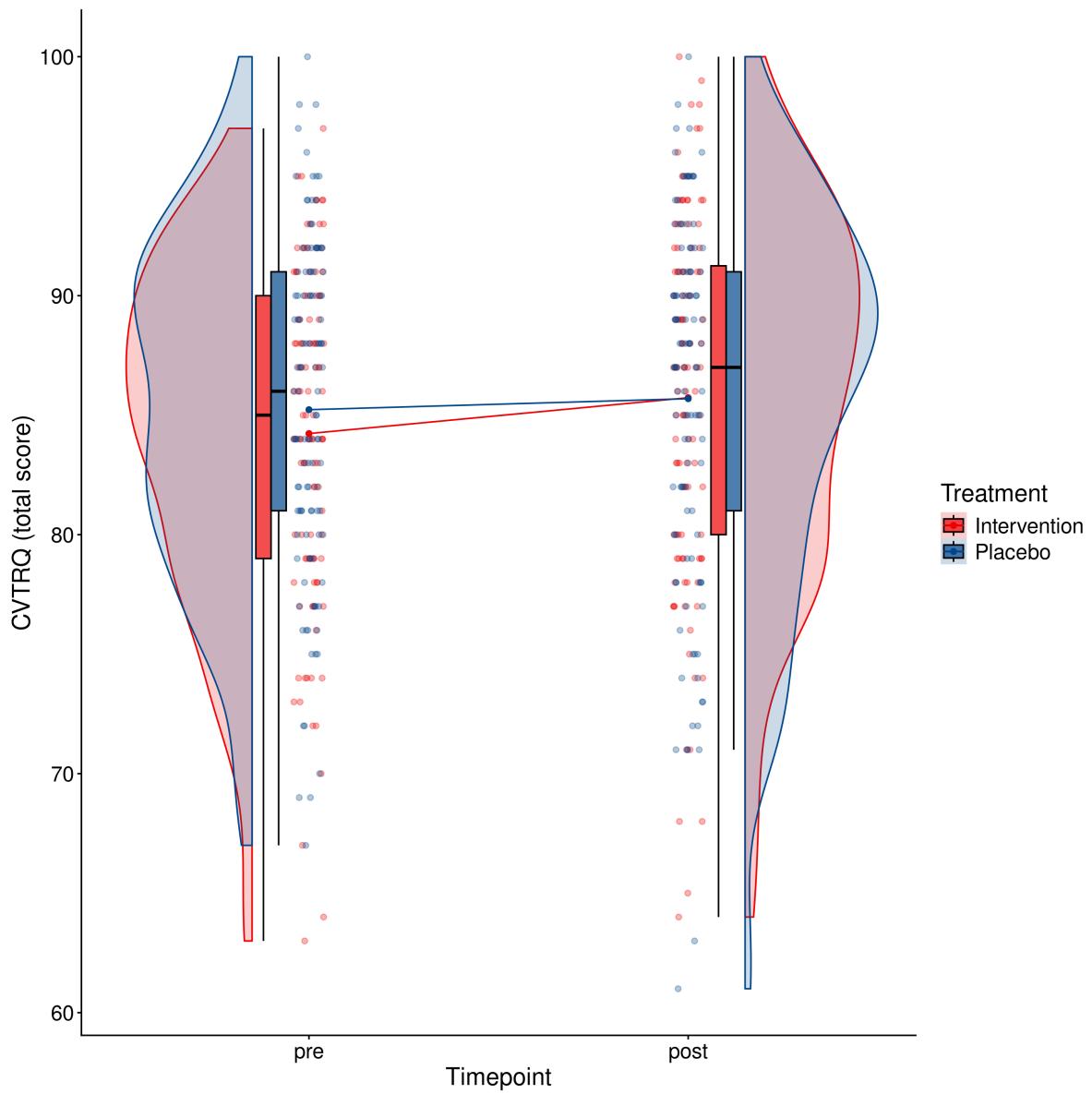


Figure 17: Rainplot of the Corrections Victoria Treatment Readiness Questionnaire (CVTRQ, total score).

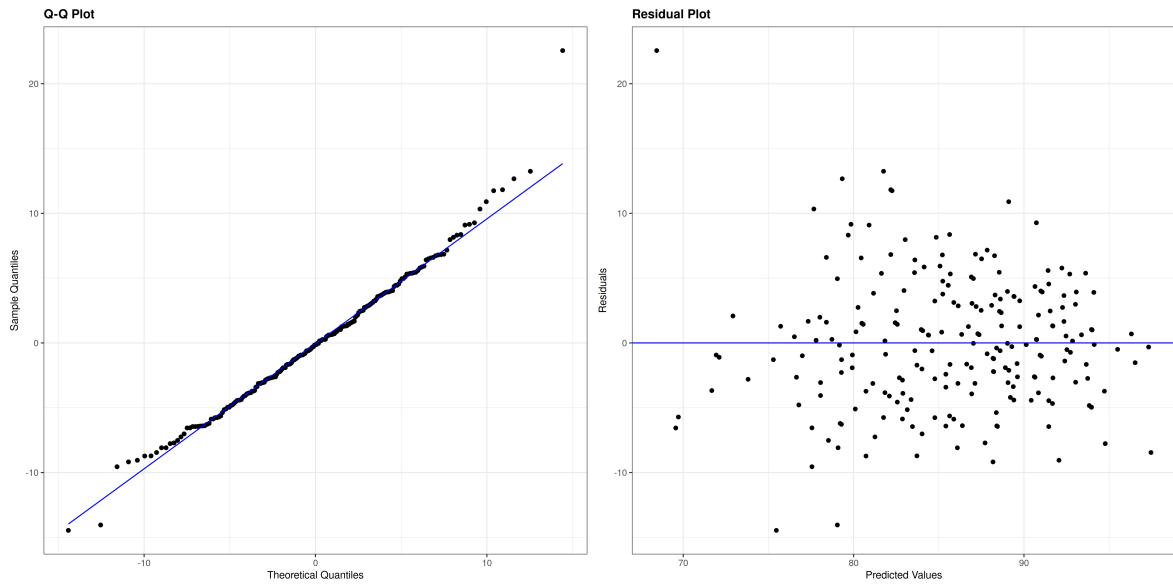


Figure 18: Diagnostic plots for the Corrections Victoria Treatment Readiness Questionnaire (CVTRQ, total score).

7.3 RCQ

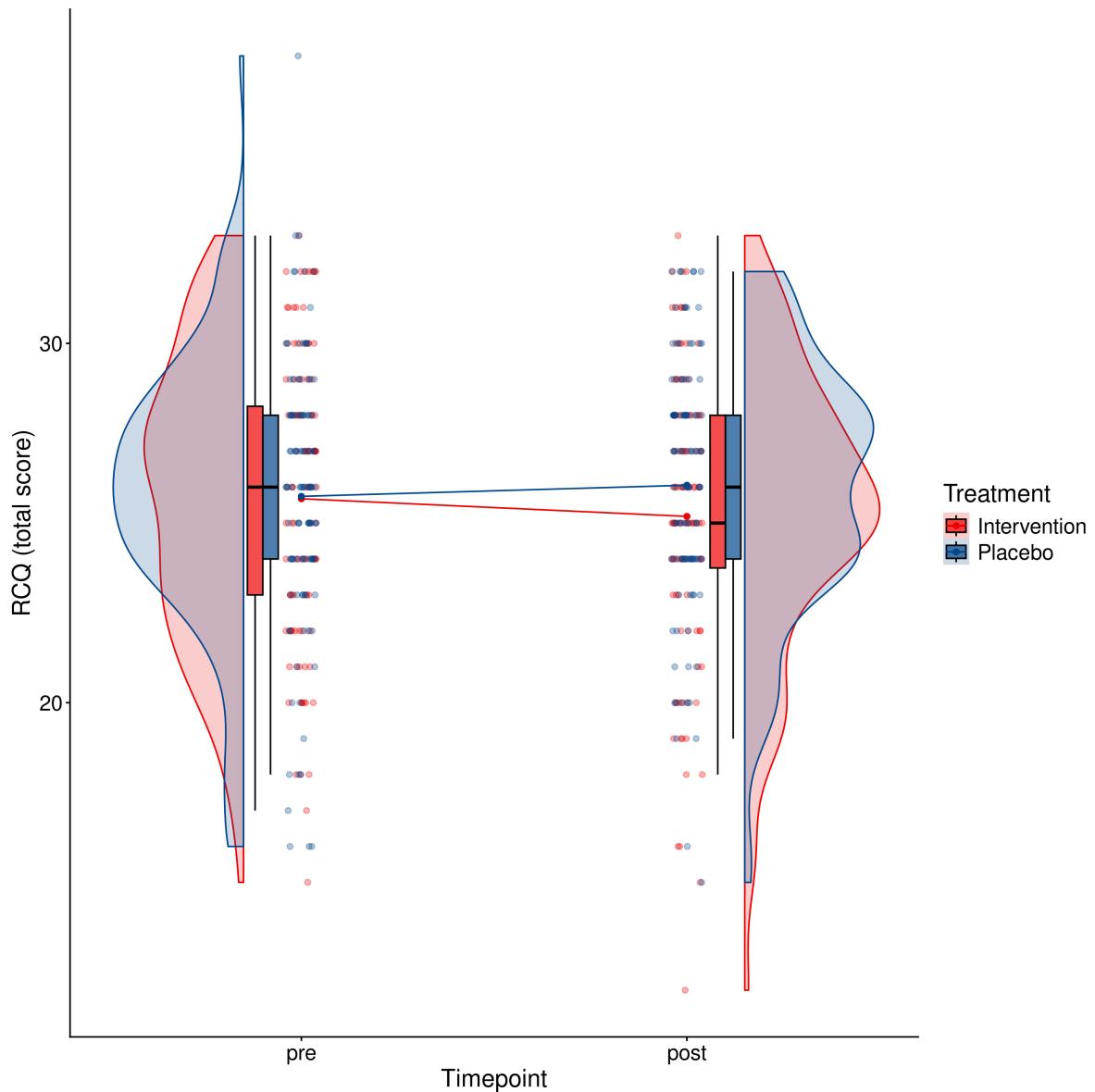


Figure 19: Rainplot of the Readiness to Change Questionnaire (RCQ, total score).

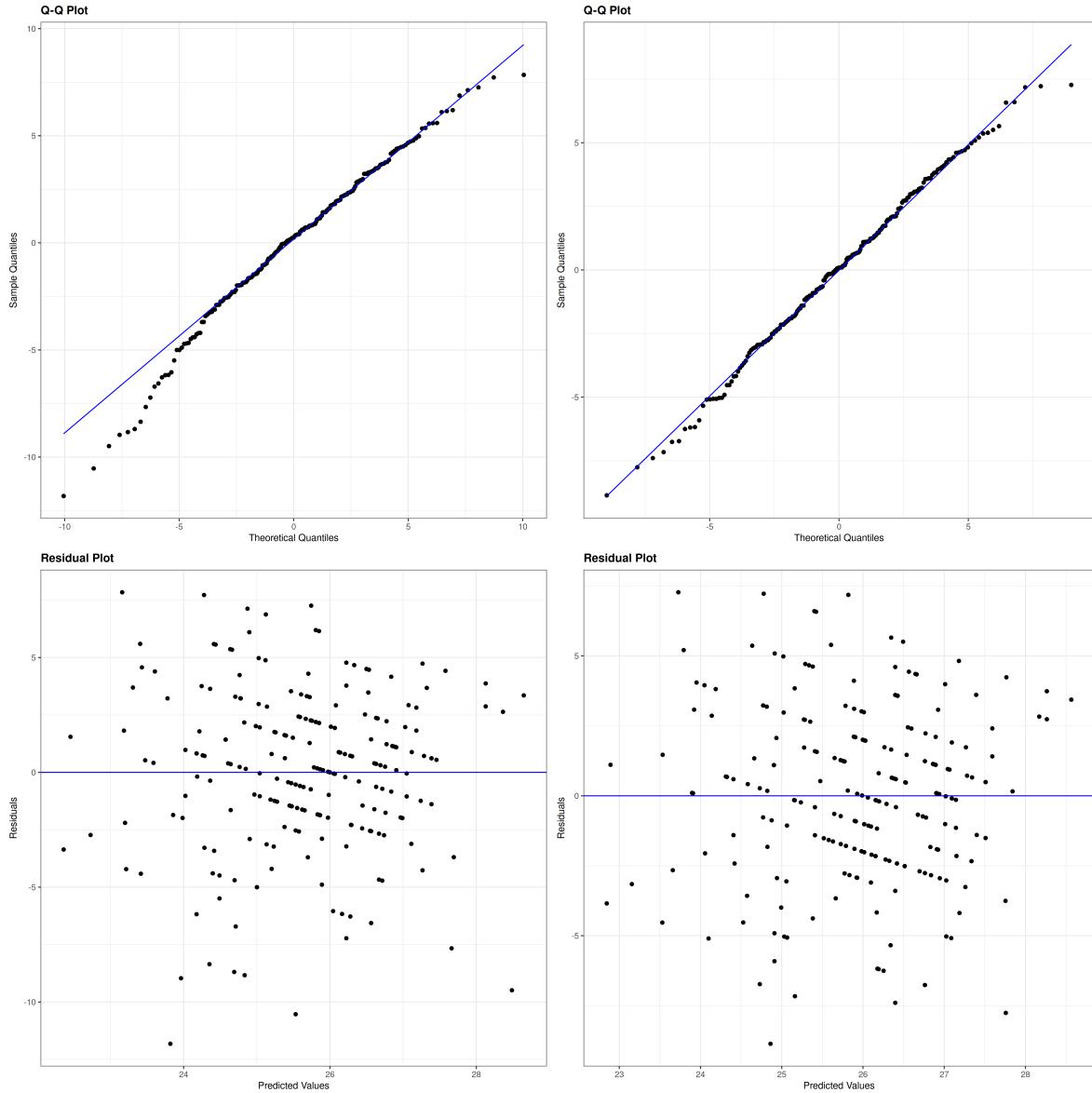


Figure 20: Diagnostic plots for the Readiness to Change Questionnaire (RCQ, total score). The plots on the left correspond to the model fitted with all data points, the plots on the right correspond to the model after the removal of identified outliers. Both models showed similar results.

Table 20: Linear regression model for the Readiness to Change Questionnaire RCQ (total score); Number of observations: 221).

Variable	N	95% CI	p-value	adjusted p value
(Intercept)	221	19	16, 23	<0.001
Baseline value (pre)	221	0.27	0.15, 0.40	<0.001
Treatment				
Placebo	113	0.00	Ref.	
Intervention	108	-0.89	-1.8, 0.07	0.069
				0.826
Offense type				
Hands-on (at least one conviction §176 ff StGB)	89	0.00	Ref.	
Hands-off (only §184b StGB)	132	0.12	-0.95, 1.2	0.8
				1.000
Type of supervision				
community supervision	176	0.00	Ref.	
post-release supervision	45	-0.84	-2.2, 0.48	0.2
				1.000
Additional treatment				
No	148	0.00	Ref.	
Yes	73	-0.40	-1.4, 0.64	0.4
				1.000
Static recidivism risk (baseline)	221	-0.08	-0.53, 0.36	0.7
				1.000

Abbreviation: CI = Confidence Interval

7.4 F-Soz-U

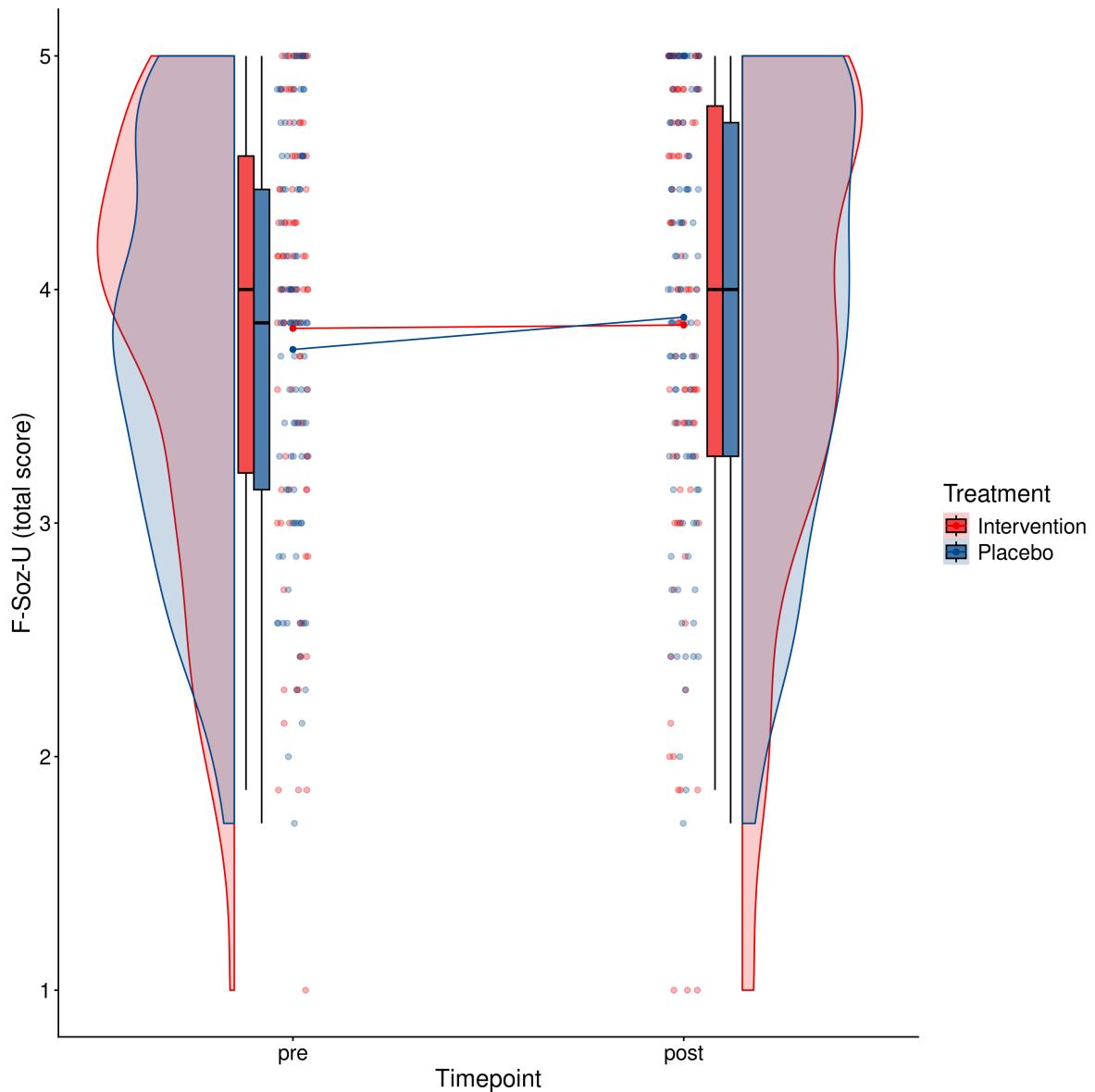


Figure 21: Rainplot of the Social Support Questionnaire (F-Soz-U, total score).

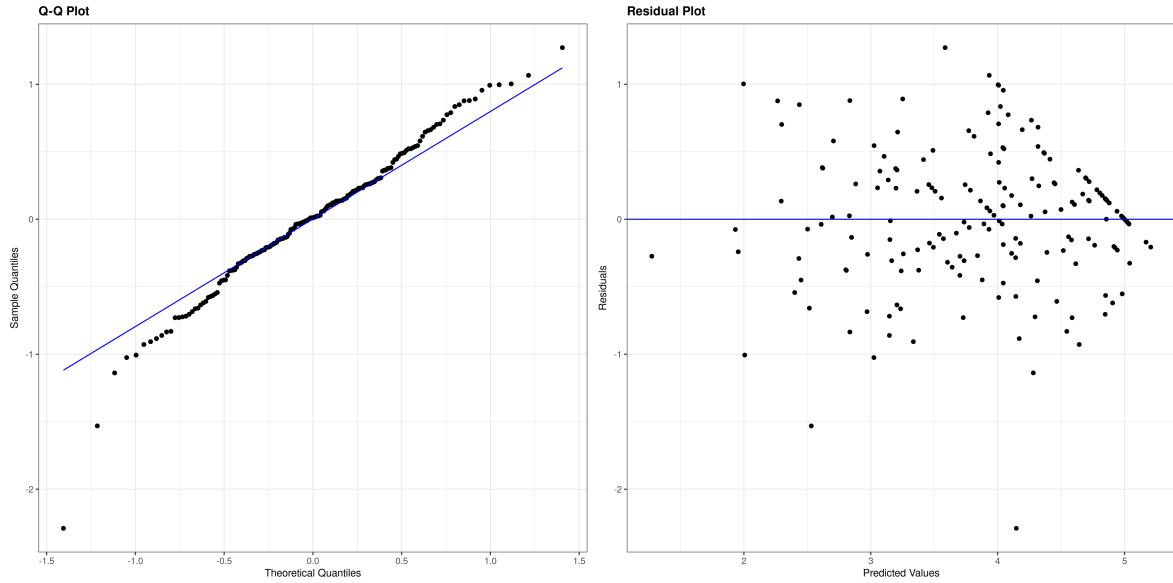


Figure 22: Diagnostic plots for the Social Support Questionnaire (F-Soz-U, total score).

Table 21: Linear regression model for the Social Support Questionnaire F-Soz-U (total score); Number of observations: 185.

Variable	N	95% CI	p-value	adjusted p value
(Intercept)	185	0.50	0.12, 0.89	0.010
Baseline value (pre)	185	0.94	0.85, 1.0	<0.001
Treatment				
Placebo	94	0.00	Ref.	
Intervention	91	-0.13	-0.28, 0.02	0.092
Offense type				
Hands-on (at least one conviction §176 ff StGB)	72	0.00	Ref.	
Hands-off (only §184b StGB)	113	-0.05	-0.22, 0.12	0.5
Type of supervision				
community supervision	151	0.00	Ref.	
post-release supervision	34	-0.04	-0.25, 0.17	0.7
Additional treatment				
No	122	0.00	Ref.	
Yes	63	-0.10	-0.26, 0.06	0.2
Static recidivism risk (baseline)	185	-0.04	-0.11, 0.03	0.3

Abbreviation: CI = Confidence Interval

Table 22: Linear regression model for the Questionnaire for the Measurement of Psychological Reactance OQMPr (total score) with n = 185 observations.

Variable	N	95% CI	p-value	adjusted p value
(Intercept)	185	1.8	-2.0, 5.6	0.3
Baseline value (pre)	185	0.89	0.79, 1.0	<0.001
Treatment				
Placebo	94	0.00	Ref.	
Intervention	91	-0.36	-2.0, 1.3	0.7
Offense type				
Hands-on (at least one conviction §176 ff StGB)	72	0.00	Ref.	
Hands-off (only §184b StGB)	113	-1.1	-2.9, 0.68	0.2
Type of supervision				
community supervision	151	0.00	Ref.	
post-release supervision	34	-0.44	-2.7, 1.8	0.7
Additional treatment				
No	122	0.00	Ref.	
Yes	63	-0.13	-1.9, 1.6	0.9
Static recidivism risk (baseline)	185	0.68	-0.09, 1.4	0.083

Abbreviation: CI = Confidence Interval

7.5 OQMPr

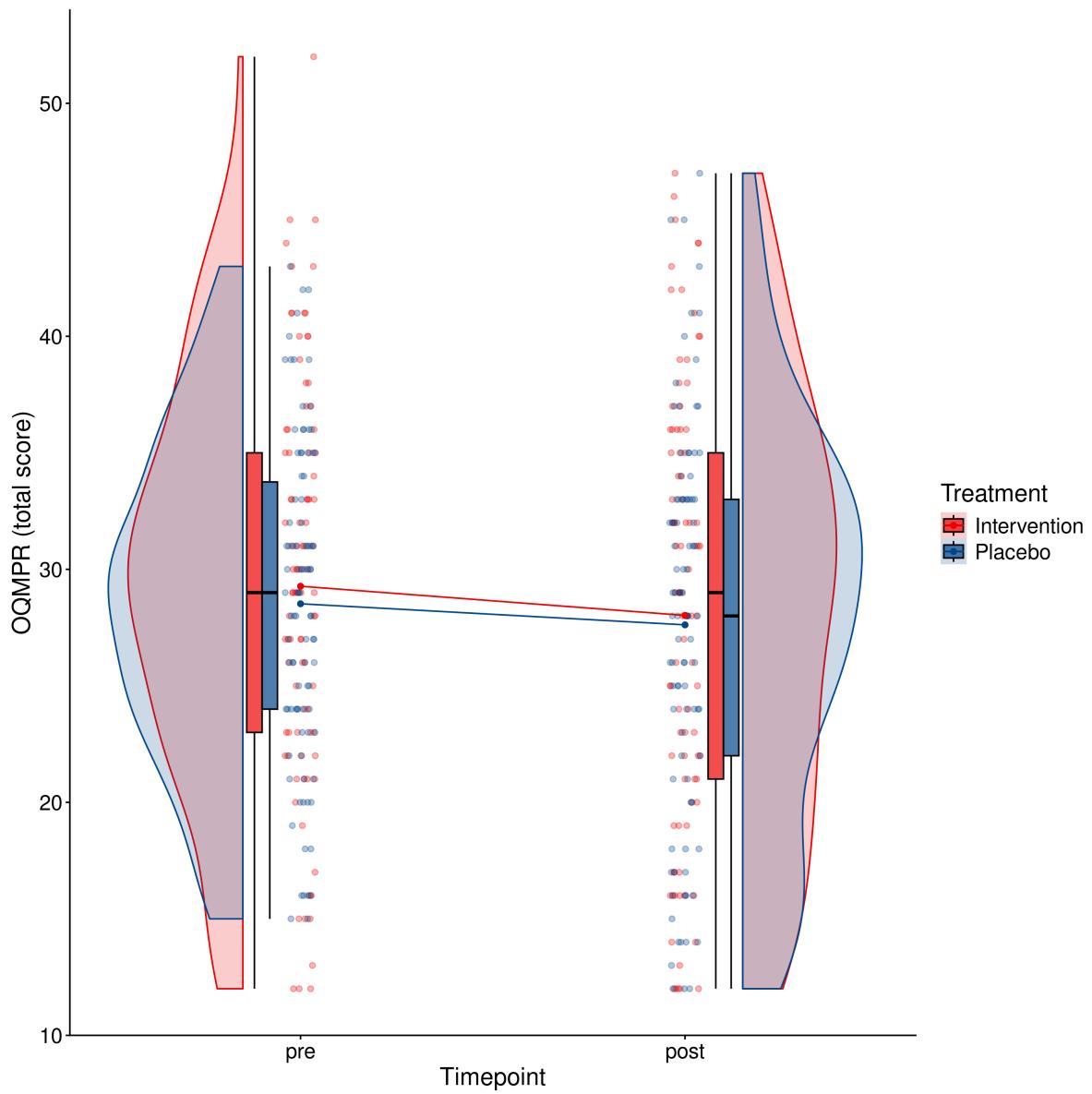


Figure 23: Rainplot of the Questionnaire for the Measurement of Psychological Reactance (OQMPR, total score).

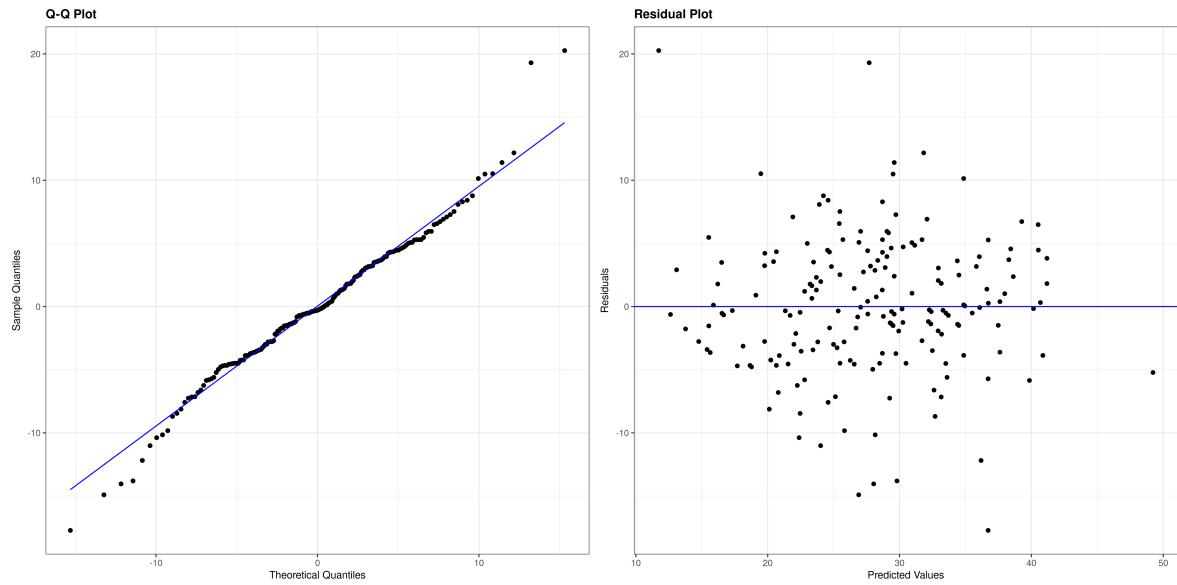


Figure 24: Diagnostic plots for the Questionnaire for the Measurement of Psychological Reactance (OQMMPR, total score).

7.6 SOI-R

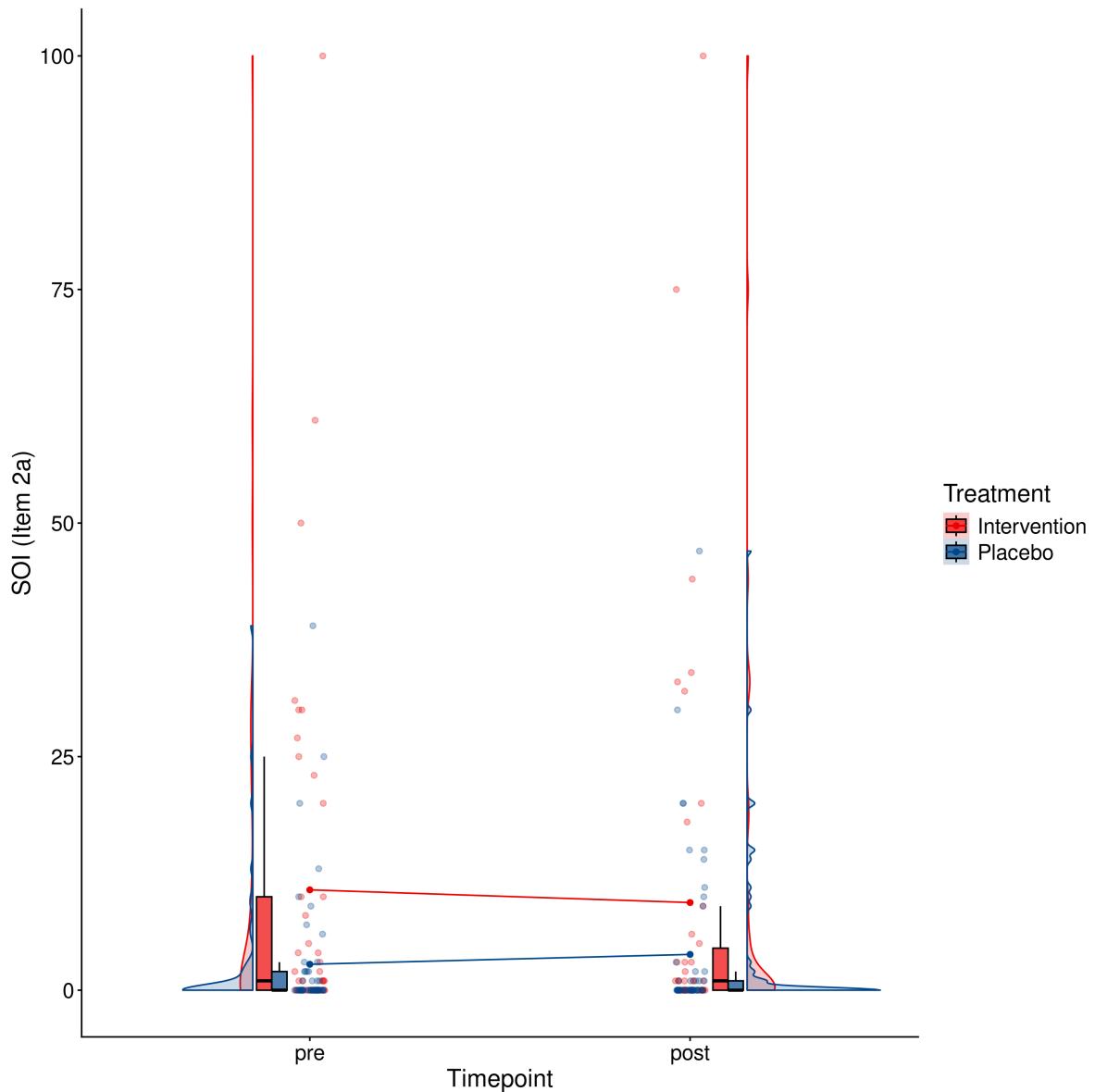


Figure 25: Rainplot of the Sexual Outlet Inventory revised (SOI-R, Item 2a).

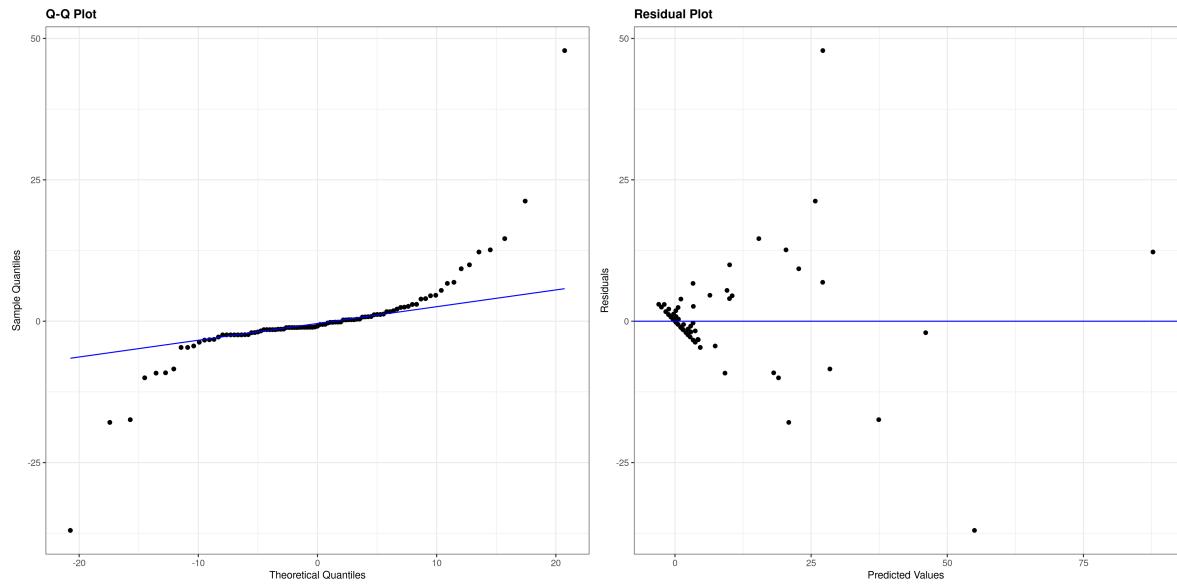


Figure 26: Diagnostic plots for the Sexual Outlet Inventory revised (SOI-R, Item 2a).

7.7 UCLA

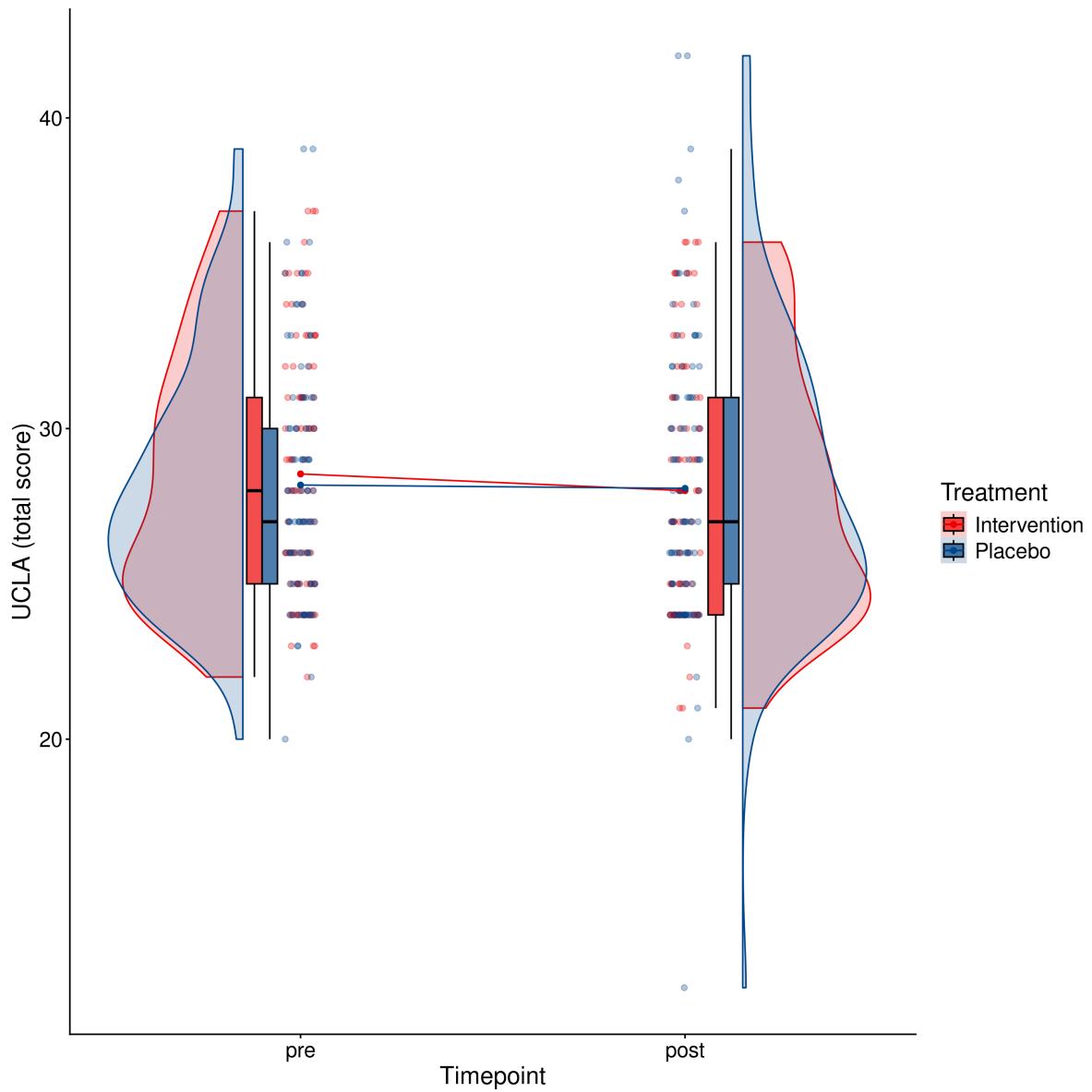


Figure 27: Rainplot of the UCLA Loneliness Scale (UCLA, total score).

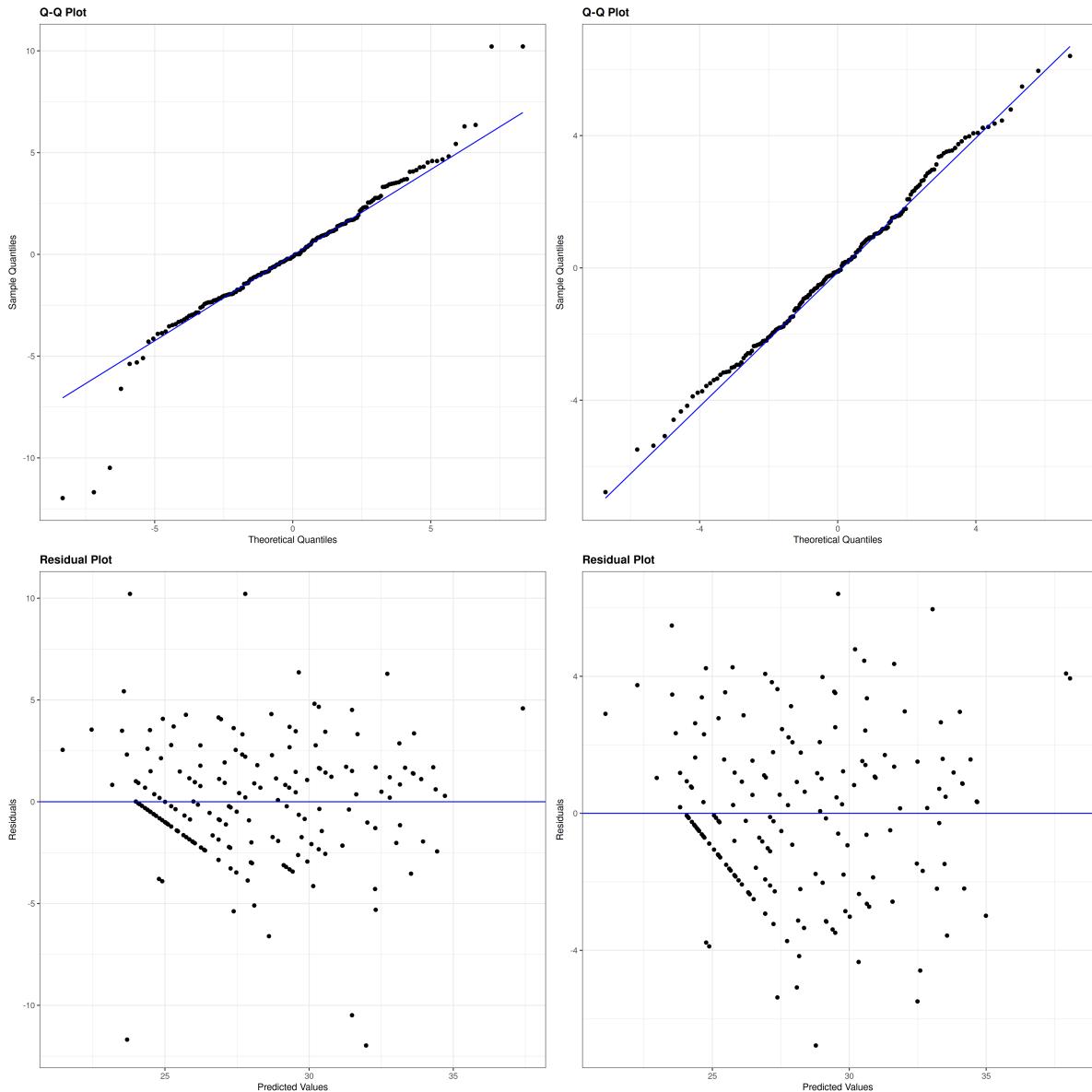


Figure 28: Diagnostic plots for the UCLA Loneliness Scale (UCLA, total score). The plots on the left correspond to the model fitted with all data points, the plots on the right correspond to the model after the removal of identified outliers. Both models showed similar results.

Table 23: Linear regression model for the UCLA Loneliness Scale UCLA (total score) with n = 185 observations.

Variable	N	95% CI	p-value	adjusted p value
(Intercept)	185	5.2 1.8, 8.6	0.003	
Baseline value (pre)	185	0.82 0.70, 0.94	<0.001	
Treatment				
Placebo	94	0.00 Ref.		
Intervention	91	-0.40 -1.3, 0.49	0.4	1.00
Offense type				
Hands-on (at least one conviction §176 ff StGB)	72	0.00 Ref.		
Hands-off (only §184b StGB)	113	-0.62 -1.6, 0.37	0.2	1.00
Type of supervision				
community supervision	151	0.00 Ref.		
post-release supervision	34	-0.47 -1.7, 0.77	0.5	1.00
Additional treatment				
No	122	0.00 Ref.		
Yes	63	0.08 -0.88, 1.0	0.9	1.00
Static recidivism risk (baseline)	185	0.11 -0.31, 0.52	0.6	1.00

Abbreviation: CI = Confidence Interval

7.8 BIS-15

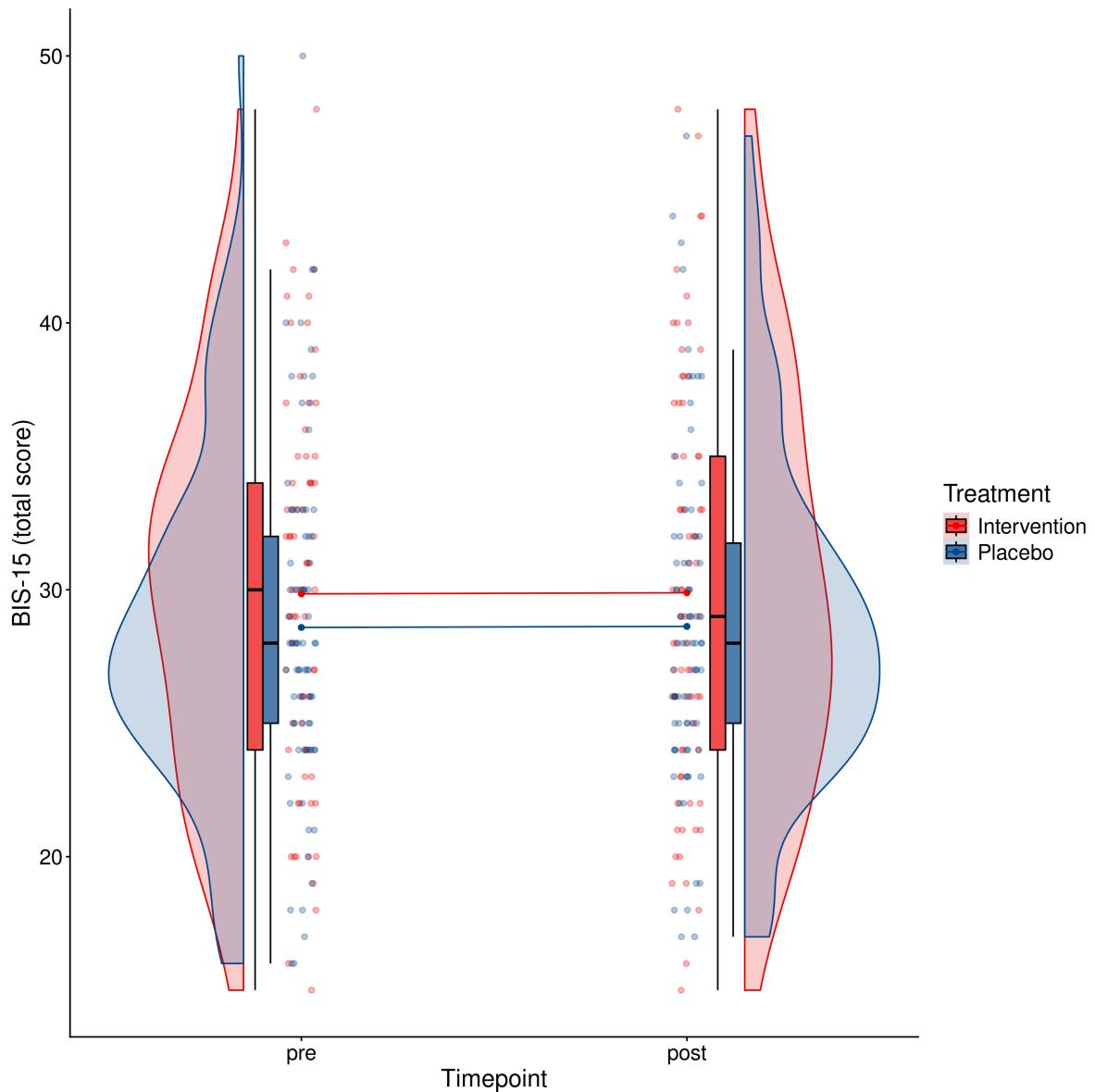


Figure 29: Rainplot of the Barratt Impulsiveness Scale-15 (BIS-15, total score).

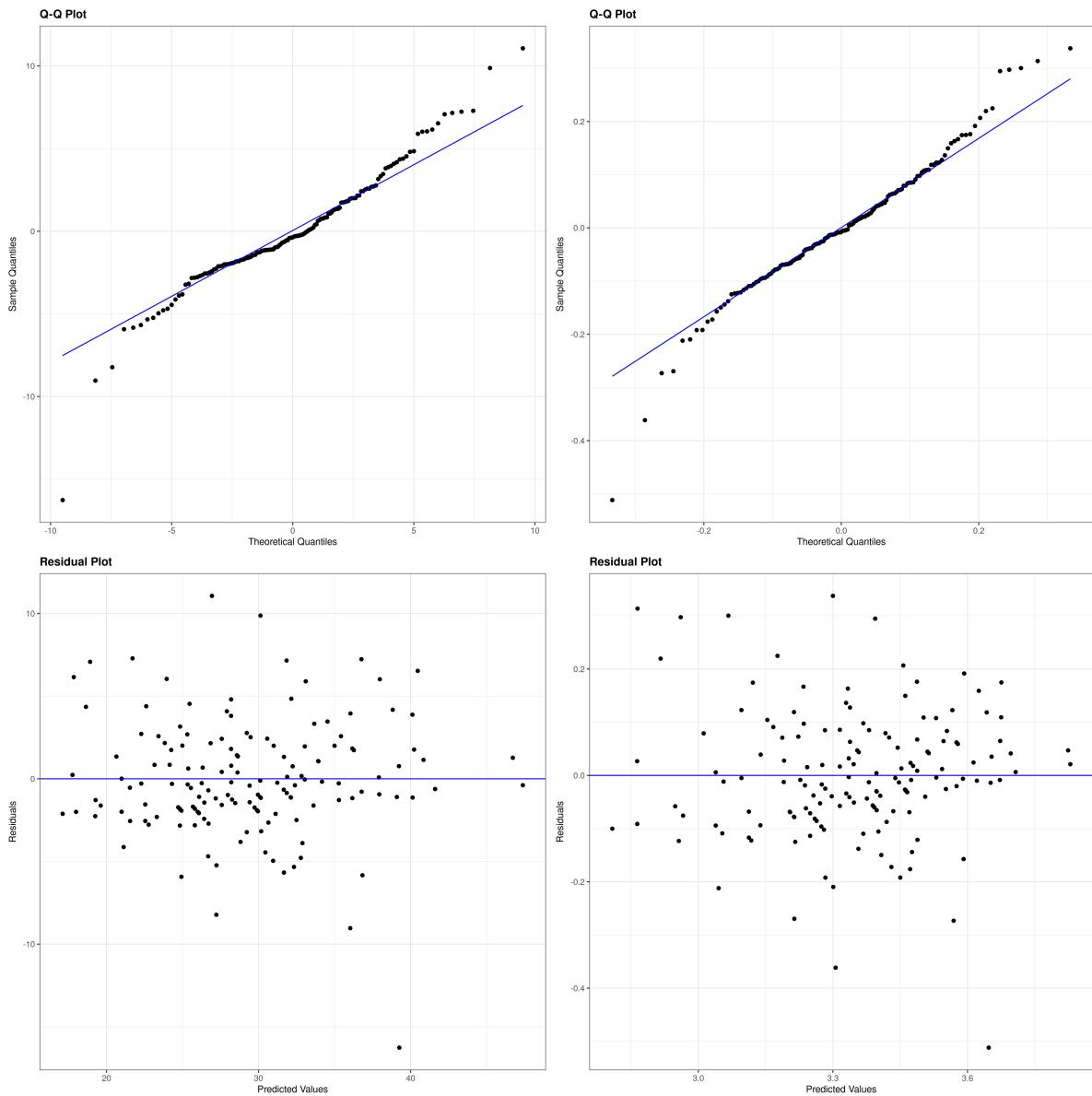


Figure 30: Diagnostic plots for the Barratt Impulsiveness Scale-15 (BIS-15, total score). The plots on the left correspond to an ordinary linear model, the plots on the right correspond to the model where the bis score was log-transformed.

Table 24: Linear regression model for the Barratt Impulsiveness Scale-15 BIS-15 (total score) with n = 150 observations.

Variable	N	95% CI	p-value	adjusted p value
(Intercept)	150	3.2	0.41, 6.0	0.025
Baseline value (pre)	150	0.88	0.80, 0.97	<0.001
Treatment				
Placebo	78	0.00	Ref.	
Intervention	72	0.15	-1.0, 1.3	0.8
Offense type				
Hands-on (at least one conviction §176 ff StGB)	58	0.00	Ref.	
Hands-off (only §184b StGB)	92	0.77	-0.56, 2.1	0.3
Type of supervision				
community supervision	120	0.00	Ref.	
post-release supervision	30	0.67	-0.93, 2.3	0.4
Additional treatment				
No	95	0.00	Ref.	
Yes	55	0.01	-1.2, 1.2	>0.9
Static recidivism risk (baseline)	150	-0.27	-0.83, 0.28	0.3

Abbreviation: CI = Confidence Interval

Table 25: Linear regression model for CUSI (total score) with n = 150 observations.

Variable	N	95% CI	p-value	adjusted p value
(Intercept)	150	3.5	-0.24, 7.3	0.066
Baseline value (pre)	150	0.82	0.71, 0.94	<0.001
Treatment				
Placebo	78	0.00	Ref.	
Intervention	72	-0.09	-2.0, 1.8	>0.9
Offense type				
Hands-on (at least one conviction §176 ff StGB)	58	0.00	Ref.	
Hands-off (only §184b StGB)	92	1.8	-0.33, 4.0	0.10
Type of supervision				
community supervision	120	0.00	Ref.	
post-release supervision	30	1.9	-0.68, 4.4	0.15
Additional treatment				
No	95	0.00	Ref.	
Yes	55	0.41	-1.6, 2.4	0.7
Static recidivism risk (baseline)	150	0.29	-0.58, 1.2	0.5

Abbreviation: CI = Confidence Interval

7.9 CUSI

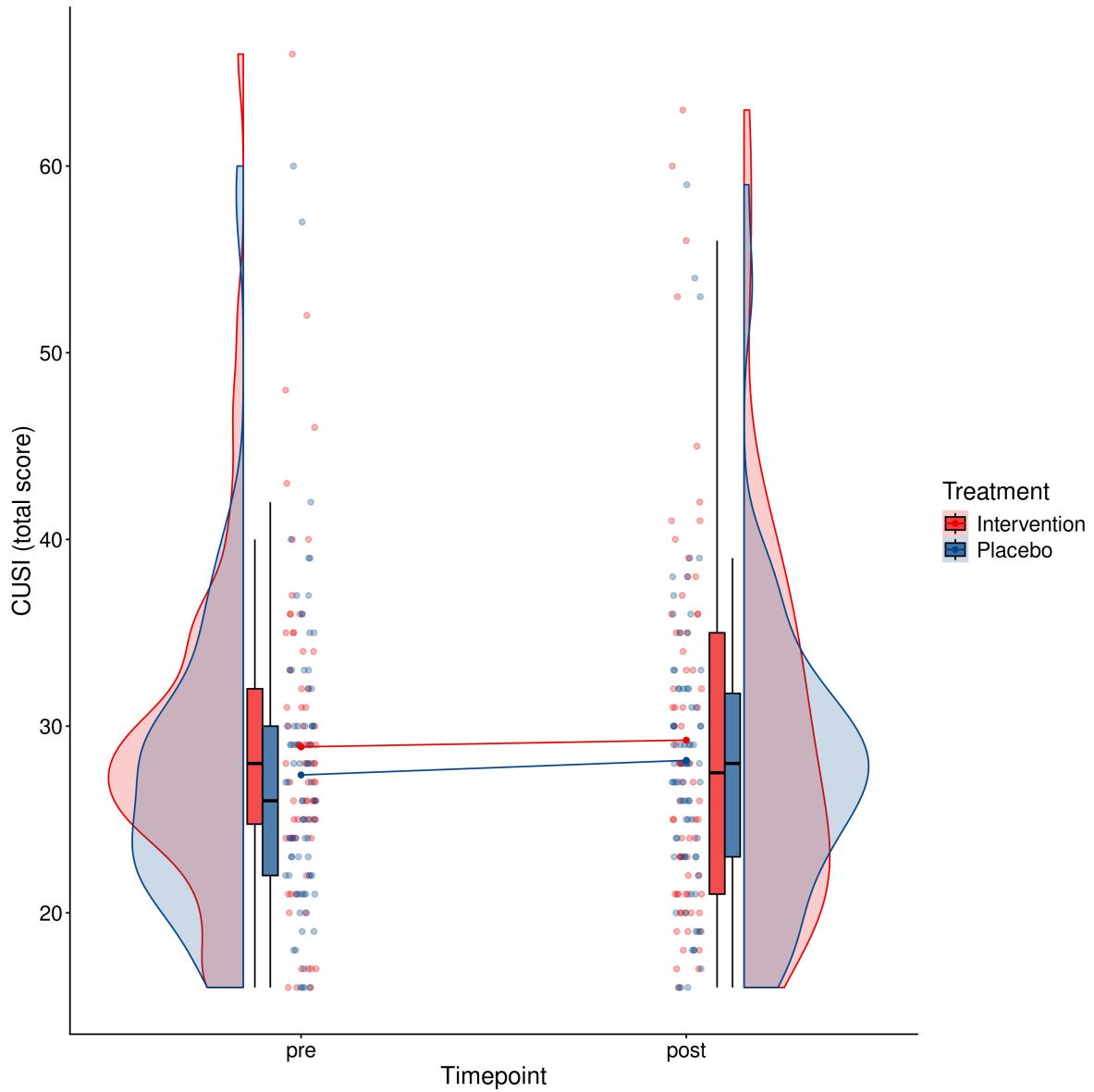


Figure 31: Rainplot of the Coping Using Sex Inventory (CUSI, total score).

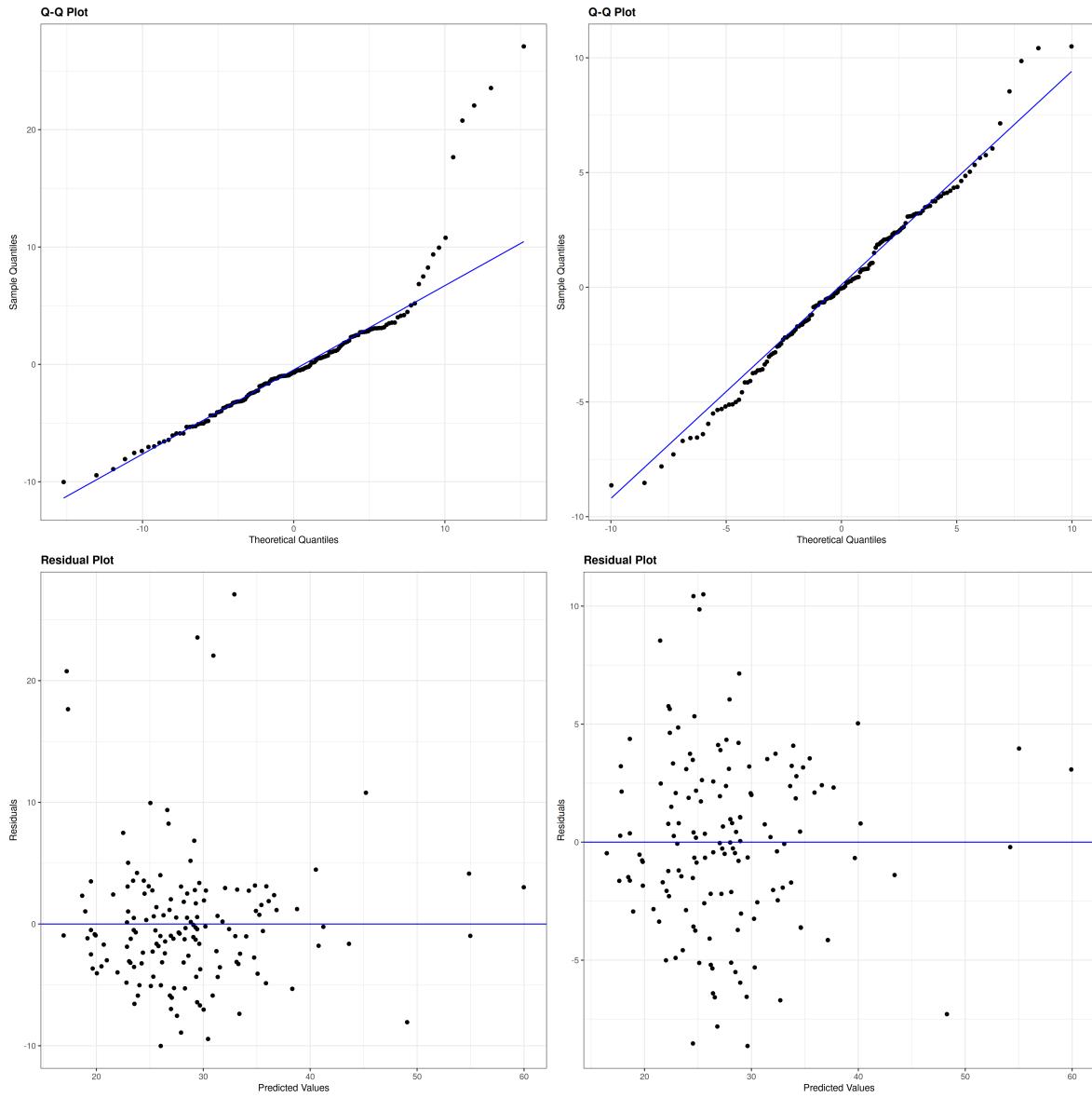


Figure 32: Diagnostic plots for Coping Using Sex Inventory (CUSI, total score). The plots on the left correspond to the model fitted with all data points, the plots on the right correspond to the model after the removal of identified outliers. Both models showed similar results.

Table 26: Linear regression model for Social Problem-Solving Inventory Revised SPSI-R (total score) with n = 134 observations.

Variable	N	95% CI	p-value	adjusted p value
(Intercept)	134	1.3	-0.24, 2.8	0.10
Baseline value (pre)	134	0.90	0.81, 1.0	<0.001
Treatment				
Placebo	72	0.00	Ref.	
Intervention	62	0.03	-0.51, 0.56	>0.9
Offense type				
Hands-on (at least one conviction §176 ff StGB)	53	0.00	Ref.	
Hands-off (only §184b StGB)	81	0.08	-0.52, 0.67	0.8
Type of supervision				
community supervision	109	0.00	Ref.	
post-release supervision	25	-0.44	-1.2, 0.30	0.2
Additional treatment				
No	88	0.00	Ref.	
Yes	46	0.23	-0.34, 0.81	0.4
Static recidivism risk (baseline)	134	0.01	-0.23, 0.25	>0.9

Abbreviation: CI = Confidence Interval

7.10 SPSI-R

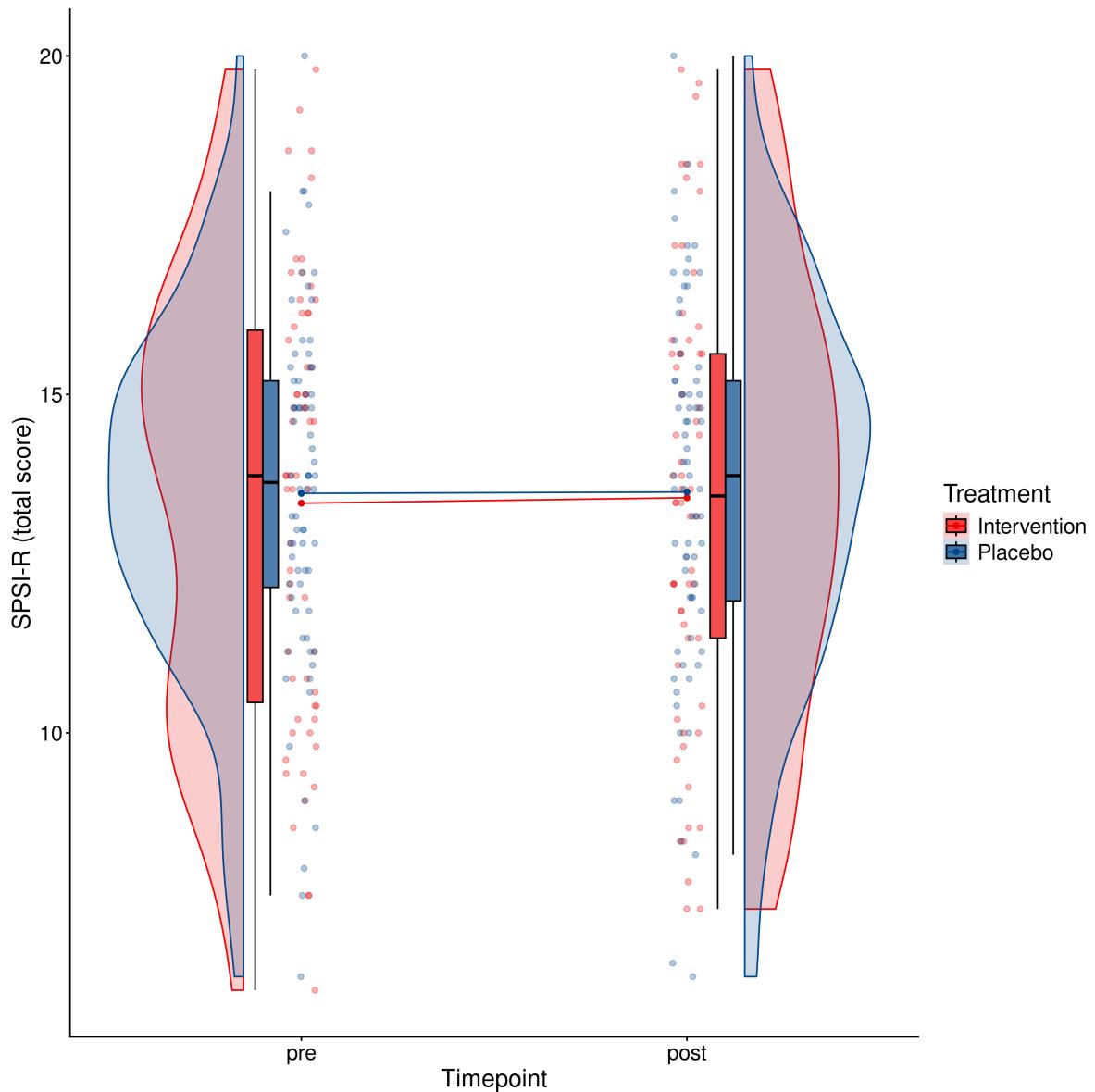


Figure 33: Rainplot of the Social Problem-Solving Inventory Revised (SPSI-R, total score).

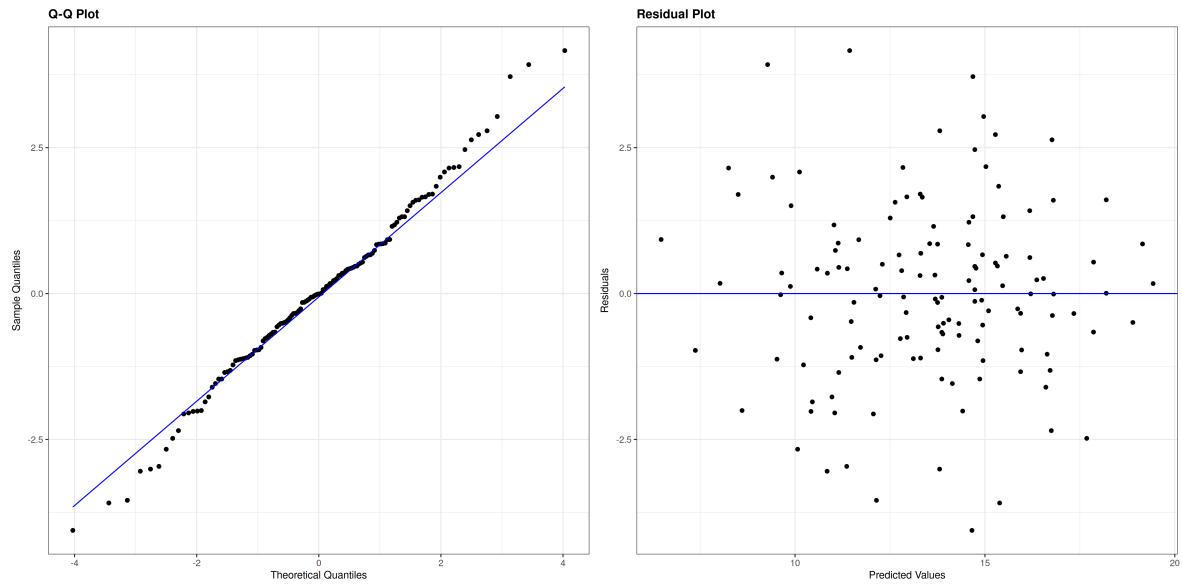


Figure 34: Diagnostic plots for the Social Problem-Solving Inventory Revised (SPSI-R, total score).

7.11 EKK-R

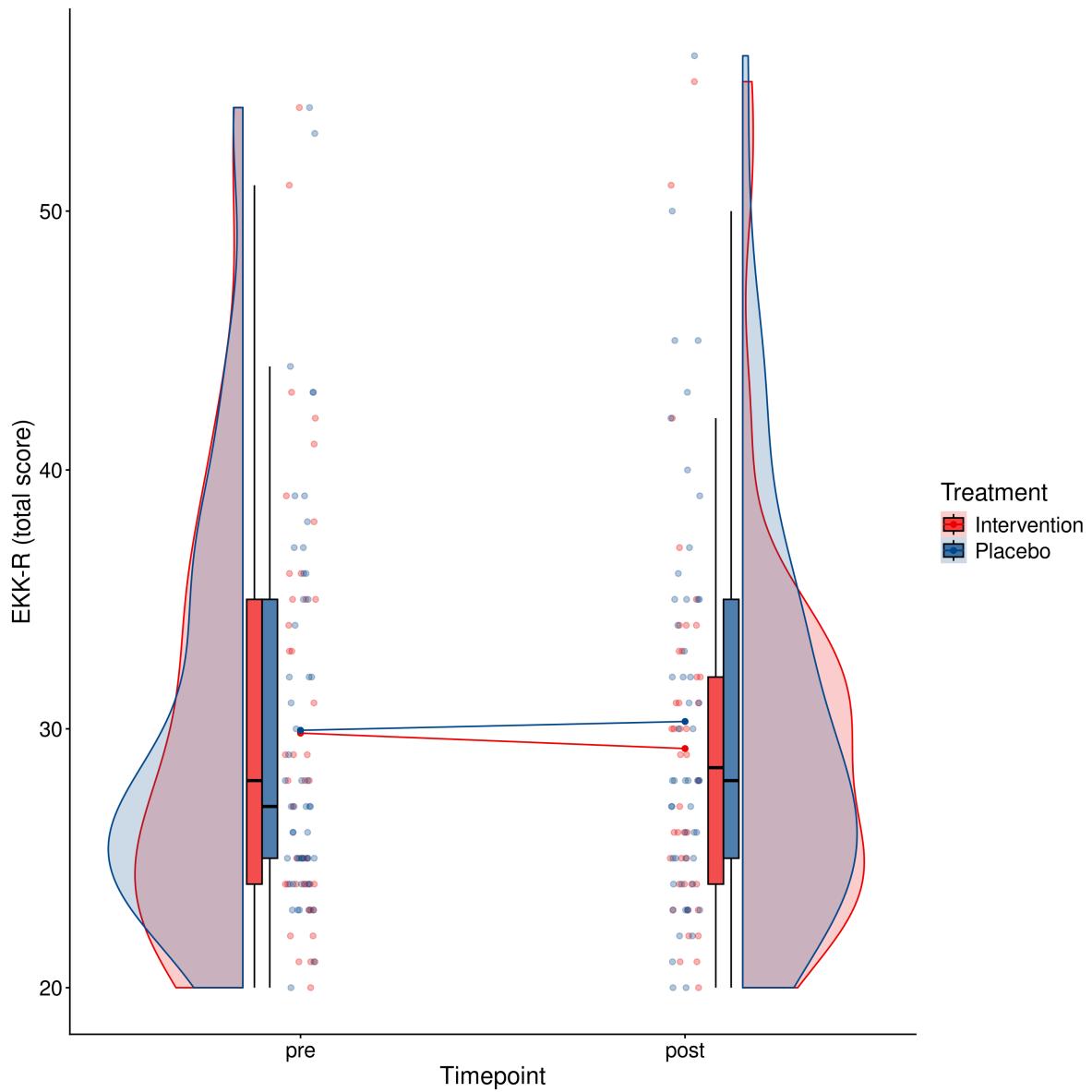


Figure 35: Rainplot of the Emotional Congruence with Children-Revised (EKK-R, total score).

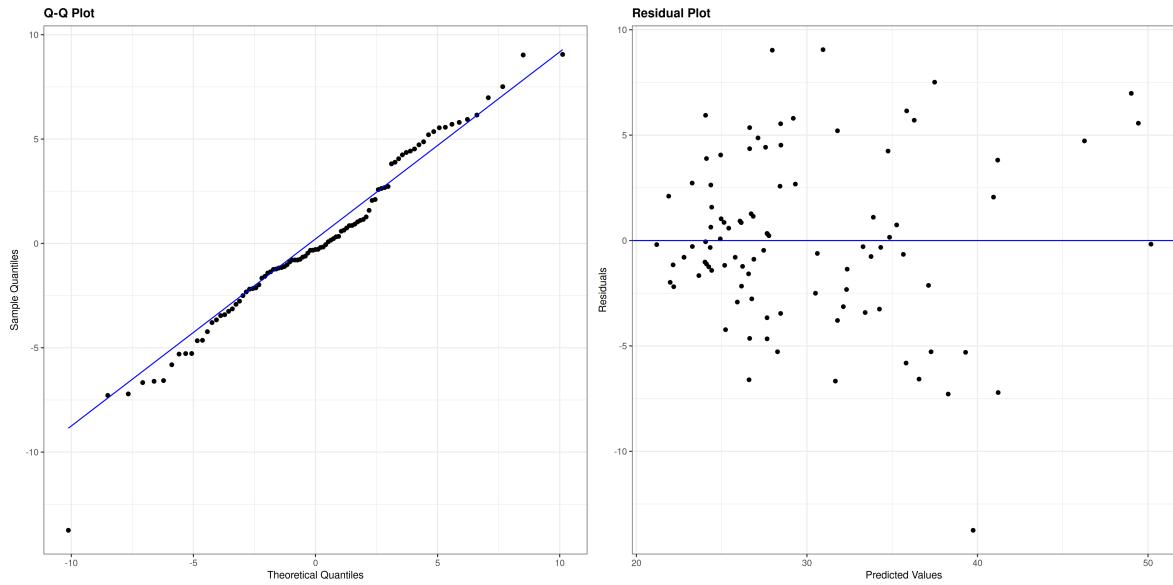


Figure 36: Diagnostic plots for Emotional Congruence with Children-Revised (EKK-R, total score).

Table 27: Linear regression model for EKK-R (total score) with n = 95 observations.

Variable	N	95% CI	p-value	adjusted p value
(Intercept)	95	5.4	1.6, 9.3	0.006
Baseline value (pre)	95	0.81	0.70, 0.92	<0.001
Treatment				
Placebo	53	0.00	Ref.	
Intervention	42	-0.90	-2.6, 0.81	0.3
Offense type				1.00
Hands-on (at least one conviction §176 ff StGB)	38	0.00	Ref.	
Hands-off (only §184b StGB)	57	0.27	-1.7, 2.2	0.8
Type of supervision				1.00
community supervision	77	0.00	Ref.	
post-release supervision	18	0.40	-2.0, 2.8	0.7
Additional treatment				1.00
No	59	0.00	Ref.	
Yes	36	0.57	-1.3, 2.5	0.6
Static recidivism risk (baseline)	95	0.04	-0.78, 0.86	>0.9

Abbreviation: CI = Confidence Interval

Table 28: Ordinal regression model for the Difficulties in Emotion Regulation Scale DERS (subscore impulsivity) with n = 150 observations. For the dependent variable, values greater than 9 were merged into a single category.

Variable	N	OR	95% CI	p-value	adjusted p value
Baseline value (pre) treatment	150	2 · 08	1 · 70, 2 · 59	< 0.001	
Placebo	78	1 · 00	Ref.		
Intervention	72	2 · 35	1 · 22, 4 · 61	0.012	0.176
Offense type					
Hands-on (at least one conviction §176 ff StGB)	58	1 · 00	Ref.		
Hands-off (only §184b StGB)	92	1 · 02	0 · 47, 2 · 29	0.955	1.000
Type of supervision					
community supervision	120	1 · 00	Ref.		
post-release supervision	30	1 · 87	0 · 74, 4 · 75	0.182	1.000
Additional treatment					
No	95	1 · 00	Ref.		
Yes	55	0 · 69	0 · 33, 1 · 41	0.313	1.000
Static recidivism risk (baseline)	150	0 · 89	0 · 64, 1 · 21	0.459	1.000

Abbreviations: CI = Confidence Interval, OR = Odds Ratio

7.12 DERS

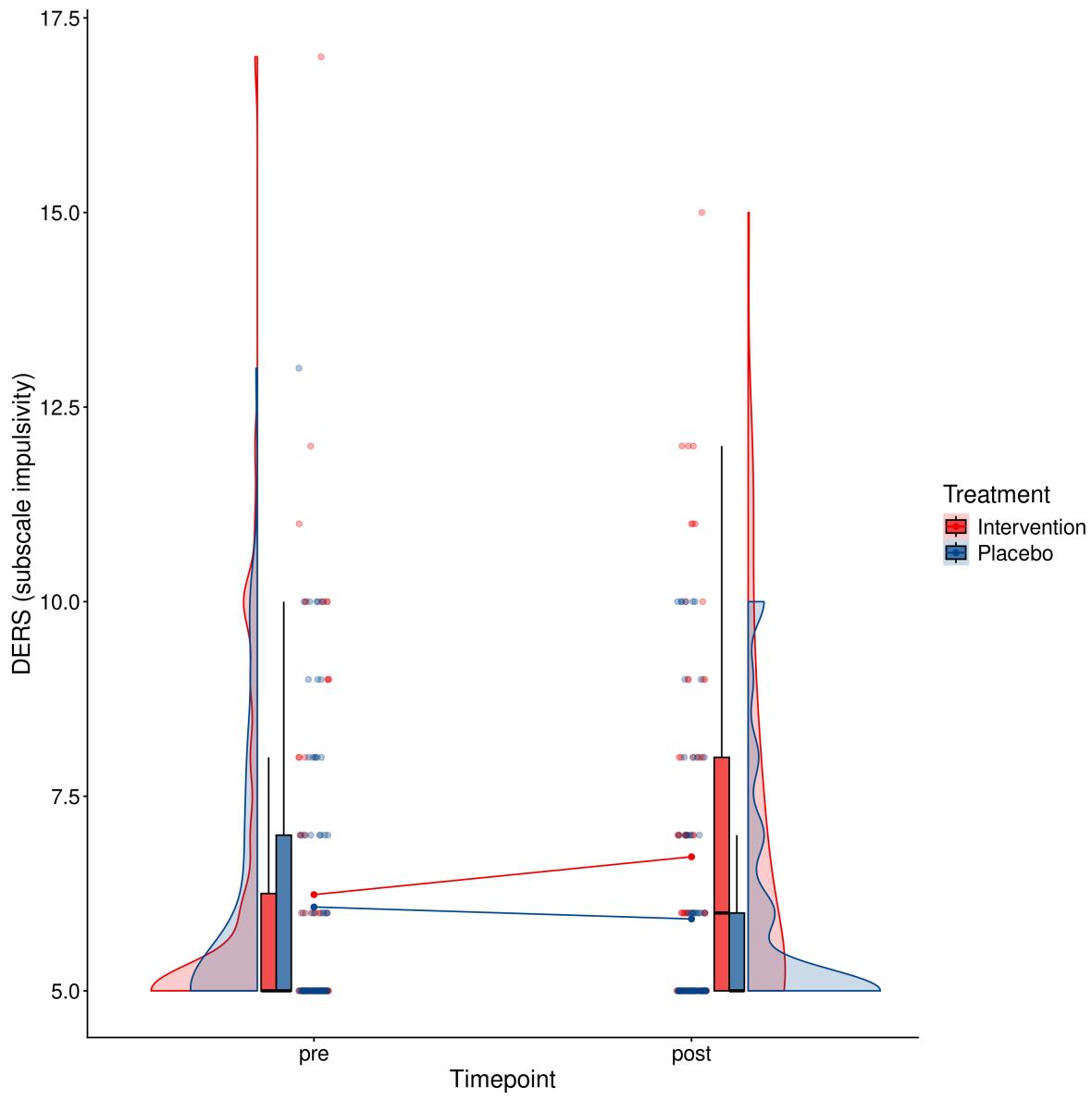


Figure 37: Rainplot of the Difficulties in Emotion Regulation Scale (DERS, subscale impulsivity). Note, that for the dependent variable, values greater than 9 were merged into a single category.

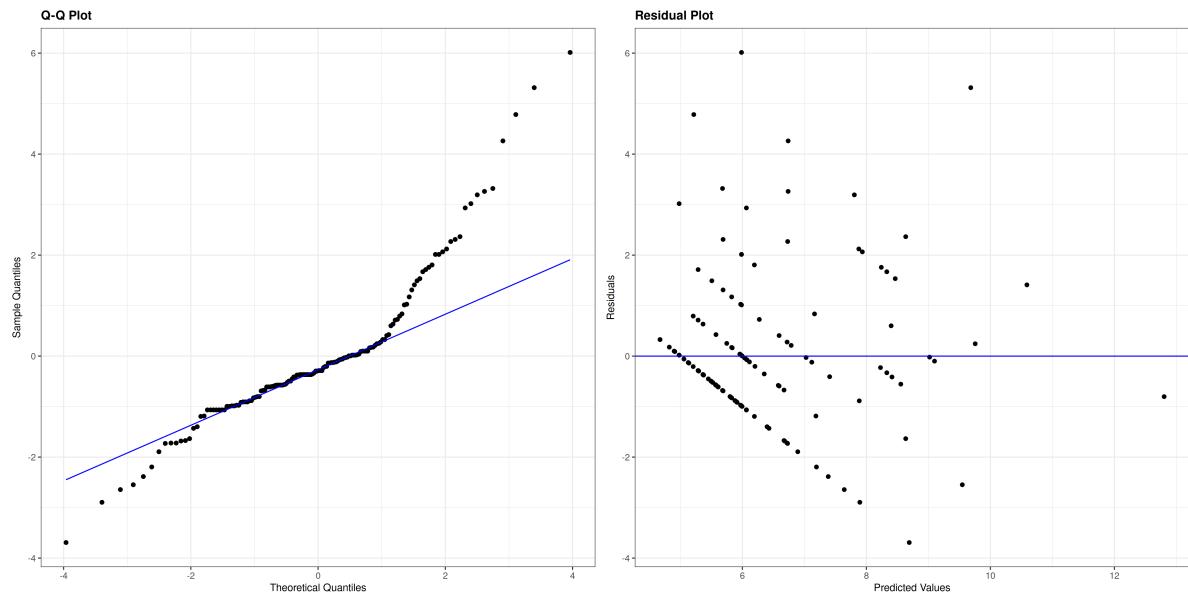


Figure 38: Diagnostic plots for Difficulties in Emotion Regulation Scale (DERS, subscale impulsivity).
Note, that for the dependent variable, values greater than 9 were merged into a single category..

7.13 NARQ

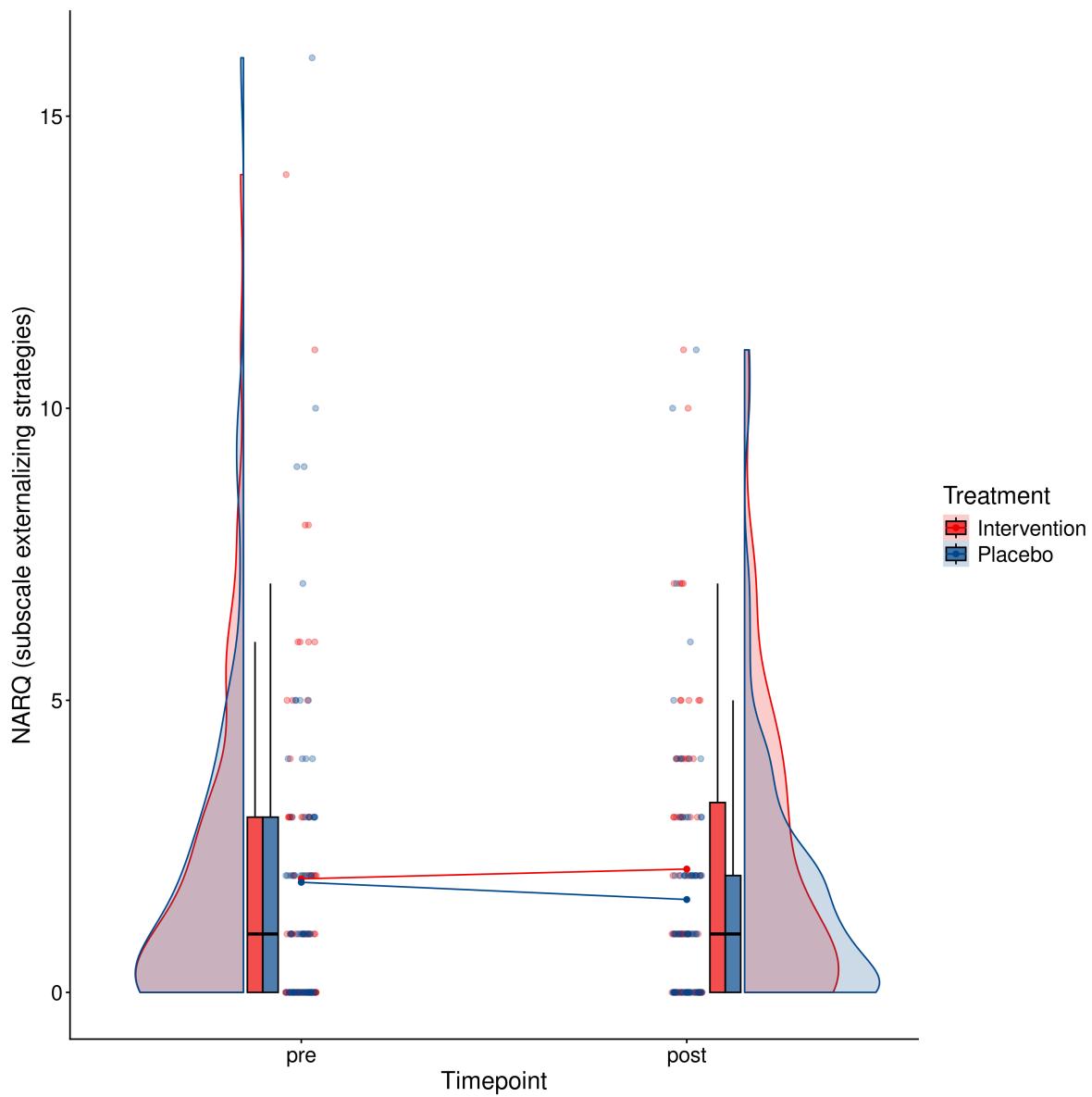


Figure 39: Rainplot of the Negative Affect Repair Questionnaire (NARQ, subscale externalizing strategies). Note, that for the dependent variable, values greater than 4 were merged into a single category.

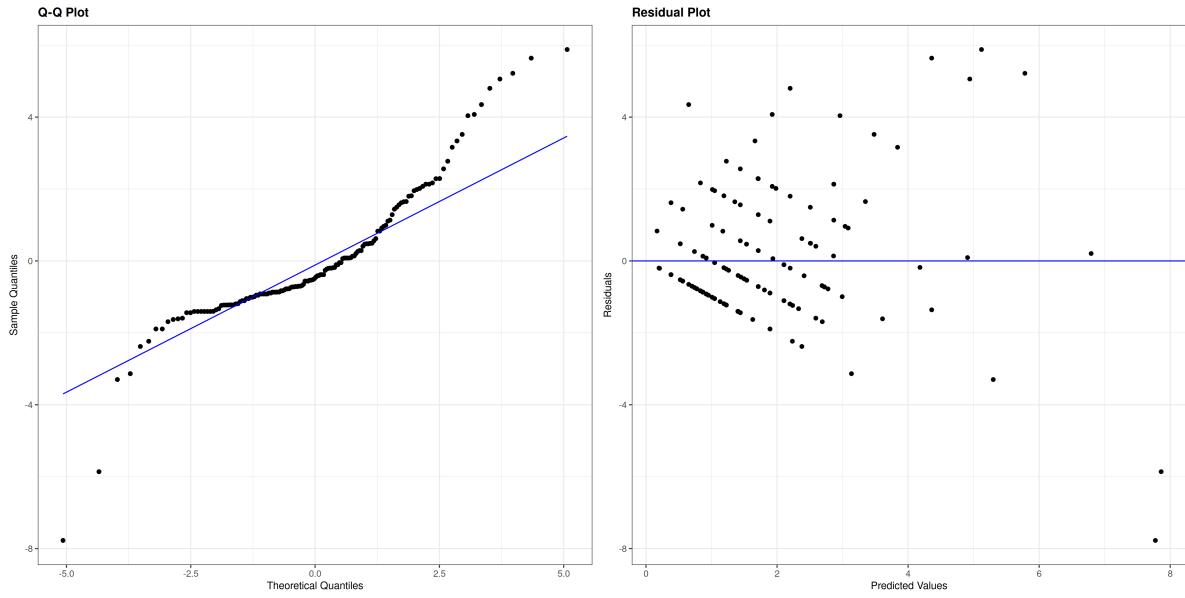


Figure 40: Diagnostic plots for the Negative Affect Repair Questionnaire (NARQ, subscale externalizing strategies). Note, that for the dependent variable, values greater than 4 were merged into a single category

Table 29: Ordinal regression model for the Negative Affect Repair Questionnaire NARQ (subscale externalizing strategies) with n = 150 observations. Note, that for the dependent variable, values greater than 4 were merged into a single category.

Variable	N	OR	95% CI	p-value	adjusted p value
Baseline value (pre treatment)	150	1 · 69	1 · 44, 2 · 02	< 0.001	
Placebo	78	1 · 00	Ref.		
Intervention	72	1 · 46	0 · 81, 2 · 66	0.212	1.00
Offense type					
Hands-on (at least one conviction §176 ff StGB)	58	1 · 00	Ref.		
Hands-off (only §184b StGB)	92	1 · 35	0 · 68, 2 · 71	0.395	1.00
Type of supervision					
community supervision	120	1 · 00	Ref.		
post-release supervision	30	1 · 04	0 · 47, 2 · 32	0.922	1.00
Additional treatment					
No	95	1 · 00	Ref.		
Yes	55	0 · 98	0 · 52, 1 · 84	0.952	1.00
Static recidivism risk (baseline)	150	0 · 85	0 · 63, 1 · 13	0.273	1.00

Abbreviations: CI = Confidence Interval, OR = Odds Ratio

Table 30: Ordinal regression model for the Explicit Sexual Interest Questionnaire ESIQ (total score) with n = 95 observations. Note, that the dependent variable has 3 ordered categories: worsening (increase), no change, improvement (decrease).

Variable	N	OR	95% CI	p-value	adjusted p value
treatment					
Placebo	53	1 · 00	Ref.		
Intervention	42	4 · 36	1 · 93, 10 · 3	< 0.001	0.009
Offense type					
Hands-on (at least one conviction §176 ff StGB)	38	1 · 00	Ref.		
Hands-off (only §184b StGB)	57	0 · 87	0 · 34, 2 · 19	0.764	1.000
Type of supervision					
community supervision	77	1 · 00	Ref.		
post-release supervision	18	0 · 88	0 · 29, 2 · 75	0.822	1.000
Additional treatment					
No	59	1 · 00	Ref.		
Yes	36	1 · 19	0 · 49, 2 · 95	0.701	1.000
Static recidivism risk (baseline)					
	95	0 · 98	0 · 66, 1 · 45	0.904	1.000

Abbreviations: CI = Confidence Interval, OR = Odds Ratio

7.14 ESIQ

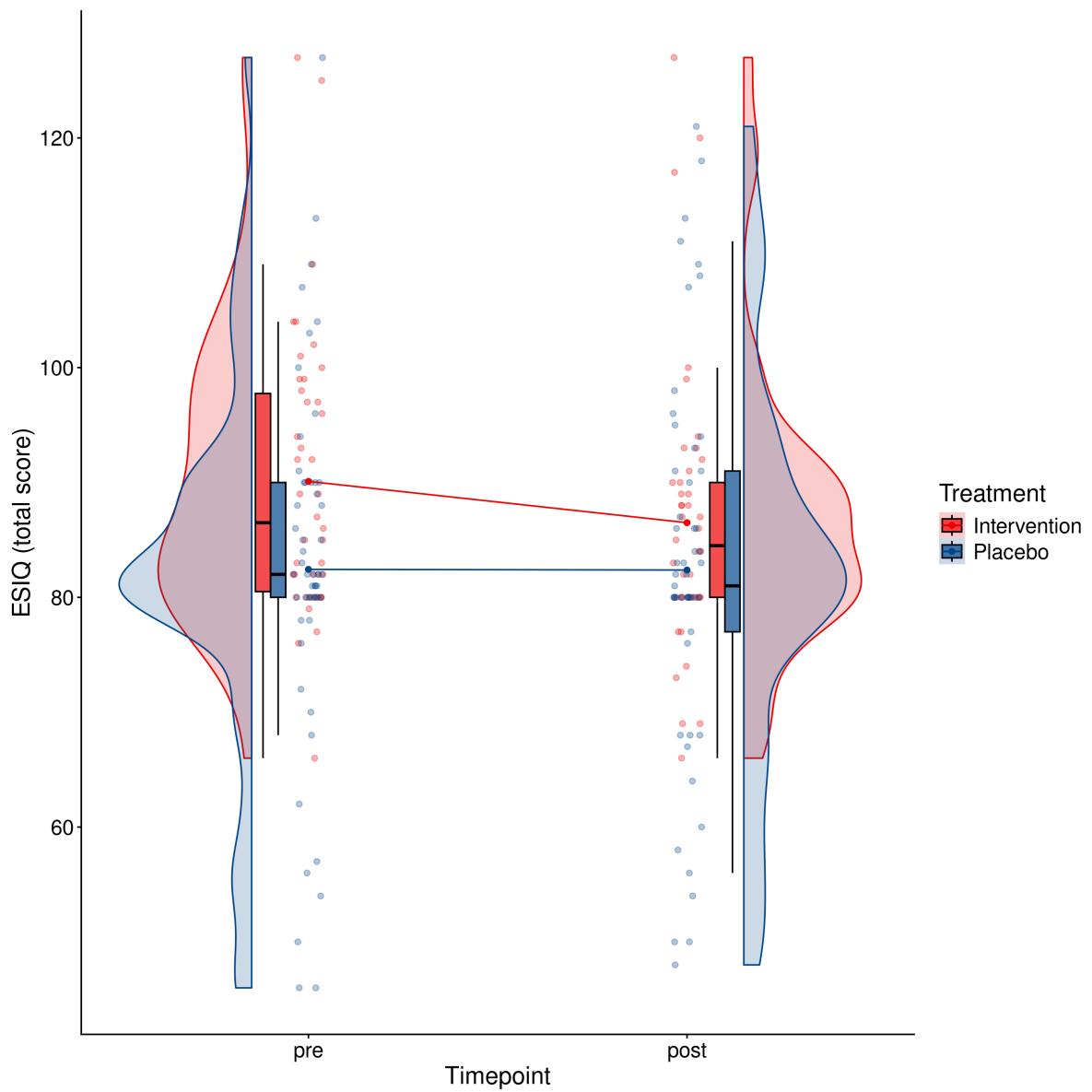


Figure 41: Rainplot of the Explicit Sexual Interest Questionnaire (ESIQ, total score). Note, that the dependent variable has 3 ordered categories: worsening (increase), no change, improvement (decrease).

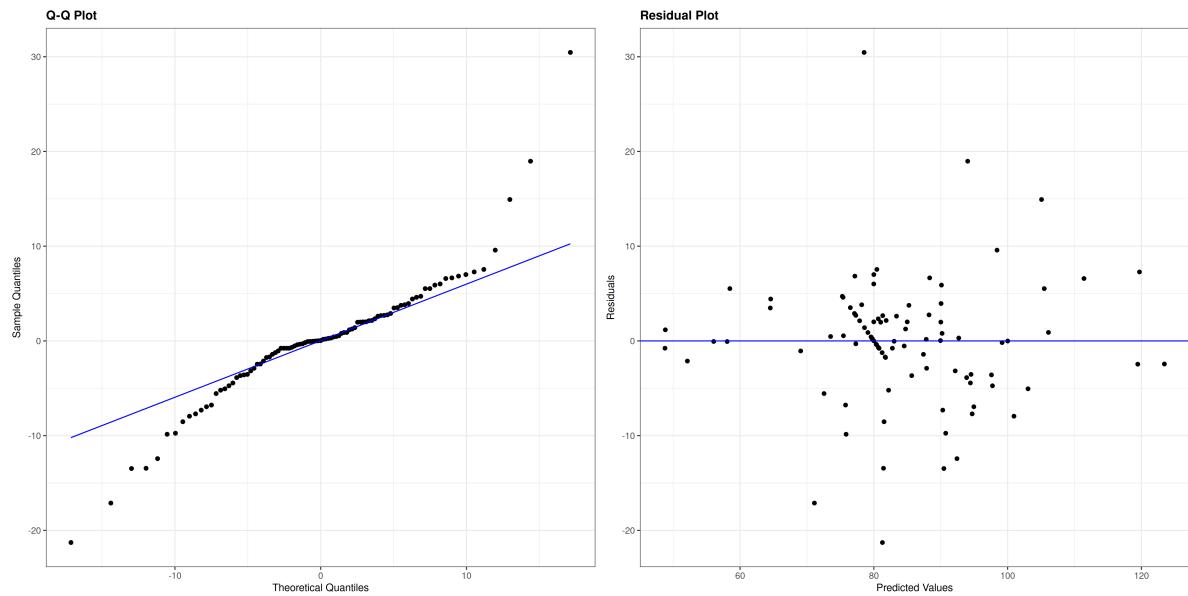


Figure 42: Diagnostic plots for the Explicit Sexual Interest Questionnaire (ESIQ, total score). Note, that the dependent variable has 3 ordered categories: worsening (increase), no change, improvement (decrease)

7.15 BMS

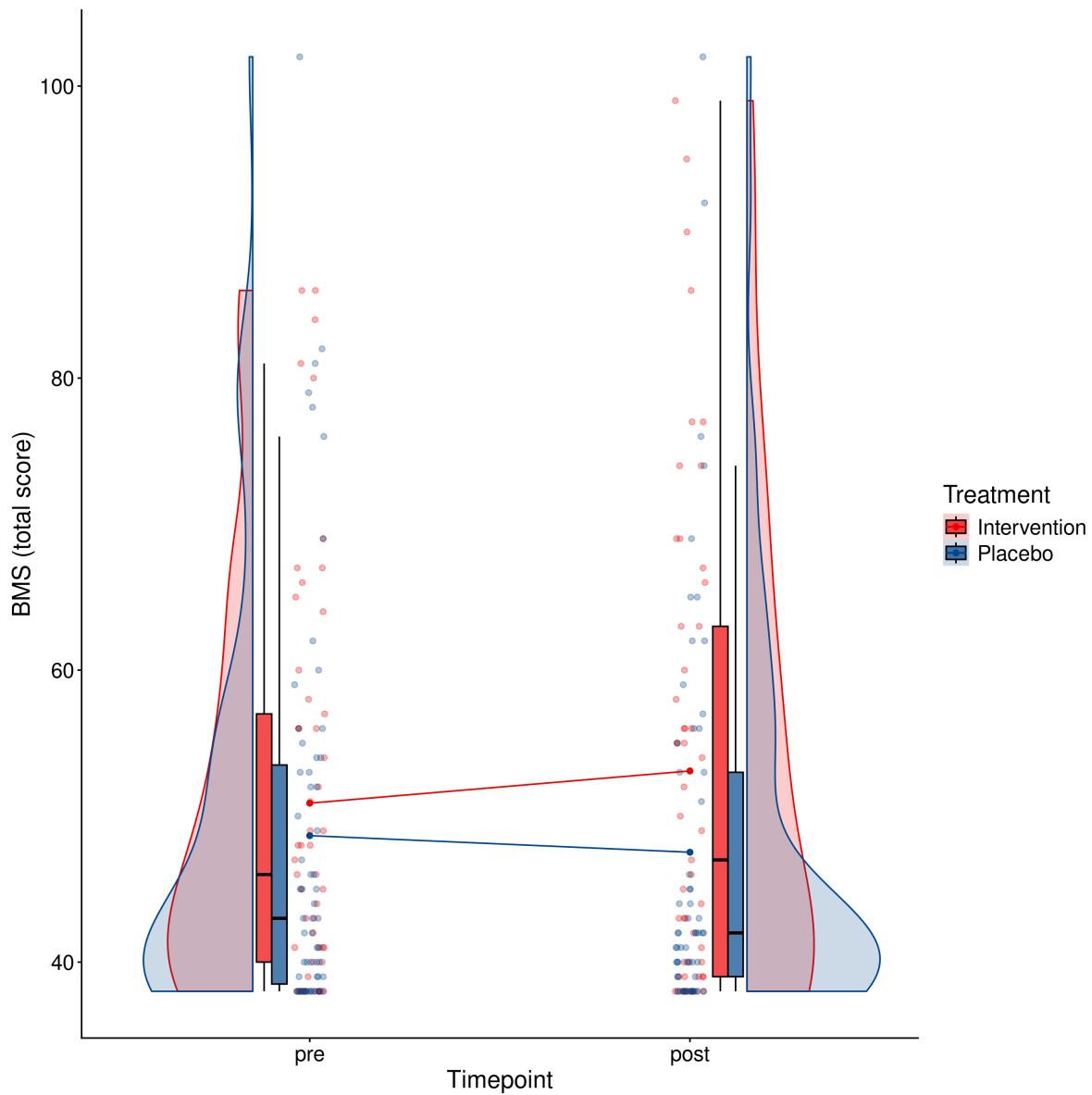


Figure 43: Rainplot of the Bumby Molest Scale (BMS, total score). Note, that for the dependent variable, values were dichotomized based on the cut-off greater or equal to 43.

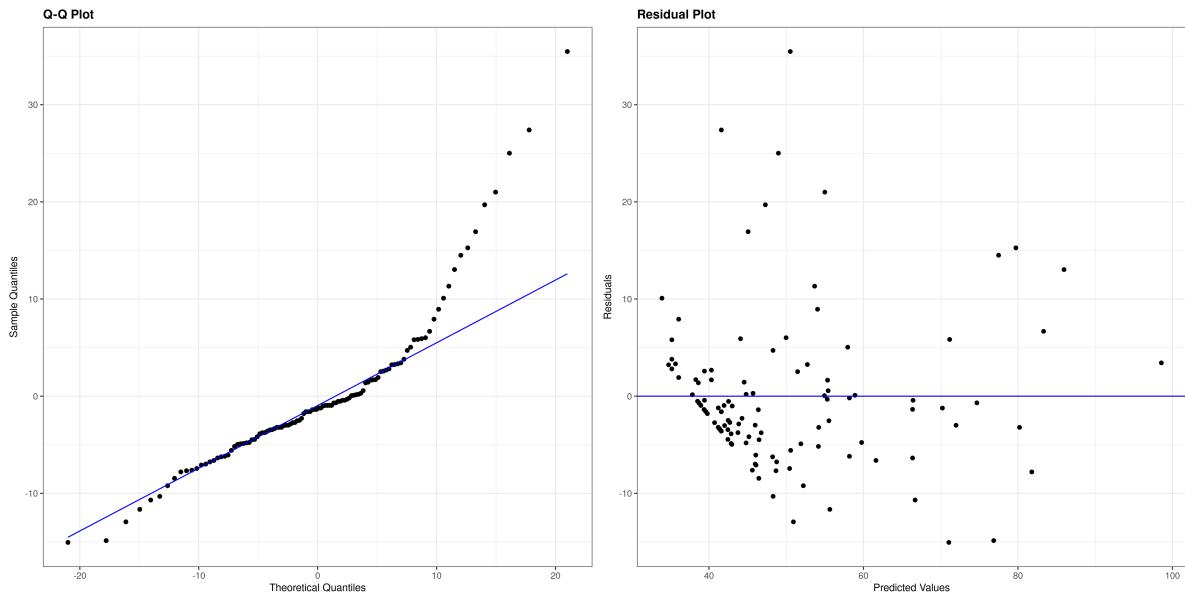


Figure 44: Diagnostic plots for the Bumby Molest Scale (BMS, total score). Note, that for the dependent variable, values were dichotomized based on the cut-off greater or equal to 43.

Table 31: Logistic regression model for the Bumby Molest Scale BMS (total score) with $n = 112$ observations. Note, that for the dependent variable, values were dichotomized based on the cut-off greater or equal to 43.

Variable	N	Event N	OR	95% CI	p-value	adjusted p value
Baseline value (pre)						
FALSE	48	6	1 · 00	Ref.		
TRUE	64	51	31 · 2	10 · 7, 111	< 0.001	
Treatment						
Placebo	59	24	1 · 00	Ref.		
Intervention	53	33	3 · 51	1 · 22, 11 · 3	0.025	0.348
Offense type						
Hands-on (at least one conviction §176 ff StGB)	43	27	1 · 00	Ref.		
Hands-off (only §184b StGB)	69	30	0 · 47	0 · 13, 1 · 60	0.229	1.000
Type of supervision						
community supervision	91	45	1 · 00	Ref.		
post-release supervision	21	12	1 · 27	0 · 28, 6 · 28	0.764	1.000
Additional treatment						
No	70	32	1 · 00	Ref.		
Yes	42	25	1 · 60	0 · 52, 5 · 08	0.415	1.000
Static recidivism risk (baseline)	112	57	1 · 05	0 · 61, 1 · 77	0.866	1.000

Abbreviations: CI = Confidence Interval, OR = Odds Ratio

Table 32: Logistic regression model for the Hypersexual Behavior Inventory-19 HBI-19 (total score) with n = 95 observations. Note, that for the dependent variable, values were dichotomized based on the cut-off greater or equal to 24.

Variable	N	Event N	OR	95% CI	p-value	adjusted p value
Baseline value (pre)						
FALSE	50	13	1 · 00	Ref.		
TRUE	45	38	16 · 9	5 · 95, 55 · 5	< 0.001	
Treatment						
Placebo	53	24	1 · 00	Ref.		
Intervention	42	27	2 · 84	0 · 98, 8 · 84	0.060	0.782
Offense type						
Hands-on (at least one conviction §176 ff StGB)	38	18	1 · 00	Ref.		
Hands-off (only §184b StGB)	57	33	2 · 48	0 · 74, 8 · 92	0.148	1.000
Type of supervision						
community supervision	77	42	1 · 00	Ref.		
post-release supervision	18	9	0 · 65	0 · 14, 2 · 97	0.572	1.000
Additional treatment						
No	59	29	1 · 00	Ref.		
Yes	36	22	2 · 63	0 · 85, 8 · 73	0.102	1.000
Static recidivism risk (baseline)	95	51	0 · 72	0 · 43, 1 · 17	0.190	1.000

Abbreviations: CI = Confidence Interval, OR = Odds Ratio

7.16 HBI-19

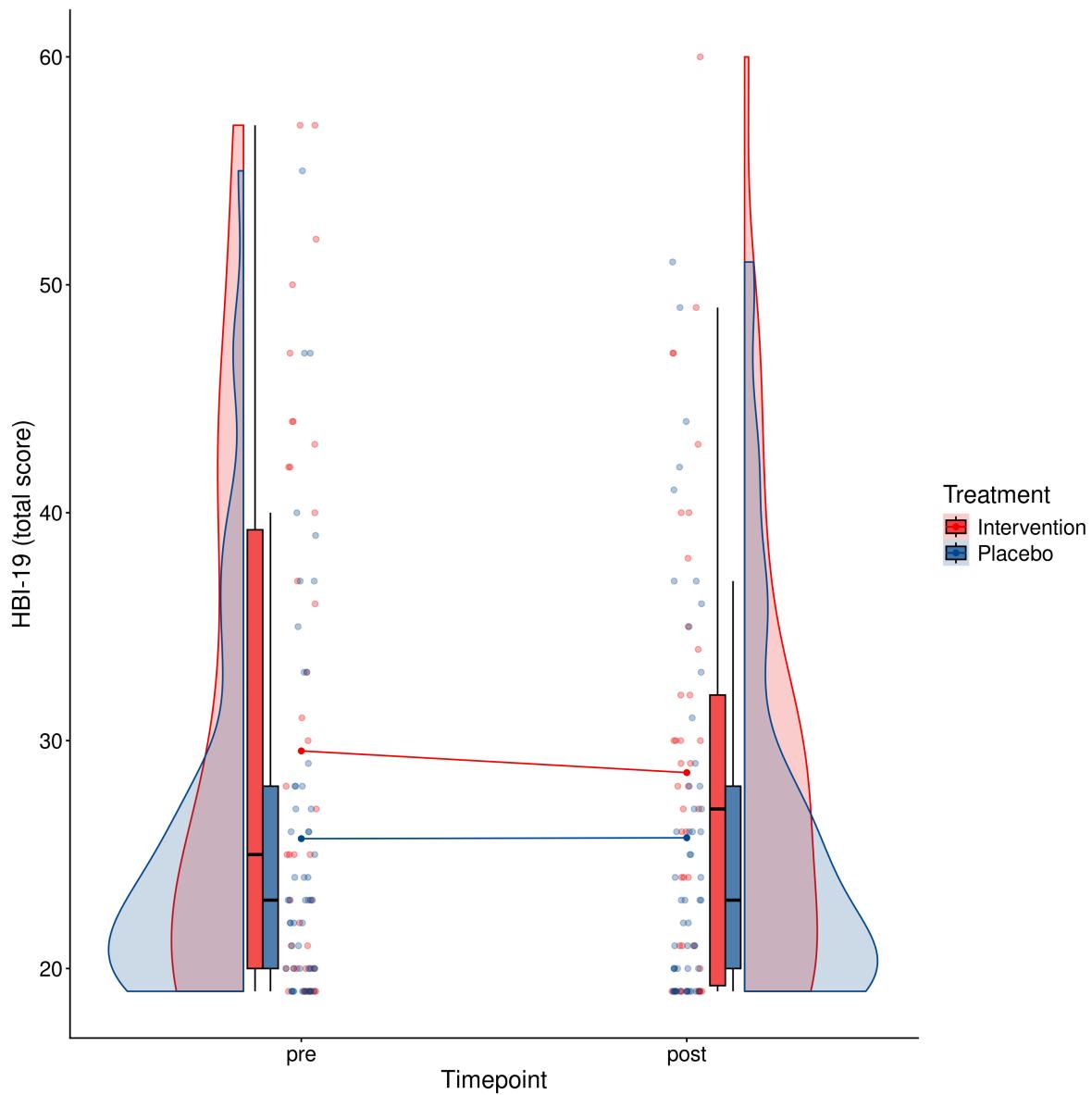


Figure 45: Rainplot of the Hypersexual Behavior Inventory-19 (HBI-19, total score). Note, that for the dependent variable, values were dichotomized based on the cut-off greater or equal to 24.

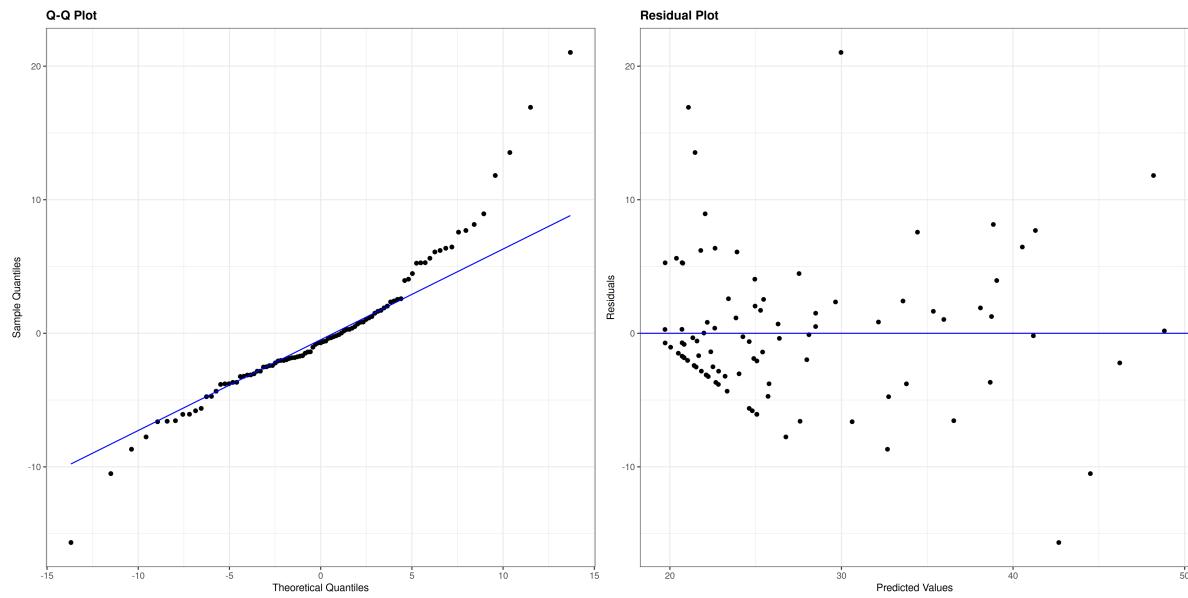


Figure 46: Diagnostic plots for the Hypersexual Behavior Inventory-19 (HBI-19, total score). Note, that for the dependent variable, values were dichotomized based on the cut-off greater or equal to 24.

7.17 SSIC

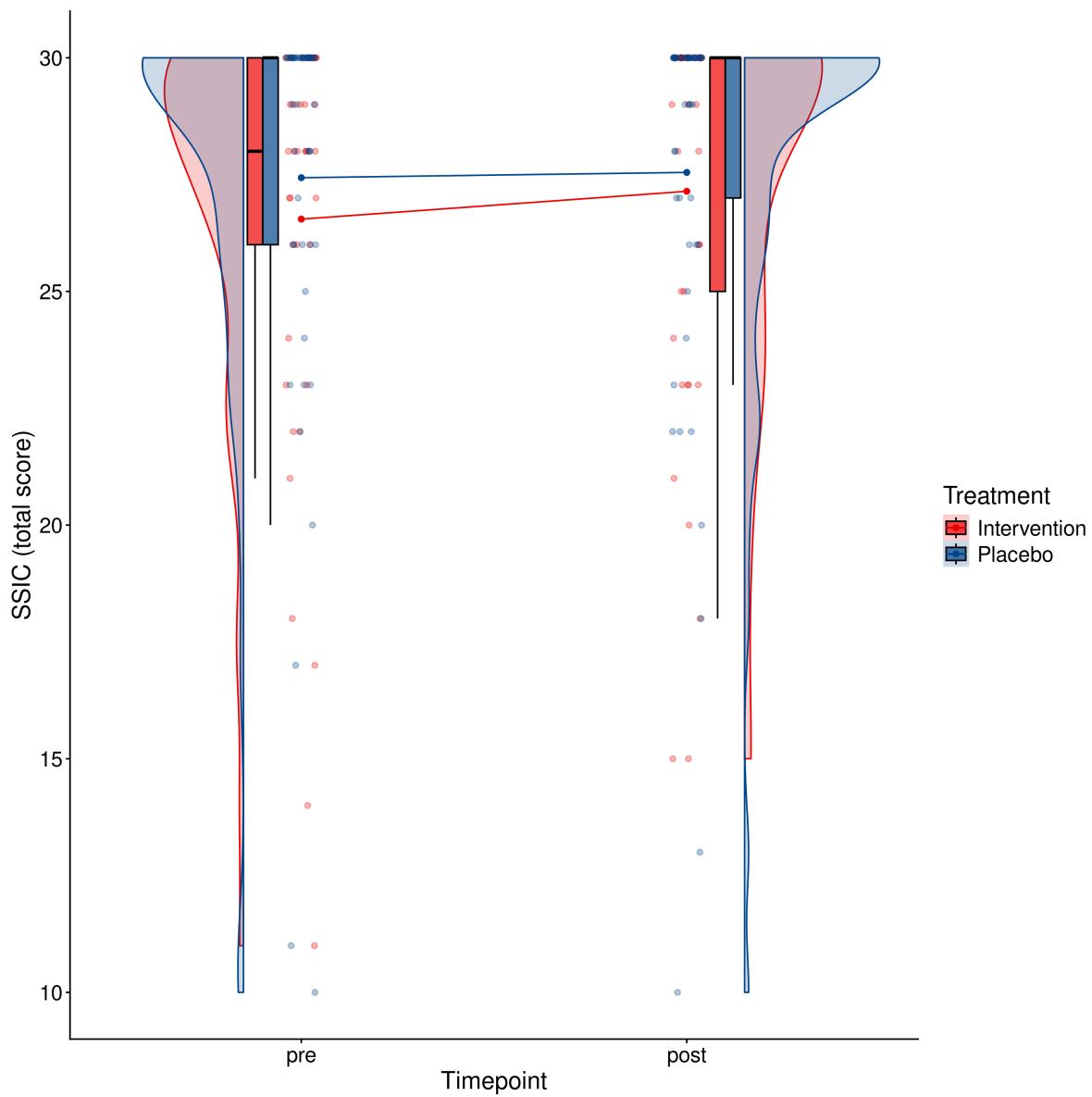


Figure 47: Rainplot of the Specific self-efficacy for modifying Sexual Interest in Children (SSIC, total score). Note, that for the dependent variable, values were dichotomized based on the cut-off greater or equal to 30.

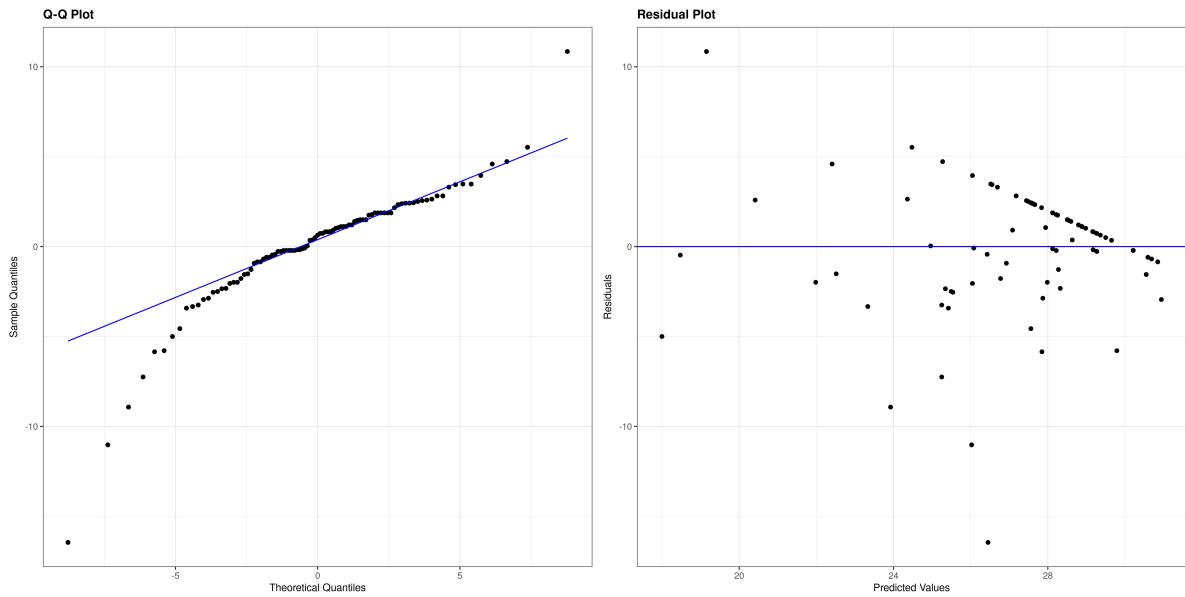


Figure 48: Diagnostic plots for the Specific self-efficacy for modifying Sexual Interest in Children (SSIC, total score). Note, that for the dependent variable, values were dichotomized based on the cut-off greater or equal to 30.

Table 33: Logistic regression model for the Specific self-efficacy for modifying Sexual Interest in Children SSIC (total score) with n = 95 observations. Note, that for the dependent variable, values were dichotomized based on the cut-off greater or equal to 30.

Variable	N	Event N	OR	95% CI	p-value	adjusted p value
Baseline value (pre)						
FALSE	51	18	1 · 00	Ref.		
TRUE	44	36	10 · 4	3 · 77, 33 · 0	< 0.001	
Treatment						
Placebo	53	31	1 · 00	Ref.		
Intervention	42	23	1 · 51	0 · 57, 4 · 24	0.411	1.00
Offense type						
Hands-on (at least one conviction §176 ff StGB)	38	23	1 · 00	Ref.		
Hands-off (only §184b StGB)	57	31	0 · 51	0 · 16, 1 · 54	0.239	1.00
Type of supervision						
community supervision	77	42	1 · 00	Ref.		
post-release supervision	18	12	0 · 86	0 · 21, 3 · 46	0.834	1.00
Additional treatment						
No	59	31	1 · 00	Ref.		
Yes	36	23	1 · 15	0 · 40, 3 · 25	0.797	1.00
Static recidivism risk (baseline)	95	54	1 · 04	0 · 66, 1 · 64	0.868	1.00

Abbreviations: CI = Confidence Interval, OR = Odds Ratio

8 Adverse events

8.1 Mood and Risk Questionnaire (MRQ)

Table 34: Frequencies of adverse events during working on the WBI. The MRQ Post questionnaire explicitly asks for the time during working on one session. Shown are the frequency of participants reported at least once the highest possible score (= adverse event) on one of the measurement time-points (after each session of the WBI).

MRQ subscale	Overall N = 221 ¹	Intervention N = 108 ¹	Placebo N = 113 ¹	Cramer's V	p	add_stat_1	q-value
Contact planning	2 (0 · 9%)	1 (0 · 9%)	1 (0 · 9%)	0 · 002	>0 · 99		>0 · 99
Contact preparation	0 (0%)	0 (0%)	0 (0%)				
Urge for CSA	0 (0%)	0 (0%)	0 (0%)				
Urge for CSAM	0 (0%)	0 (0%)	0 (0%)				
Sexual tension	2 (0 · 9%)	2 (1 · 9%)	0 (0%)	0 · 098	0 · 24		>0 · 99
Control of sexual thoughts and activity	1 (0 · 5%)	0 (0%)	1 (0 · 9%)	0 · 066	>0 · 99		>0 · 99
Unable to cope with the mental burden	9 (4 · 1%)	7 (6 · 5%)	2 (1 · 8%)	0 · 119	0 · 10		0 · 57
Suicidal ideations	2 (0 · 9%)	2 (1 · 9%)	0 (0%)	0 · 098	0 · 24		>0 · 99
Mental crisis	15 (6 · 8%)	11 (10%)	4 (3 · 5%)	0 · 132	0 · 062		0 · 44
Very bad current mood	12 (5 · 4%)	8 (7 · 4%)	4 (3 · 5%)	0 · 085	0 · 24		>0 · 99

¹ n (%)

Abbreviation: p_{adj} = Holm-Bonferroni adjusted p.

Table 35: Frequencies of adverse events between working on sessions of the WBI. The MRQ pre questionnaire explicitly asks for the last week before beginning a new session. Shown are the frequency of participants reported at least once the highest possible score (= adverse event) on one of the measurement time-points (after each session of the WBI).

MRQ subscale	Overall N = 221 ¹	Intervention N = 108 ¹	Placebo N = 113 ¹	Cramer's V	p	add_stat_1	q-value
Contact planning	3 (1 · 4%)	3 (2 · 8%)	0 (0%)	0 · 120	0 · 12		>0 · 99
Contact preparation	1 (0 · 5%)	1 (0 · 9%)	0 (0%)	0 · 069	0 · 49		>0 · 99
Urge for CSA	0 (0%)	0 (0%)	0 (0%)				
Urge for CSAM	2 (0 · 9%)	1 (0 · 9%)	1 (0 · 9%)	0 · 002	>0 · 99		>0 · 99
Sexual tension	8 (3 · 6%)	6 (5 · 6%)	2 (1 · 8%)	0 · 101	0 · 16		>0 · 99
Control of sexual thoughts and activity	2 (0 · 9%)	2 (1 · 9%)	0 (0%)	0 · 098	0 · 24		>0 · 99
Unable to cope with the mental burden	20 (9 · 0%)	11 (10%)	9 (8 · 0%)	0 · 039	0 · 64		>0 · 99
Suicidal ideations	4 (1 · 8%)	1 (0 · 9%)	3 (2 · 7%)	0 · 065	0 · 62		>0 · 99
Mental crisis	16 (7 · 2%)	11 (10%)	5 (4 · 4%)	0 · 111	0 · 12		>0 · 99
Very bad current mood	8 (3 · 6%)	5 (4 · 6%)	3 (2 · 7%)	0 · 053	0 · 49		>0 · 99

¹ n (%)

Abbreviation: p_{adj} = Holm-Bonferroni adjusted p.

Table 36: Frequencies of exclusion reasons reported by the supervision officers. Exclusion reasons have been assessed in a structured manner by the Exclusion Questionnaire (EQ).

Exclusion reason	Overall N = 124 ¹	Intervention N = 66 ¹	Placebo N = 58 ¹	Cramer's V	p	q-value
Other reasons	17 (14%)	10 (15%)	7 (12%)	0 · 040	0 · 79	>0 · 99
Withdrawal of informed consent	7 (5 · 6%)	4 (6 · 1%)	3 (5 · 2%)	0 · 020	>0 · 99	>0 · 99
Probationary supervision expired	22 (18%)	8 (12%)	14 (24%)	0 · 160	0 · 10	0 · 91
CSA offense, Other offense (not CSA or CSEM)	1 (0 · 8%)	0 (0%)	1 (1 · 7%)	0 · 100	0 · 47	>0 · 99
CSA and CSEM offense	2 (1 · 6%)	1 (1 · 5%)	1 (1 · 7%)	0 · 010	>0 · 99	>0 · 99
Other offense (not CSA or CSEM)	3 (2 · 4%)	3 (4 · 5%)	0 (0%)	0 · 150	0 · 25	>0 · 99
CSA offense	1 (0 · 8%)	0 (0%)	1 (1 · 7%)	0 · 100	0 · 47	>0 · 99
CSEM offense	1 (0 · 8%)	1 (1 · 5%)	0 (0%)	0 · 080	>0 · 99	>0 · 99
End of RCT	70 (56%)	39 (59%)	31 (53%)	0 · 060	0 · 59	>0 · 99

¹ n (%)

Abbreviation: p_{adj} = Holm-Bonferroni adjusted p.

8.2 Exclusion reasons

9 Ancillary Outcomes

9.1 WHO-5

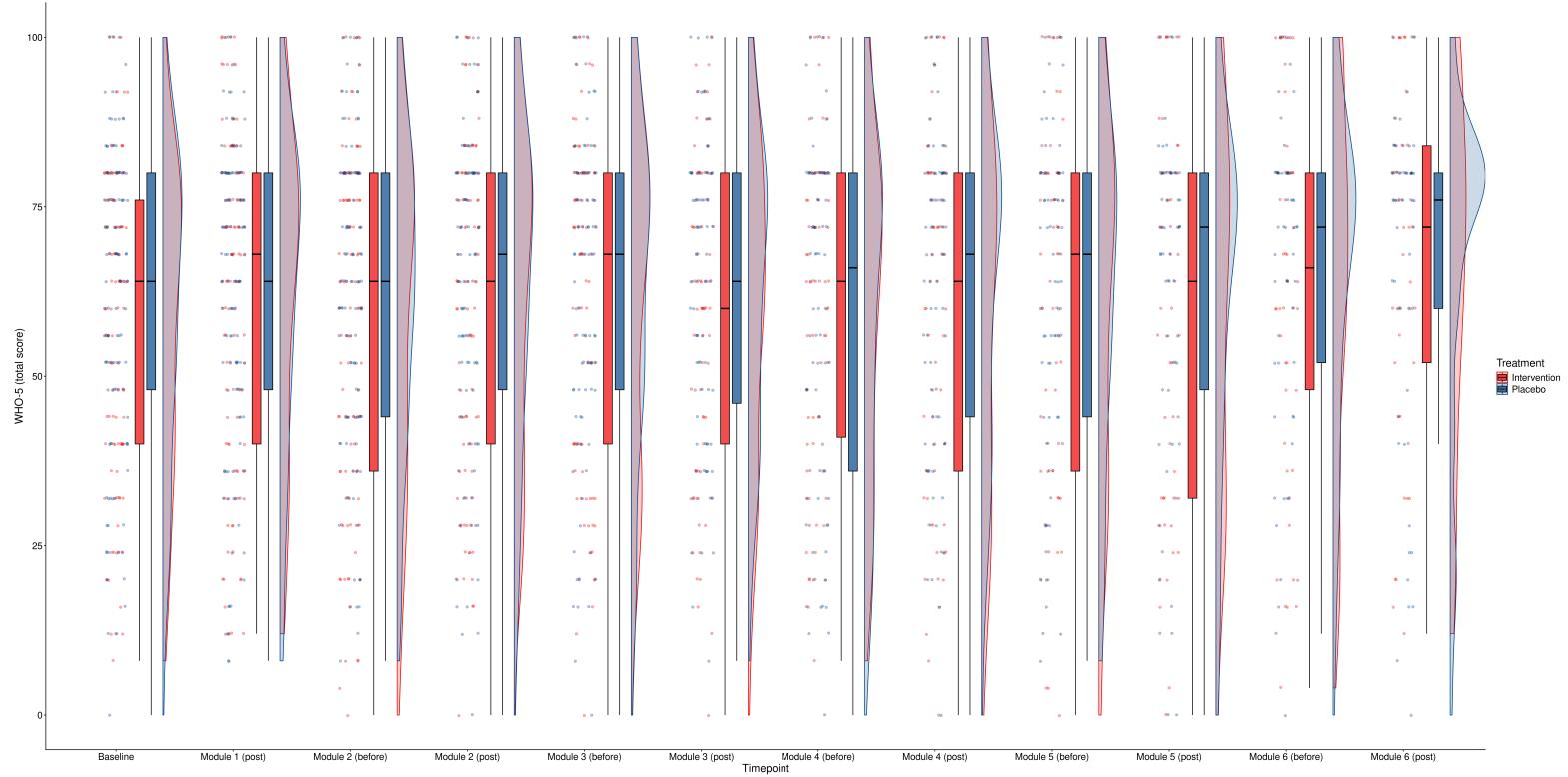


Figure 49: Rainplot for the WHO-5 Well-Being Index (WHO-5, total score).

Table 37: Mixed model for repeated measures (MMRM) for the WHO-5 Well-Being Index (WHO-5, total score).

Variable	Beta	95% CI	p-value
(Intercept)	22	14 to 29	<0 · 0001
Type of supervision			
community supervision	0 · 00	Ref.	
post-release supervision	-2 · 9	-8 · 2 to 2 · 4	0 · 29
Additional treatment			
No	0 · 00	Ref.	
Yes	0 · 25	-3 · 9 to 4 · 4	0 · 91
Offense type			
Hands-off (only §184b StGB)	0 · 00	Ref.	
Hands-on (at least one conviction §176 ff StGB)	1 · 9	-2 · 5 to 6 · 2	0 · 40
Static recidivism risk (baseline)	-0 · 94	-2 · 7 to 0 · 84	0 · 30
Timepoint * treatment			
Module 1 (post) * Intervention	1 · 7	-3 · 1 to 6 · 5	0 · 49
Module 2 (before) * Intervention	-1 · 8	-6 · 4 to 2 · 7	0 · 43
Module 2 (post) * Intervention	-1 · 6	-6 · 6 to 3 · 3	0 · 52
Module 3 (before) * Intervention	-0 · 86	-6 · 3 to 4 · 6	0 · 76
Module 3 (post) * Intervention	-4 · 5	-10 to 1 · 1	0 · 11
Module 4 (before) * Intervention	0 · 10	-5 · 9 to 6 · 1	0 · 97
Module 4 (post) * Intervention	-4 · 3	-10 to 1 · 6	0 · 15
Module 5 (before) * Intervention	-2 · 4	-8 · 8 to 4 · 1	0 · 47
Module 5 (post) * Intervention	-2 · 9	-10 to 4 · 4	0 · 43
Module 6 (before) * Intervention	-1 · 9	-9 · 2 to 5 · 3	0 · 60
Module 6 (post) * Intervention	-0 · 15	-7 · 6 to 7 · 3	0 · 97
Timepoint			
Module 1 (post)	0 · 00	Ref.	
Module 2 (before)	-0 · 22	-3 · 1 to 2 · 6	0 · 88
Module 2 (post)	0 · 72	-2 · 6 to 4 · 0	0 · 67
Module 3 (before)	-0 · 02	-3 · 9 to 3 · 9	>0 · 99
Module 3 (post)	1 · 8	-2 · 0 to 5 · 5	0 · 35
Module 4 (before)	-2 · 4	-6 · 4 to 1 · 6	0 · 24
Module 4 (post)	1 · 1	-2 · 8 to 5 · 0	0 · 57
Module 5 (before)	0 · 90	-3 · 4 to 5 · 2	0 · 68
Module 5 (post)	-0 · 31	-5 · 2 to 4 · 6	0 · 90
Module 6 (before)	1 · 4	-3 · 7 to 6 · 5	0 · 58
Module 6 (post)	3 · 1	-2 · 2 to 8 · 5	0 · 24
Baseline value (WHO-5)	0 · 68	0 · 60 to 0 · 77	<0 · 0001

Abbreviations: CI = Confidence Interval, CI = Confidence interval, StGB = German penalty law.

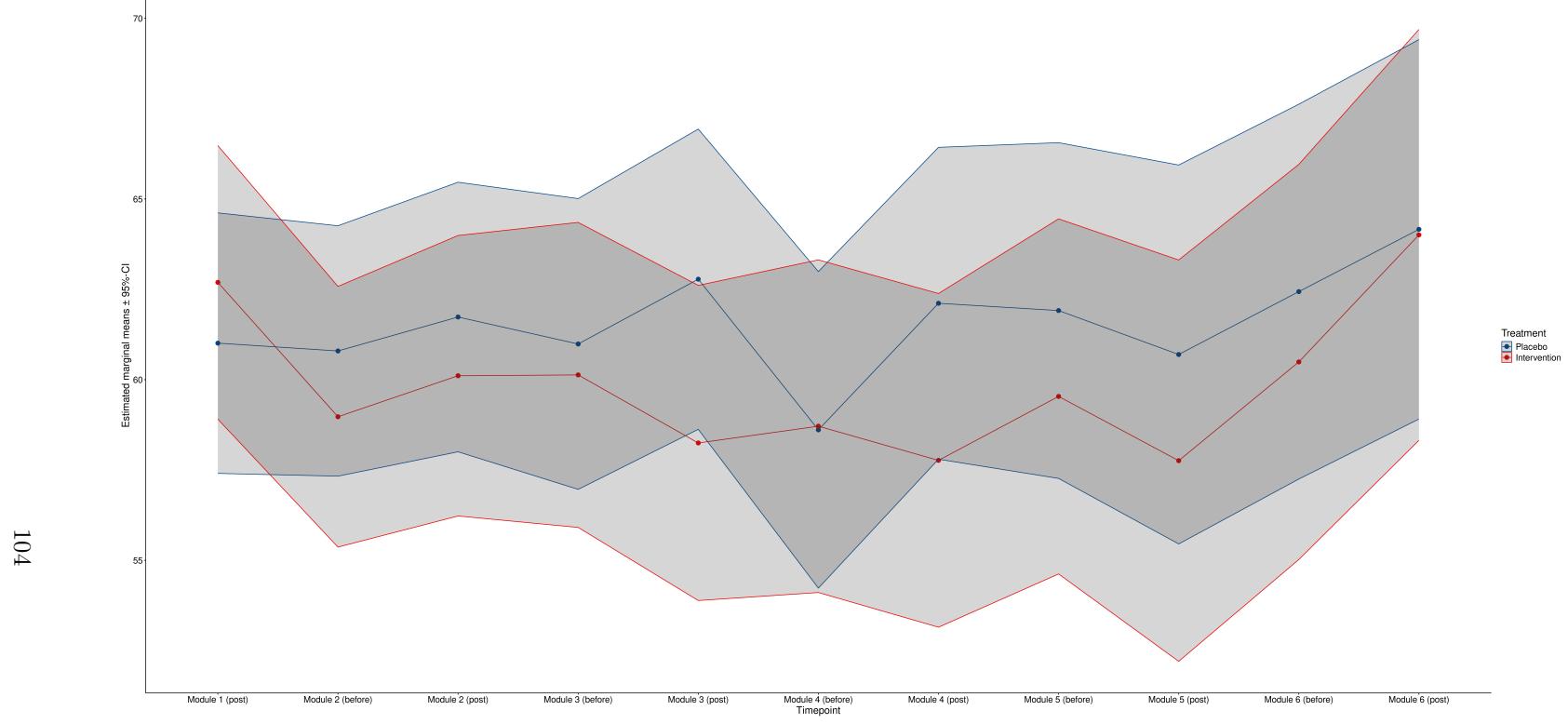


Figure 50: Estimated marginal means of the MMRM for the WHO-5 Well-Being Index (WHO-5 total score).

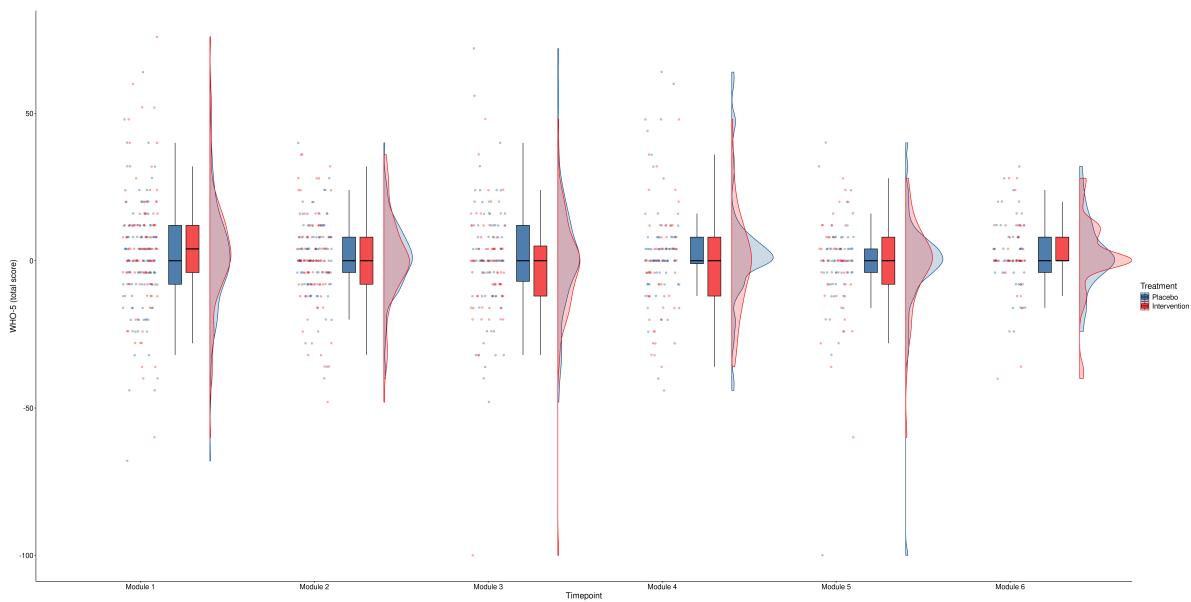


Figure 51: Rainplot for the WHO-5 Well-Being Index (WHO-5 difference from baseline).

Table 38: Pairwise comparison of the WHO-5 Well-Being Index (WHO-5, total score, pre-post module) for the intervention arm.

Module	pre N = 106	post N = 106	p	q-value
Module 1			0 · 054	0 · 27
Mean (SD)	58 (23)	62 (24)		
Median (Q1, Q3)	64 (40, 76)	68 (40, 80)		
Min, Max	8, 100	12, 100		
Module 2			0 · 52	>0 · 99
Mean (SD)	63 (23)	62 (23)		
Median (Q1, Q3)	70 (44, 80)	64 (44, 80)		
Min, Max	12, 100	0, 100		
Missing	22	22		
Module 3			0 · 041	0 · 25
Mean (SD)	63 (24)	60 (23)		
Median (Q1, Q3)	68 (44, 80)	62 (42, 80)		
Min, Max	0, 100	0, 100		
Missing	38	38		
Module 4			0 · 77	>0 · 99
Mean (SD)	60 (24)	58 (27)		
Median (Q1, Q3)	62 (38, 80)	64 (36, 80)		
Min, Max	0, 100	0, 100		
Missing	46	46		
Module 5			0 · 48	>0 · 99
Mean (SD)	59 (27)	60 (28)		
Median (Q1, Q3)	62 (36, 80)	66 (32, 80)		
Min, Max	0, 100	0, 100		
Missing	56	56		
Module 6			0 · 13	0 · 50
Mean (SD)	61 (29)	66 (26)		
Median (Q1, Q3)	70 (32, 80)	72 (52, 84)		
Min, Max	0, 100	12, 100		
Missing	64	64		

Abbreviation: Q1 = 25th percentile, Q3 = 75th percentile, p_{adj} = Holm-Bonferroni adjusted p.

Table 39: Pairwise comparison of the WHO-5 Well-Being Index (WHO-5, total score, pre-post module) for the placebo arm.

Module	pre N = 106	post N = 106	p	q-value
Module 1			0 · 054	0 · 27
Mean (SD)	58 (23)	62 (24)		
Median (Q1, Q3)	64 (40, 76)	68 (40, 80)		
Min, Max	8, 100	12, 100		
Module 2			0 · 52	>0 · 99
Mean (SD)	63 (23)	62 (23)		
Median (Q1, Q3)	70 (44, 80)	64 (44, 80)		
Min, Max	12, 100	0, 100		
Missing	22	22		
Module 3			0 · 041	0 · 25
Mean (SD)	63 (24)	60 (23)		
Median (Q1, Q3)	68 (44, 80)	62 (42, 80)		
Min, Max	0, 100	0, 100		
Missing	38	38		
Module 4			0 · 77	>0 · 99
Mean (SD)	60 (24)	58 (27)		
Median (Q1, Q3)	62 (38, 80)	64 (36, 80)		
Min, Max	0, 100	0, 100		
Missing	46	46		
Module 5			0 · 48	>0 · 99
Mean (SD)	59 (27)	60 (28)		
Median (Q1, Q3)	62 (36, 80)	66 (32, 80)		
Min, Max	0, 100	0, 100		
Missing	56	56		
Module 6			0 · 13	0 · 50
Mean (SD)	61 (29)	66 (26)		
Median (Q1, Q3)	70 (32, 80)	72 (52, 84)		
Min, Max	0, 100	12, 100		
Missing	64	64		

Abbreviation: Q1 = 25th percentile, Q3 = 75th percentile, p_{adj} = Holm-Bonferroni adjusted p.

Table 40: Pairwise comparisons of the WHO-5 Well-Being Index (WHO-5, total score, difference score).

Characteristic	Placebo N = 113	Intervention N = 108	p	q-value
Module 1			0 · 31	>0 · 99
Mean (SD)	1 (18)	4 (20)		
Median (Q1, Q3)	0 (-8, 12)	4 (-4, 12)		
Min, Max	-68, 64	-60, 76		
Module 2			0 · 86	>0 · 99
Mean (SD)	1 (12)	0 (16)		
Median (Q1, Q3)	0 (-4, 8)	0 (-8, 8)		
Min, Max	-40, 40	-48, 36		
Missing	19	17		
Module 3			0 · 26	>0 · 99
Mean (SD)	1 (19)	-2 (19)		
Median (Q1, Q3)	0 (-8, 12)	0 (-12, 6)		
Min, Max	-48, 72	-100, 48		
Missing	35	36		
Module 4			0 · 15	0 · 89
Mean (SD)	4 (18)	-1 (17)		
Median (Q1, Q3)	0 (-2, 8)	0 (-12, 8)		
Min, Max	-44, 64	-36, 48		
Missing	41	46		
Module 5			0 · 95	>0 · 99
Mean (SD)	-1 (18)	-2 (15)		
Median (Q1, Q3)	0 (-4, 4)	0 (-8, 8)		
Min, Max	-100, 40	-60, 28		
Missing	54	55		
Module 6			0 · 63	>0 · 99
Mean (SD)	2 (12)	2 (13)		
Median (Q1, Q3)	0 (-4, 8)	0 (0, 8)		
Min, Max	-24, 32	-40, 28		
Missing	60	66		

Abbreviation: Q1 = 25th percentile, Q3 = 75th percentile, p_{adj} = Holm-Bonferroni adjusted p.

Table 41: Mixed model for repeated measures (MMRM) for the Working Alliance Inventory-Short Revised (WAI-SR COACH).

Variable	Beta	95% CI	p-value
(Intercept)	15	11 to 20	<0 · 0001
Type of supervision			
community supervision	0 · 00	Ref.	
post-release supervision	1 · 2	-1 · 6 to 4 · 0	0 · 39
Additional treatment			
No	0 · 00	Ref.	
Yes	1 · 1	-1 · 1 to 3 · 3	0 · 34
Offense type			
Hands-off (only §184b StGB)	0 · 00	Ref.	
Hands-on (at least one conviction §176 ff StGB)	-1 · 2	-3 · 4 to 1 · 1	0 · 31
Static recidivism risk (baseline)	-0 · 59	-1 · 5 to 0 · 35	0 · 22
Timepoint * treatment			
Module 1 * Intervention	2 · 4	0 · 30 to 4 · 5	0 · 025
Module 2 * Intervention	3 · 8	1 · 2 to 6 · 5	0 · 0044
Module 3 * Intervention	3 · 4	0 · 96 to 5 · 9	0 · 0067
Module 4 * Intervention	2 · 1	-0 · 63 to 4 · 9	0 · 13
Module 5 * Intervention	4 · 3	1 · 3 to 7 · 3	0 · 0051
Timepoint			
Module 1	0 · 00	Ref.	
Module 2	-0 · 96	-2 · 3 to 0 · 35	0 · 15
Module 3	0 · 39	-0 · 92 to 1 · 7	0 · 56
Module 4	0 · 46	-1 · 1 to 2 · 0	0 · 55
Module 5	1 · 1	-0 · 68 to 2 · 9	0 · 22
Baseline value (WAI-SR COACH (total score))	0 · 67	0 · 58 to 0 · 77	<0 · 0001

Abbreviations: CI = Confidence Interval, CI = Confidence interval, StGB = German penalty law.

9.2 WAI-SR: Online-Coach

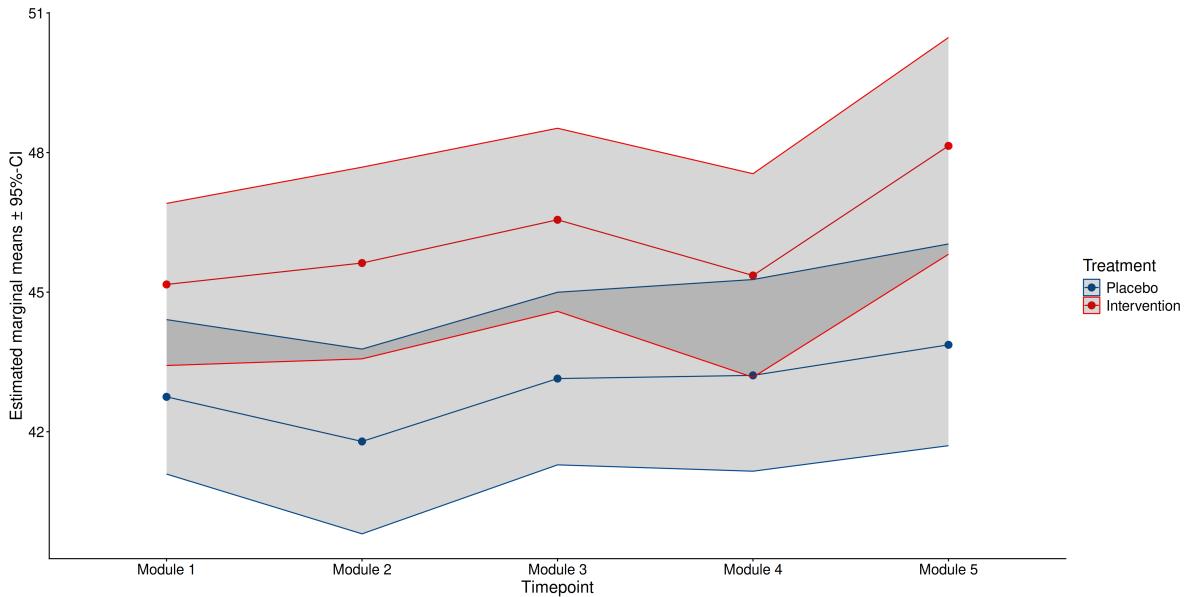


Figure 52: Estimated marginal means from the MMRM of the Working Alliance Inventory-Short Revised (WAI-SR COACH).

Table 42: Pairwise comparison of the Working Alliance Inventory-Short Revised (WAI-SR COACH total score) for the intervention arm.

Module	pre N = 97	post N = 97	p	q-value
Module 1 (post)			0 · 0020	0 · 010
Mean (SD)	42 (11)	44 (9)		
Median (Q1, Q3)	43 (34, 49)	47 (39, 51)		
Min, Max	14, 60	22, 60		
Module 2 (post)			0 · 21	0 · 62
Mean (SD)	45 (9)	46 (10)		
Median (Q1, Q3)	47 (41, 52)	48 (39, 54)		
Min, Max	23, 60	12, 60		
Missing	15	15		
Module 3 (post)			0 · 61	0 · 62
Mean (SD)	47 (9)	48 (9)		
Median (Q1, Q3)	48 (40, 55)	48 (42, 55)		
Min, Max	27, 60	24, 60		
Missing	29	29		
Module 4 (post)			0 · 24	0 · 62
Mean (SD)	47 (9)	46 (10)		
Median (Q1, Q3)	48 (42, 55)	47 (39, 56)		
Min, Max	24, 60	23, 60		
Missing	39	39		
Module 5 (post)			0 · 0040	0 · 016
Mean (SD)	46 (11)	49 (9)		
Median (Q1, Q3)	47 (37, 57)	49 (41, 57)		
Min, Max	23, 60	23, 60		
Missing	50	50		

Abbreviation: Q1 = 25th percentile, Q3 = 75th percentile, p_{adj} = Holm-Bonferroni adjusted p.

Table 43: Pairwise comparison of the Working Alliance Inventory-Short Revised (WAI-SR COACH) for the placebo arm.

Module	pre N = 98	post N = 98	p	q-value
Module 1 (post)			0 · 59	>0 · 99
Mean (SD)	42 (11)	42 (11)		
Median (Q1, Q3)	43 (34, 51)	44 (34, 50)		
Min, Max	12, 60	13, 60		
Module 2 (post)			0 · 52	>0 · 99
Mean (SD)	42 (12)	41 (13)		
Median (Q1, Q3)	44 (34, 50)	42 (30, 53)		
Min, Max	13, 60	12, 60		
Missing	16	16		
Module 3 (post)			0 · 28	>0 · 99
Mean (SD)	42 (14)	43 (13)		
Median (Q1, Q3)	44 (30, 53)	46 (35, 53)		
Min, Max	12, 60	12, 60		
Missing	24	24		
Module 4 (post)			0 · 56	>0 · 99
Mean (SD)	43 (13)	43 (13)		
Median (Q1, Q3)	46 (35, 53)	47 (34, 54)		
Min, Max	12, 60	12, 60		
Missing	32	32		
Module 5 (post)			0 · 40	>0 · 99
Mean (SD)	44 (13)	44 (13)		
Median (Q1, Q3)	48 (34, 54)	48 (34, 55)		
Min, Max	12, 60	14, 60		
Missing	43	43		

Abbreviation: Q1 = 25th percentile, Q3 = 75th percentile, p_{adj} = Holm-Bonferroni adjusted p.

Table 44: Pairwise comparisons of the pre-post differences of the Working Alliance Inventory-Short Revised (WAI-SR COACH).

Characteristic	Placebo N = 98	Intervention N = 97	p	q-value
Module 1 (post)			0 · 025	0 · 13
Median (Q1 – Q3)	1 (-3 – 5)	3 (-2 – 8)		
Module 2 (post)			0 · 059	0 · 18
Median (Q1 – Q3)	0 (-4 – 3)	1 (-2 – 4)		
Missing	16	15		
Module 3 (post)			0 · 39	0 · 39
Median (Q1 – Q3)	0 · 0 (-2 · 0 – 4 · 0)	0 · 0 (-3 · 0 – 2 · 0)		
Missing	24	29		
Module 4 (post)			0 · 19	0 · 39
Median (Q1 – Q3)	0 · 0 (-2 · 0 – 4 · 0)	-1 · 0 (-5 · 0 – 2 · 0)		
Missing	32	39		
Module 5 (post)			0 · 036	0 · 14
Median (Q1 – Q3)	0 · 0 (-2 · 0 – 4 · 0)	2 · 0 (0 · 0 – 6 · 0)		
Missing	43	50		

Abbreviation: Q1 = 25th percentile, Q3 = 75th percentile, SD = Standard deviance, p_{adj} = Holm-Bonferroni adjusted p.

Table 45: Mixed model for repeated measures (MMRM) for the Working Alliance Inventory-Short Revised (WAI-SR SUPERVISOR).

Variable	Beta	95% CI	p-value
(Intercept)	15	10 to 19	<0 · 0001
Type of supervision			
community supervision	0 · 00	Ref.	
post-release supervision	1 · 6	-0 · 61 to 3 · 9	0 · 15
Additional treatment			
No	0 · 00	Ref.	
Yes	1 · 0	-0 · 75 to 2 · 8	0 · 26
Offense type			
Hands-off (only §184b StGB)	0 · 00	Ref.	
Hands-on (at least one conviction §176 ff StGB)	-1 · 2	-3 · 0 to 0 · 62	0 · 20
Static recidivism risk (baseline)	-0 · 66	-1 · 4 to 0 · 09	0 · 086
Timepoint * treatment			
Module 1 * Intervention	2 · 3	0 · 64 to 4 · 1	0 · 0073
Module 2 * Intervention	3 · 0	0 · 87 to 5 · 2	0 · 0061
Module 3 * Intervention	3 · 2	1 · 2 to 5 · 3	0 · 0022
Module 4 * Intervention	2 · 1	-0 · 16 to 4 · 3	0 · 068
Module 5 * Intervention	3 · 6	1 · 2 to 6 · 0	0 · 0038
Timepoint			
Module 1	0 · 00	Ref.	
Module 2	-0 · 14	-1 · 3 to 0 · 96	0 · 80
Module 3	0 · 72	-0 · 36 to 1 · 8	0 · 19
Module 4	1 · 1	-0 · 04 to 2 · 3	0 · 059
Module 5	1 · 8	0 · 30 to 3 · 3	0 · 019
Baseline value (WAI-SR SUPERVISOR (total score))	0 · 70	0 · 62 to 0 · 79	<0 · 0001

Abbreviations: CI = Confidence Interval, CI = Confidence interval, StGB = German penalty law.

9.3 WAI-SR: Community supervisor

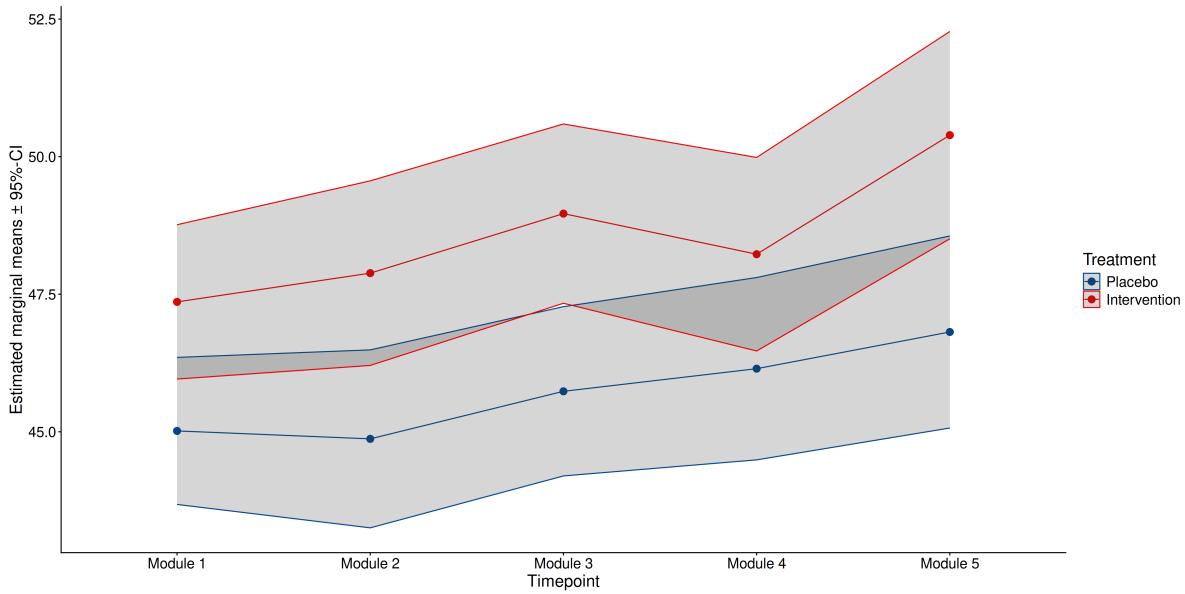


Figure 53: Estimated marginal means from the MMRM of Working Alliance Inventory-Short Revised (WAI-SR SUPERVISOR).

Table 46: Pairwise comparison of the Working Alliance Inventory-Short Revised (WAI-SR SUPERVISOR) for the intervention arm.

Module	pre N = 97	post N = 97	p	q-value
Module 1 (post)			<0 · 0001	<0 · 0001
Mean (SD)	43 (9)	46 (8)		
Median (Q1, Q3)	44 (36, 49)	47 (42, 53)		
Min, Max	22, 60	24, 60		
Module 2 (post)			0 · 14	0 · 35
Mean (SD)	47 (8)	47 (9)		
Median (Q1, Q3)	48 (42, 53)	48 (42, 54)		
Min, Max	24, 60	12, 60		
Missing	15	15		
Module 3 (post)			0 · 12	0 · 35
Mean (SD)	49 (8)	50 (8)		
Median (Q1, Q3)	51 (43, 55)	51 (46, 57)		
Min, Max	27, 60	27, 60		
Missing	29	29		
Module 4 (post)			0 · 44	0 · 44
Mean (SD)	49 (8)	49 (9)		
Median (Q1, Q3)	51 (46, 57)	49 (41, 57)		
Min, Max	27, 60	31, 60		
Missing	39	39		
Module 5 (post)			0 · 046	0 · 18
Mean (SD)	49 (9)	50 (8)		
Median (Q1, Q3)	49 (41, 57)	52 (46, 57)		
Min, Max	31, 60	33, 60		
Missing	50	50		

Abbreviation: Q1 = 25th percentile, Q3 = 75th percentile, p_{adj} = Holm-Bonferroni adjusted p.

Table 47: Pairwise comparisons of the Working Alliance Inventory-Short Revised (WAI-SR SUPERVISOR) for the placebo arm.

Module	pre N = 98	post N = 98	p	q-value
Module 1 (post)			0 · 32	>0 · 99
Mean (SD)	44 (10)	45 (10)		
Median (Q1, Q3)	45 (38, 52)	47 (37, 51)		
Min, Max	16, 60	19, 60		
Module 2 (post)			0 · 72	>0 · 99
Mean (SD)	44 (10)	44 (11)		
Median (Q1, Q3)	47 (37, 52)	47 (37, 53)		
Min, Max	19, 60	14, 60		
Missing	16	16		
Module 3 (post)			0 · 21	>0 · 99
Mean (SD)	45 (11)	45 (11)		
Median (Q1, Q3)	48 (37, 53)	48 (36, 55)		
Min, Max	14, 60	16, 60		
Missing	24	24		
Module 4 (post)			0 · 45	>0 · 99
Mean (SD)	46 (11)	46 (11)		
Median (Q1, Q3)	48 (38, 54)	49 (39, 54)		
Min, Max	16, 60	16, 60		
Missing	32	32		
Module 5 (post)			0 · 59	>0 · 99
Mean (SD)	47 (10)	47 (10)		
Median (Q1, Q3)	49 (40, 55)	48 (40, 57)		
Min, Max	20, 60	23, 60		
Missing	43	43		

Abbreviation: Q1 = 25th percentile, Q3 = 75th percentile, p_{adj} = Holm-Bonferroni adjusted p.

Table 48: Differences from baseline of the Working Alliance Inventory-Short Revised (WAI-SR SUPERVISOR).

Characteristic	Placebo N = 98	Intervention N = 97	p	q-value
Module 1 (post)			0 · 0090	0 · 045
Median (Q1 – Q3)	1 (-3 – 4)	4 (-1 – 6)		
Module 2 (post)			0 · 19	0 · 40
Median (Q1 – Q3)	0 · 0 (-3 · 0 – 2 · 0)	1 · 0 (-1 · 0 – 4 · 0)		
Missing	16	15		
Module 3 (post)			0 · 51	0 · 51
Median (Q1 – Q3)	1 · 0 (-2 · 0 – 3 · 0)	0 · 0 (-2 · 0 – 3 · 0)		
Missing	24	29		
Module 4 (post)			0 · 13	0 · 40
Median (Q1 – Q3)	0 · 0 (-2 · 0 – 3 · 0)	-0 · 5 (-4 · 0 – 2 · 0)		
Missing	32	39		
Module 5 (post)			0 · 061	0 · 24
Median (Q1 – Q3)	0 · 0 (-2 · 0 – 2 · 0)	1 · 0 (0 · 0 – 4 · 0)		
Missing	43	50		

Abbreviation: Q1 = 25th percentile, Q3 = 75th percentile, SD = Standard deviance, p_{adj} = Holm-Bonferroni adjusted p.

10 CONSORT checklist

Table 49: CONSORT 2025 checklist. Adapted from⁷⁴.

Section	Subsection	Number	Description	Reference
Title and abstract	Title and structured abstract	1a 1b	Identification as a randomised trial Structured summary of the trial design, methods, results, and conclusions	p. 2 (Summary) p. 2 (Summary)
Open science	Trial registration	2	Name of trial registry, identifying number (with URL) and date of registration	p. 5 (Section 2.1)
	Protocol and statistical analysis plan	3	Where the trial protocol and statistical analysis plan can be accessed	appendix section 1 (trial protocol); appendix section 11 (SAP)
	Data sharing	4	Where and how the individual de-identified participant data (including data dictionary), statistical code, and any other materials can be accessed	p. 13 (Data sharing)
	Funding and conflicts of interest	5a 5b	Sources of funding and other support (eg, supply of drugs), and role of funders in the design, conduct, analysis, and reporting of the trial Financial and other conflicts of interest of the manuscript authors	Summary; p. 8 (Section 2.6); p. 13 (Acknowledgements) p. 13 (Declaration of interests)
Introduction	Background and rationale	6	Scientific background and rationale	p. 4-5 (Section 1)
	Objectives	7	Specific objectives related to benefits and harms	p. 5 (Section 1)
Methods	Patient and public involvement	8	Details of patient or public involvement in the design, conduct, and reporting of the trial	p. 5 Section 2.1)
	Trial design	9	Description of trial design including type of trial (eg, parallel group, crossover), allocation ratio, and framework (eg, superiority, equivalence, non-inferiority, or exploratory)	p. 5 (Section 2.1)
	Changes to trial protocol	10	Important changes to the trial after it commenced including any outcomes or analyses that were not prespecified, with reason	p. 6 (Section 2.4)
	Trial setting	11	Settings (eg, community, hospital) and locations (eg, countries, sites) where the trial was conducted	p. 5 (Section 2.1)
	Eligibility criteria	12a 12b	Eligibility criteria for participants If applicable, eligibility criteria for sites and for individuals delivering the interventions (eg, surgeons, physiotherapists)	p. 5 (Section 2.1) not applicable
	Intervention and comparator	13	Intervention and comparator with sufficient details to allow replication. If relevant, where additional materials	p. 5 (Section 2.3)

11 Statistical analysis plan

Statistical Analysis Plan

Evaluation of the effectiveness of an online intervention for discharged childabuser

@myTabu

Studienprotokoll

EudraCT-Nr.	—
Clinicaltrials.gov-Nr./DRKS - Nr.	DRKS00021256
Interne Protokoll-ID-Nr.	—
NCT-Nr.	—
SAP Version	0.1/13.09.2024
SAP-Revisionen	—
Entwicklungsphase	—
Sponsor	Universitätsmedizin Göttingen (UMG) Vertreten durch das Studienzentrum der UMG Robert-Koch-Str. 40 37075 Göttingen, Deutschland
Coordinating Investigator	Dr. Peter Fromberger Klinik für Psychiatrie und Psychotherapie Forensische Psychiatrie Universitätsmedizin Göttingen Rosdorfer Weg 70 37081 Göttingen, Deutschland

Approval of the Statistical Analysis Plan

@myTabu

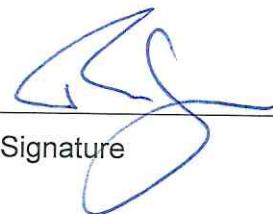
Protocol version No.:

Koordinierender Prüfarzt
Dr. Peter Fromberger

18.11.24

Date

Signature



Biostatistiker
Dr. Andreas Leha

Date

Signature

Table of Contents

Approval of the Statistical Analysis Plan	2
Table of Contents.....	3
1 Introduction	5
2 Study Methods	7
3 Statistical Principles	9
4 Trial Population.....	9
5 Analysis	10
6 References.....	16

Abbreviations

BIS-15	Barratt Impulsiveness Scale – 15
BMS	Bumby Molest Scale
CI	Confidence intervals
CMC	Checklist of Behavioral Misconduct
CS	community supervision
CsaU	CS as Usual
CUSI	Coping Using Sex Inventory
CTE	Checklist of Treatment Effectiveness
CVTRQ	Corrections Victoria Treatment Readiness Questionnaire
CAs	Child abusers
DERS	Difficulties in Emotion Regulation Scale - Sub-scale Impulsivity
EKK-R	Emotional Congruence with Children – Revised
ESIQ	The Explicit Sexual Interest Questionnaire
F-SozU K-7	Social Support Questionnaire
HBI-19	Hypersexual Behavior Inventory – 19
IoD	Index of Desistance
IfSuFP	Institute for Sex Research and Forensic Psychiatry
ITT	intention-to-treat
NARQ	Negative Affect Repair Questionnaire
OQMPr	Optimized Questionnaire for the Measurement of Psychological Reactance
RCQ	Readiness to Change Questionnaire
SO	supervision officer
SOI-R	Outlet Inventory – Revised
SPSI-R	Social Problem-Solving Inventory Revised
SSIC	Specific self-efficacy for modifying Sexual Interest in Children
UCLA	University of California Los Angeles Loneliness Scale
UMG	University Medical Center Göttingen

1 Introduction

This document has been written based on the study protocol. A design paper has been published in *Frontiers in Psychiatry*; DOI: 10.3389/fpsy.2020.575464

1.1 Background and Rationale

Treatment programs for child abusers are able to significantly reduce the risk to re-offend. In spite of some established treatment projects in primary prevention, treatment offers for child abusers conditionally released from prison are scarce – especially in rural areas. This is alarming, since most re-offences take place shortly after release from prison.

In addition to psychiatric or psychological treatment efforts, community supervision is effective in reducing re-offenses but only with high financial effort. Online interventions provide a cost-efficient additional treatment option and can be applied smoothly in rural areas.

1.2 Objectives and Endpoints

The current study aims to evaluate the effectiveness of (1) @myTabu in reducing dynamic risk factors compared to a placebo condition (2) the individual online modules within @myTabu with regard to reducing psychologically meaningful dynamic risk factors. It investigates the primary hypothesis that child abusers (CAs) who take part in @myTabu show lower values on dynamic risk factors (Index of Desistance, IoD) from baseline to inter-, and post-treatment, compared to an online-based placebo. The second hypothesis assumes that each individual online module within @myTabu is able to specifically address and reduce the dynamic risk factors which correspond to the psychotherapeutic content of the module. See table 1 for an overview of hypotheses.

Table 1 Objectives and related endpoints

	Objective	Endpoint
Primary	Evaluate the effectiveness of @myTabu in reducing dynamic risk factors compared to a placebo condition	Dynamic risk factors (Index of Desistance, IoD) from baseline to inter-, and post-treatment, compared to an online-based placebo
	Evaluate the effectiveness of @myTabu in reducing the number of officially recorded re-offenses compared to a placebo condition	Number of officially recorded re-offenses five years after last patient out
Secondary	Evaluate the Motivation Module individually with regard to its targeted risk factors low therapy motivation and low readiness for change	Differences in the Corrections Victoria Treatment Readiness Questionnaire, CVTRQ, pre and post Motivation Module
	Evaluate the Motivation Module individually with regard to its targeted risk factors low therapy motivation and missing readiness for change	Differences in the Readiness to Change Questionnaire, RCQ, pre and post Motivation Module
	Evaluate the Community-Supervision Module individually with regard to its	Differences in the Optimized Questionnaire for the Measurement of

	Objective	Endpoint
	targeted risk factors negative social influences and resistance to rules and supervision	Psychological Reactance, OQMPR, pre and post Community-Supervision Module
	Evaluate the Community-Supervision Module individually with regard to its targeted risk factors negative social influences and resistance to rules and supervision	Differences in the Social Support Questionnaire, F-SozU K-7, pre and post Community-Supervision Module
	Evaluate the Community-Supervision Module individually with regard to its targeted risk factors negative social influences and resistance to rules and supervision	Differences in the University of California Los Angeles Loneliness Scale, UCLA, pre and post Community-Supervision Module
	Evaluate the Emotion Management Module individually with regard to its targeted risk factors lifestyle impulsiveness and dysfunctional coping	Differences in the Difficulties in Emotion Regulation Scale - Sub-scale Impulsivity, DERS, pre and post the Emotion Management Module
	Evaluate the Emotion Management Module individually with regard to its targeted risk factors lifestyle impulsiveness and dysfunctional coping	Differences in the Negative Affect Repair Questionnaire, NARQ, pre and post the Emotion Management Module
	Evaluate the Emotion Management Module individually with regard to its targeted risk factors lifestyle impulsiveness and dysfunctional coping	Differences in the Barratt Impulsiveness Scale – 15, BIS-15, pre and post the Emotion Management Module
	Evaluate the Emotion Management Module individually with regard to its targeted risk factors lifestyle impulsiveness and dysfunctional coping	Differences in the Coping Using Sex Inventory, CUSI, pre and post the Emotion Management Module
	Evaluate the Problem Solving Module individually with regard to its targeted risk factor poor (social) problem solving	Differences in the Social Problem-Solving Inventory Revised, SPSI-R, pre and post the Problem Solving Module
	Evaluate the Offense-supportive Attitudes Module individually with regard to its targeted risk factor offense-supportive attitudes	Differences in the Bumby Molest Scale, BMS, pre and post the Offense-supportive Attitudes Module
	Evaluate the Sexuality Module individually with regard to its targeted risk factors sexual preoccupation, lack of emotionally intimate relationships with adults and emotional congruence with children	Differences in the Hypersexual Behavior Inventory – 19, HBI-19, pre and post the Sexuality Module
	Evaluate the Sexuality Module individually with regard to its targeted risk factors sexual preoccupation, lack of emotionally intimate relationships	Differences in the Specific self-efficacy for modifying Sexual Interest in Children, SSIC, pre and post the Sexuality Module

	Objective	Endpoint
	with adults and emotional congruence with children	
	Evaluate the Sexuality Module individually with regard to its targeted risk factors sexual preoccupation, lack of emotionally intimate relationships with adults and emotional congruence with children	Differences in the Outlet Inventory – Revised, SOI-R, pre and post the Sexuality Module
	Evaluate the Sexuality Module individually with regard to its targeted risk factors sexual preoccupation, lack of emotionally intimate relationships with adults and emotional congruence with children	Differences in Questionnaire on Emotional Congruence with Children – Revised, EKK-R, pre and post the Sexuality Module
	Evaluate the Sexuality Module individually with regard to its targeted risk factors sexual preoccupation, lack of emotionally intimate relationships with adults and emotional congruence with children	Differences in The Explicit Sexual Interest Questionnaire, ESIQ, pre and post the Sexuality Module
Safety	Assessment of safety	Number of violations of directives and re-offences

1.3 Primary objectives and endpoints

Officially recorded re-offenses in the experimental condition compared to the placebo condition are the first primary outcome. Officially recorded re-offenses are obtained five years after last patient out.

To evaluate the effectiveness of @myTabu in reducing dynamic risk factors compared to a placebo condition the Index of Desistance, IoD, is analyzed. The classical full IoD will be used. The IoD will be used in its reduced form without considering the CMC scale in an exploratory analysis.

1.4 Secondary objective and endpoints

Secondary endpoints are listed in Table 1.

1.5 Safety endpoints

Safety endpoint include violations of directives and re-offences.

2 Study Methods

2.1 Trial design

This trial is a concealed, prospective, parallel group, placebo-control trial with 1:1 randomization. Due to security reasons, blinding of supervision officers (SOs) and supporting therapists is not

possible. Nevertheless, subjects, outcome adjudicators and data analysts (observer) are blinded until the end of the data analysis.

2.2 Randomization

@myTabu comprises two parallel groups (online intervention group and online Placebo group). Subjects are randomly allocated to treatment using an allocation ratio of 1:1. The randomization lists are centrally generated using a computerized system stratified by the offense type, type of community supervision (CS), external treatment in addition to CS as Usual (CSaU), and risk to re-offend at baseline. At screening, each subject receives the next consecutive screening number. At randomization, each subject eligible for study participation receives the next consecutive randomization number according to his stratum from a block of randomization numbers (block size: 4). The randomization list is kept in safe and confidential custody at the "Stabstelle Klinische Studien" of the UMG.

2.3 Sample Size

@myTabu assumes that CAs who take part in @myTabu show significantly lower values on dynamic risk factors (IoD) from baseline to inter-, and post-treatment, compared to an online-based placebo condition. Within sex offender programs, higher levels of motivation and better treatment engagement were associated with decreased attrition. @myTabu addresses this in a direct manner (see sub-project 1). In addition, compliance of CAs in taking part of scientific studies is a critical point. Based on the experiences of the consortium, who have successfully finished numerous scientific studies with child abusers, @myTabu assumes a conservative rate of 50% of participants taking voluntarily part in the clinical trial. @myTabu tries to enhance the motivation to take part in the clinical trial by providing an adequate allowance. It is expected to see 1165 subjects eligible for participation during the recruitment period of which 50% (582) are expected to participate in the study. Taking an expected dropout of 26.4% into account, @myTabu expect to see 428 subjects completing the online intervention. With 214 subjects in both groups (treatment and placebo) it is possible to detect small effects ($d = .27$) with a power of 80% at a significance level $\alpha = 5\%$ (calculated using nQuery Advisor 7.0).

2.4 Framework

All endpoints are tested for superiority of the intervention over placebo control.

2.5 Stopping Guidance

The individual subject is excluded from the clinical trial if the responsible SO or the supporting therapist reports clear evidence for sexual recidivism or concrete preparations for any sexual recidivism. Stopping rules for participating centres are not applicable, since the UMG serves as the only participating centre.

2.6 Timing of the final analysis

The final analysis will take place when all outcomes have been collected and the database is cleaned and locked.

2.7 Timing of Outcome Assessments

Officially recorded re-offenses are obtained five years after last patient out.

The IoD is collected at the beginning of each module and at the end of the last module.

The secondary endpoints are collected at the beginning and at the end of the corresponding module (see Table 1).

2.8 Methods against Bias

@myTabu tries to minimize a potential volunteer-bias by providing an adequate allowance.

3 Statistical Principles

3.1 Confidence intervals and p-values

All tests will be performed two-sided with 5% significance level. Confidence intervals (CI) will be reported with 95% confidence level.

3.2 Adherence and protocol deviations

Feedback to subjects is automatized, elsewise manualized. Regular meetings between therapists and an experienced supervisor provided by IfSuFP will promote treatment consistency and regular supervision. Sessions are recorded and assessed.

3.3 Analysis populations

For all analyses on the primary endpoints the intention-to-treat (ITT) population will be used. The ITT population consists of all participants who have completed at least one module.

Analyses on the secondary endpoints will be carried out on the specific per-protocol (PP) population. The PP population for a specific module consists of all participants who have completed this module.

4 Trial Population

4.1 Screening Data

Screening data will be reported and described within a CONSORT flowchart.

4.2 Eligibility

All sexual offenders under CS (§§56, 57, 68 StGB), which are sentenced for at least one case of child abuse (§§176 StGB) or child pornography (§184b StGB) in Germany (besides the federal states Hamburg, Bremen and Saarland), are eligible for the trial. There exist no exclusion criteria regarding gender. Only persons who are older than 18 years are included. Exclusion criteria are (1) a probation period shorter than 6 months at enrollment, (2) no access to a PC, tablet or smart phone, (3) a severe acute psychiatric disorder (e.g., acute psychosis), (4) a severe cerebro-organic disorder, (5) a severe cognitive impairment, (6) withdrawal of the informed consent, and (7) no written informed consent. Due to a possible volunteer-effect, the representativeness of the sample for the entire population of CAs cannot be ensured.

4.3 Recruitment

Recruitment numbers will be summarized within a CONSORT flow diagram.

4.4 Drop-out/follow-up

Reasons for study drop-out will be documented using a CONSORT flowchart.

4.5 Baseline patient characteristics

All baseline patient characteristics will be summarized using descriptive statistics (e.g. mean, standard deviation, median, inter-quartile range, minimum and maximum value for continuous variables and frequencies (percentages) for categorical variables) and appropriate graphical methods (e.g. boxplots, bar plots) depending on the data type.

5 Analysis

5.1 Outcome definitions

A full list of outcomes and their timing is described in Section 2.7.

Officially recorded re-offenses Officially recorded re-offenses in the experimental condition compared to the placebo condition are the first primary outcome. Officially recorded re-offenses are obtained 5 years after last patient out.

Index of Desistance (IoD) The IoD represents the individual risk to re-offend assessed through self-report. This composite measurement reflects improvements with regard to meaningful psychological risk factors, offense behaviors as well as offense-related behaviors that have not been registered or convicted yet. The IoD comprises three self-report instruments, which were developed and have been evaluated by members of the @myTabu-consortium (Nitsche et al., 2022): (1) The Acute-2007-SR is an online and self-report adaption of the Acute-2007 (Hanson & Harris 2007) risk-assessment tool in order to assess treatment induced short-term changes in recidivism risk factors. (2) The Checklist of Treatment Effectiveness (CTE) is developed to assess the differential treatment-induced changes in every individual treatment module. Thus, the CTE assesses the constructs that are targeted by the six modules of the intervention. (3) The Checklist

of Behavioral Misconduct (CMC) measures offense-related behavior as well as low- and high-level new offenses. Offense-related behavior includes behaviors which are relevant for individuals previously convicted of sexual offenses prior to a new offense (recidivistic behavior, e.g., if a person convicted of sexual abuse of minors visits a playground or a school). The classical IoD will provide the individual risk level to re-offend by summing up the scores of the Acute-2007-SR, CTE, and CMC. The reduced IoD is the sum of Acute-2007-SR and CTE.

Corrections Victoria Treatment Readiness Questionnaire (CVTRQ) The Corrections Victoria Treatment Readiness Questionnaire [CVTRQ; Casey et al., 2007] is a self-report measure designed to assess treatment readiness in offenders who have been referred to a cognitive skills program. It comprises 20 items and consists of four scales: (1) Attitudes and motivation, (2) emotional reaction, (3) offending beliefs, and (4) efficacy. Each item has to be answered on a five-point Likert-scale. The CVTRQ shows an acceptable convergent validity, discriminant validity as well as predictive validity (Casey et al., 2007). To the best of our knowledge, there is no validated German version of the CVTRQ. @myTabu translated the CVTRQ. The translation was checked and translated back into English. Resulting differences due to adaptation were inspected and discussed in terms of item content.

Readiness to Change Questionnaire (RCQ) The Readiness to Change Questionnaire [RCQ; Rollnick et al., 1992] is a 12-item questionnaire originally designed to identify the stage of change based on the transtheoretical model (Prochaska et al, 1992) reached by individuals who excessively drink alcohol. The German version [RCQ-D; Hannöver et al., 2002] was adapted within the project by changing the questions regarding drinking of alcohol in questions regarding the offenses of ISAC and ICCSEM. Responses are made on a five-point Likert-scale. The RCQ-D as well as the RCQ shows good psychometric properties (Hannöver et al., 2002). Psychometric properties of the version adapted for ISAC and ICCSEM are not available yet.

Optimized Questionnaire for the Measurement of Psychological Reactance (OQMPr) The Questionnaire for the Measurement of Psychological Reactance [QMPr; Merz 1983] is a questionnaire for the assessment of psychological reactance defined as the theory that people resist attempts to constrain either their thoughts or their behaviors. The Optimized Questionnaire for the Measurement of Psychological Reactance [OQMPr; Herzberg 2002] is a German variant of the QMPr. The OQMPr consists of 12 statements on a five-point Likert-scale. It has a test-retest-reliability of $rtt = 0.85$ (Herzberg 2002).

Social Support Questionnaire (F-SozU K-7) The seven-item short version of the Social Support Questionnaire (F-SozU K-7) is an efficient questionnaire to assess perceived social support (Dunkel et al., 2005). Each item comprises a five-point Likert-scale. Internal consistency of the original long version showed Cronbach's α between 0.81 and 0.93 (Sommer & Fydrich, 1989).

University of California Los Angeles Loneliness Scale (UCLA) The UCLA Loneliness Scale (Russell, 1996) consists of 20 items in order to assess subjective feelings of loneliness. The German short version (Bilsky & Hosser, 1998) consists of 12 items on a four-point Likert-scale. A German study with ISAC and ICCSEM demonstrated a high reliability of $\alpha = 0.92$ (Neutze et al, 2012).

Bumby Molest Scale (BMS) The Bumby Molest Scale [BMS; (Bumby, 1996)] is a self-report questionnaire which assesses cognitive distortion in ISAC [German Version: Feelgood et al., 2009]. It consists of 38 items on a four-point Likert-scale. The German version shows good construct validity, internal consistency, and test-retest reliability (Feelgood et al., 2009)..

Social Problem-Solving Inventory Revised (SPSI-R) The Social Problem-Solving Inventory Revised [SPSI-R; D'Zurilla et al., 2002] is a self-report questionnaire for the assessment of the five dimensions in the social problem-solving model: (1) positive problem orientation, (2) negative problem orientation, (3) rational problem solving, (4) impulsivity/carelessness style, and (5) avoidance style. The SPSI-R consists of 52 items on a five-point Likert-scale. It shows good psychometric properties for individuals who have sexually offended (Nezu et al., 2005). The validated German version is a short form consisting of 25 items (Graf, 2003).

Difficulties in Emotion Regulation Scale—Sub-scale Impulsivity (DERS) The Difficulties in Emotion Regulation Scale [DERS; Gratz & Roemer, 2004] is a self-report questionnaire to assess emotion dysregulation. The sub-scale “impulse control difficulties” is used as outcome measure and comprises five items on a five-point Likert-scale. The German version is used, which shows a good internal consistency, construct and predictive validity (Ehring et al., 2008).

Negative Affect Repair Questionnaire (NARQ) The Negative Affect Repair Questionnaire [NARQ; Scherer et al., 2013] is a self-report questionnaire to assess strategies to regulate negative affect in a systematic manner. It consists of 17 items on a five-point Likert-scale and provides a good construct validity. Reliability scores (Cronbach's α) for the three NARQ scales ranged between 0.71 and 0.80 (Scherer et al., 2013).

Barratt Impulsiveness Scale-15 (BIS-15) The Barratt Impulsiveness Scale [BIS; Patton et al., 1995] is a questionnaire developed to assess the personality/behavioral construct of impulsiveness. It consists of 30 items on a four-point Likert-scale. The German short version [BIS-15; Meule et al., 2011] consisting of 15 items is used in the clinical trial. The BIS-15 is an efficient measure of impulsiveness with good internal consistency [$\alpha = 0.81$; Meule et al., 2011].

Coping Using Sex Inventory (CUSI) The Coping Using Sex Inventory [CUSI; Cortoni & Marshall, 2001] is a questionnaire to assess the presence of and the degree to which sex was used to deal with problematic situations. It consists of 16 items on a five-point Likert-scale with a satisfying internal consistency (Cortoni & Marshall, 2001). The translation was checked and translated back into English by the authors. Resulting differences due to adaptation were inspected and discussed in terms of item content. For another German translation, it has been shown that the CUSI is able to assess therapy-induced changes of ISAC (Beier et al., 2015).

Hypersexual Behavior Inventory-19 (HBI-19) Sexual preoccupation is assessed by the Hypersexual Behavior Inventory-19 [HBI-19; Klein et al., 2014]. The HBI-19 is a three-factor measure (coping, control, and consequences) developed to assess hypersexual behavior. The instrument consists of 19 items (e.g., “I use sex to forget sorrows of everyday life.”) answered on a scale from 1 (never) to 5 (very often). The maximum score is 95, with higher scores indicating a higher level of sexual preoccupation. The questionnaire was shown to have good reliability ($\alpha = 0.90$) and validity (Klein et al., 2014).

Questionnaire on Emotional Congruence with Children-Revised (EKK-R) Emotional congruence with children is assessed by the Questionnaire on Emotional Congruence with Children-Revised [EKK-R; Mack & Yundina, 2012] including three factors (special relationship to children, immaturity, and emotional closeness to children). Twenty items are answered on a four-point scale. The questionnaire demonstrated good reliability ($\alpha = 0.80$) and validity (Mack & Yundina, 2012).

Specific self-efficacy for modifying Sexual Interest in Children (SSIC) Specific self-efficacy for modifying sexual interest in children (SSIC) is assessed by the Self-Efficacy for Modifying Sexual Interest in Children Scale [SSIC-Scale; Tozdan et al., 2015]. Six items on the participant's conviction regarding the ability to change their sexual interest in children (e.g., "I can succeed in reducing my sexual interest in children") were answered on a scale from 1 (do not agree at all) to 5 (totally agree). The maximum score is 30, with higher scores indicating a higher level of self-efficacy. The instrument was shown to have good reliability ($\alpha = 0.87$) and validity (Tozdan et al., 2015). Four variables that were shown to be related to the SSIC (Tozdan & Briken, 2015, Tozdan & Briken, 2019) are also assessed: flexibility of sexual interest in children (three items on previous experiences concerning changes in the participant's sexual interest in children), exclusiveness of sexual interest in children (single-item question), motivation to change sexual interest in children (single-item question), and age of onset of sexual interest in children (single-item question).

The Explicit Sexual Interest Questionnaire (ESIQ) The Explicit Sexual Interest Questionnaire [ESIQ; Banse et al., 2010] directly assesses pedophilic interest. It consists of two scales measuring sexual behavior (20 items, e.g., "I enjoyed orally stimulating a man.") and sexual fantasies (20 items, e.g., "I find it attractive to imagine a little boy sexually stimulating me."). All items are answered on a scale from 1 (totally disagree) to 5 (totally agree). The reliability of the instrument ranges between 0.86 and 0.97 (Banse et al., 2010).

Outlet Inventory-Revised (SOI-R) The second measurement for sexual interest in children is the Item 2a of the Sexual Outlet Inventory-Revised [SOI-R; Briken, 2010] which assesses the desire for sexual activity involving children on a visual analog scale from 0 (desire is absent) to 100 (I have to act to satisfy the desire). Higher values on the scale indicate a stronger sexual interest in children (Briken, 2010).

5.2 Analysis methods

All patients who discontinue from the study are identified and the extent of their participation in the study reported. If discontinuation is due to clear evidence for recidivism or concrete preparation for or intention to recidivism (assessed by responsible SO or guiding therapist) the subject receives the maximal value for IoD for the remaining measurements (assuming the worst case of a failed treatment). In all other cases of drop-out, the reason is recorded.

All variables corresponding to the endpoints will be summarized using descriptive statistics (e.g. mean, standard deviation, median, inter-quartile range, minimum and maximum value for continuous variables and frequencies (percentages) for categorical variables) and appropriate graphical methods (e.g. boxplots, bar plots) depending on the data type per time point.

The efficacy of the online intervention will be analyzed by means of Gaussian linear model for repeated measures (MMRM) with treatment group, time (six time-points; after each online intervention module), treatment-by-time interaction, and the randomization strata as factors (offense type, type of CS, external treatment in addition to CSaU), baseline measurements (at clearing up) of the outcome and potentially time span from baseline as covariate and the classical IoD as dependent variable. The error terms are assumed to follow a multivariate normal distribution with unstructured covariance if all parameters of the corresponding model can be estimated and Ante-dependence covariance structure otherwise. Least squares mean changes from baseline are reported for the treatment groups with 95% confidence interval (CI) as well as the difference between the least squares treatment group means with 95% CI and p-value testing

the null hypothesis of no treatment effect. The analysis is performed on the ITT population comprising all patients who have completed at least one module.

The analysis on the ITT population will be complemented by a sensitivity analysis on imputed values. Imputation will be done using multiple reference based imputation followed by ANCOVA models (Wolbers et al., 2022). Differences between the results of the two analyses will be discussed.

The effect of each module will additionally be analyzed using the Wilcoxon signed rank test for paired data if appropriate and the Brunner-Munzel test for paired data otherwise for both the IoD and the scores corresponding to the secondary endpoints comparing the pre- and post-module scores, adjusting the p-values using the Bonferroni-Holm's method to control the family-wise error rate.

The effect of each module on the scores corresponding to the secondary endpoints will be further analyzed by fitting linear models with treatment group and the randomization strata as factors (offense type, type of CS, external treatment in addition to CSaU), pre module scores as covariate and post module scores as dependent variable if appropriate. If necessary, the scores are transformed before this analysis. If the assumptions of a linear model are not fulfilled either an ordinal regression model will be fitted analogously instead or the score is dichotomised in order to fit a logistic regression model analogously.

The proportion of officially recorded re-offenses five years after last patient out is estimated using the Kaplan-Meier method with 95% confidence interval estimated using the beta product confidence procedure for right censored data (Fay, 2013) in both arms and compared between the arms using the beta product confidence procedure for right censored data (Fay, 2014). Depending on the number of events a Cox proportional hazard regression model with treatment group and an additional appropriate set of factors will be fit and resulting hazard ratios will be reported with 95% confidence intervals.

Economic analyses will be based on analytic modeling techniques, including data about recidivism rates based on dynamic risk factors (derived from the clinical trial), tangible benefits (based on existing literature), intangible benefits (based on surveys using the contingent valuation method) and intervention costs (based on the costs incurred per participant in the study). Tangible benefits are estimated on the basis of existing literature, such as costs occurring in the German society in the context of recidivism of child abusers. After reviewing the literature, a selection of relevant studies was made. Several criteria were taken into account in the final selection of relevant studies. Since the health economic evaluation was to be modeled over a 5-year period, only studies that allowed conclusions to be drawn about annual costs were taken into account. Furthermore, only studies were included in which the cost unit (e.g. per capita or per company) was clearly understandable. Furthermore, for studies that were based on the same dataset, only the study thematically most appropriate was included. For the intangible benefits, the contingent valuation method (CVM) was used to estimate the lower and upper bounds for intangible benefits related to reduced recidivism rates. Intangible costs of CA are non-measurable costs that arise around the suffering of the victims. CVM is a survey-based, hypothetical and direct method to value intangible costs, assessed by the willingness to pay. Participants are asked how much should be paid for a program which reduces the relapse rates of sexual child offenders and thus the victims' pain. Survey participants define the maximum money amount accounting for the indicated public good in an unbiased open-answer format. The percentage of prevented recurrence of CA is increased stepwise starting with 5% up to 100%, tightly spaced in the lower ranks, representing more likely potential outcomes of an intervention. The intervention costs include costs of developing the intervention (personnel costs for content and programming), costs associated with conducting the

study and maintaining the intervention (personnel costs for therapeutic support, personnel costs for maintenance and technical support, supervision costs, travel costs, personnel costs for study information and file analysis, compensation for participants) and costs associated with future implementation of the intervention (personnel costs for therapeutic support, personnel costs for maintenance and technical support, technology costs for the provision of appropriate servers and daily backups, supervision costs). We will perform a cost-benefit analysis. Results of the analysis will be presented in terms of a net benefit (NB) estimate and a benefit-cost ratio. The net benefit will be calculated as $NB = \text{tangible benefits} + \text{intangible benefits} - \text{intervention costs}$. The benefit-cost ratio will be calculated as $\text{benefits} / \text{costs}$.

Since many cost parameters occur after the trial period, the cost-benefit analysis of the intervention will be modelled for a 5-year period. Additionally, sensitivity analysis will be conducted to examine how the NB will be influenced by the recidivism rates, intangible benefits and discount rates. Tangible costs of the control group (no intervention) are estimated on the basis of existing literature. The per-person costs found in literature will be multiplied with German prevalence data, to get an estimate for the German society. In order to account for the costs caused by relapses, this data is multiplied by the probability of re-offending (without any treatment). The tangible costs of the intervention group will be estimated taking into account the risk to reoffend in both groups (IoD (IG) / IoD (KG)). The ratio of the risk to reoffend in the intervention and control group will be multiplied by the tangible costs of the control group, assuming that if the risk to reoffend is lower in the intervention group, the costs will be decreased respectively: tangible costs of intervention group = tangible costs of control group * (IoD (IG)/IoD (KG)). The difference in costs between groups represents the tangible benefits: tangible benefits = tangible costs control group – tangible costs intervention group. Intangible benefits will be calculated by analyzing the data from the CVM survey. Within the CVM survey, the question nearest to assessed risk to reoffend (IoD) will be used as the basis of data. Then, the medium of all answers will be calculated and used for the main analysis. Lower and upper bound will be calculated and used for sensitivity analysis. In order to consider the intervention costs from a societal perspective, the costs which occurred per person during the clinical trial will be multiplied by the potential number of convicted child sexual offenders that use the intervention.

5.3 Missing values

In the main analysis missing values will not be imputed. Reference based multiple imputation will be used in a sensitivity analysis for the primary endpoint IoD.

5.4 Additional analyses

Additional analyses will be conducted on an exploratory basis. In particular the analysis of the classical IoD will be complemented by analysis of the reduced IoD.

5.5 Harms – Safety endpoints

Safety data are summarized using descriptive statistics. In particular proportions of occurrences of reoffence and concrete preparation for or intention to recidivism are estimated with 95% confidence intervals for both groups using the beta product confidence procedure for right censored data (Fay, 2013) and differences in the proportions between both groups are estimated

with 95% confidence intervals using the beta product confidence procedure for right censored data (Fay, 2014).

5.6 Statistical Software

The statistical analysis is conducted with R (version 4.2.3 or higher).

6 References

- Nitsche, K., Etzler, S., Tröger, A.-S. & Rettenberger, M. (2022). Erfassung akut- und stabildynamischer Risikofaktoren sowie devianten und delinquenter Verhaltens bei Pädosexualität: Manual zur Verwendung von Acute-SR, Stable-SR und CMC (Berichte und Materialien, Band 34). Wiesbaden: Kriminologische Zentralstelle (KrimZ). Verfügbar unter: <https://www.krimz.de/fileadmin/dateiablage/E-Publikationen/BM-Online/bm-online34.pdf>
- Hanson RK, Harris A. *Acute-2007 Scoring Guide*. Ottawa, ON: Public Safety Canada (2007).
- Casey S, Day A, Howells K, Ward T. Assessing suitability for offender rehabilitation: development and validation of the treatment readiness questionnaire. *Criminal Justice Behav.* (2007) 34:1427–40. doi: 10.1177/0093854807305827
- Rollnick S, Heather N, Gold R, Hall W. Development of a short 'readiness to change' questionnaire for use in brief, opportunistic interventions among excessive drinkers. *Br J Addict.* (1992) 87:743–54. doi: 10.1111/j.1360-0443.1992.tb02720.x
- Hannöver W, Thyrian JR, Hapke U, Rumpf HJ, Meyer C, John U. The readiness to change questionnaire in subjects with hazardous alcohol consumption, alcohol misuse and dependence in a general population survey. *Alcohol Alcohol.* (2002) 37:362–9. doi: 10.1093/alcalc/37.4.362
- Merz J. A Questionnaire for the measurement of psychological reactance. *Diagnostica.* (1983) 29:75–82.
- Herzberg PY. Zur psychometrischen Optimierung einer Reaktanzskala mittels klassischer und IRT-basierter Analysemethoden. *Diagnostica.* (2002) 48:163–71. doi: 10.1026/0012-1924.48.4.163
- Dunkel D, Antretter E, Fröhlich-Walser S, Haring C. Evaluation of the short-form social support questionnaire (SOZU-K-22) in clinical and non-clinical samples. *Psychother Psychosom Med Psychol.* (2005) 55:266–77. doi: 10.1055/s-2004-834746
- Sommer G, Fydrich T. Soziale Unterstützung. *Diagnostik, Konzepte, Fragebogen F-SozU*. Tübingen: Deutsche Gesellschaft für Verhaltenstherapie (1989).
- Russell DW. UCLA loneliness scale (version 3): reliability, validity, and factor structure. *J Pers Assess.* (1996) 66:20–40. doi: 10.1207/s15327752jpa6601_2
- Bilsky W, Hosser D. Social support and loneliness: psychometric comparison of two scales based on a nationwide representative survey. *Z Differ Diagn Psychol.* (1998) 19:130–44.
- Neutze J, Grundmann D, Scherner G, Beier KM. Undetected and detected child sexual abuse and child pornography offenders. *Int J Law Psychiatry.* (2012) 35:168–75. doi: 10.1016/j.ijlp.2012.02.004

- Bumby KM. Assessing the cognitive distortions of child molesters and rapists: Development and validation of the MOLEST and RAPE scales. *Sexual Abuse: A Journal of Research and Treatment* 8.1, (1996) 37–54.
- Feelgood S, Schaefer G, Hoyer J. *KV-M-Skala Zur Erfassung Kognitiver Verzerrungen Bei Missbrauchern* Dresden: Technische Universität. (2009).
- D'Zurilla TJ, Nezu AM, Maydeu-Olivera A. *Manual for the Social Problem-Solving Inventory-Revised*. North Tonawanda, NY: Multi-Health Systems (2002).
- Nezu CM, Nezu AM, Dudek JA, Peacock MA, Stoll JG. Social problem-solving correlates of sexual deviancy and aggression among adult child molesters. *J Sexual Aggress*. (2005) 11:27–36. doi: 10.1080/13552600512331329540
- Graf A. Psychometrische Überprüfung einer deutschsprachigen Übersetzung des SPSI-R. *Z Differ Diagn Psychol*. (2003) 24:277–91. doi: 10.1024/0170-1789.24.4.277
- Gratz KL, Roemer L. Multidimensional assessment of emotion regulation and dysregulation: development, factor structure, and initial validation of the difficulties in emotion regulation scale. *J Psychopathol Behav Assess*. (2004) 26:41–54. doi: 10.1023/B:JOBA.0000007455.08539.94
- Ehring T, Fischer S, Schnüller J, Bösterling A, Tuschen-Caffier B. Characteristics of emotion regulation in recovered depressed versus never depressed individuals. *Pers Indiv Differ*. (2008) 44:1574–84. doi: 10.1016/j.paid.2008.01.013
- Scherer A, Eberle N, Boecker M, Vögele C, Gauggel S, Forkmann T. The negative affect repair questionnaire: factor analysis and psychometric evaluation in three samples. *BMC Psychiatry*. (2013) 13:16. doi: 10.1186/1471-244X-13-16
- Patton JH, Stanford MS, Barratt ES. Factor structure of the Barratt impulsiveness scale. *J Clin Psychol*. (1995) 51:768–74.
- Prochaska, J., DiClemente, C. C., & Norcross, J. C. (1992). In search of how people change: Applications to addictive behaviors. *American Psychologist*, 47(9), 1102-1114.
- Meule A, Vögele C, Kübler A. Psychometrische Evaluation Der Deutschen Barratt Impulsiveness Scale–Kurzversion (BIS-15). *Diagnostica*. (2011) 57:126–33. doi: 10.1037/t63911-000
- Cortoni F, Marshall WL. Sex as a coping strategy and its relationship to juvenile sexual history and intimacy in sexual offenders. *Sexual Abuse*. (2001) 13:27–43. doi: 10.1177/107906320101300104
- Beier KM, Grundmann D, Kuhle LF, Scherner G, Konrad A, Amelung T. The German Dunkelfeld project: a pilot study to prevent child sexual abuse and the use of child abusive images. *J Sexual Med*. (2015) 12:529–42. doi: 10.1111/jsm.12785
- Klein V, Rettenberger M, Boom KD, Briken P. Eine Validierungsstudie der deutschen Version des Hypersexual Behavior Inventory (HBI). *Psychother Psychosom Med Psychol*. (2014) 64:136–40. doi: 10.1055/s-0033-1357133
- Mack C, Yundina E. "Emotionale Kongruenz mit der Kinderwelt als mögliches diagnostisches Merkmal von Pädophilie?" In Müller JM, Rösler M, Briken P, Fromberger P, Jordan K, editors. *EFPP Jahrbuch 2012 – Empirische Forschung in der Forensischen Psychiatrie, Psychologie und Psychotherapie*, (2012) p. 71–81.
- Tozdan S, Jakob C, Schuhmann P, Budde M, Briken P. Spezifische Selbstwirksamkeit zur Beeinflussung des sexuellen Interesses an Kindern (SSIK): Konstruktion und Validierung eines

Messinstruments. *Psychother Psychosom Med Psychol.* (2015) 65:345–52. doi: 10.1055/s-0035-1548842

Tozdan S, Briken P. The earlier, the worse? Age of onset of sexual interest in children. *J Sexual Med.* (2015) 12:1602–8. doi: 10.1111/jsm.12927

Tozdan S, Briken P. Age of onset and its correlates in men with sexual interest in children. *Sexual Med.* (2019) 7:61–71. doi: 10.1016/j.esxm.2018.10.004

Banse R, Schmidt AF, Clarbour J. Indirect measures of sexual interest in child sex offenders: a multimethod approach. *Criminal Justice Behav.* (2010) 37:319–35. doi: 10.1177/0093854809357598

Briken P. Sexual Outlet Inventory Revised. *Institute for Sex Research and Forensic Psychiatry, University Medical Center Hamburg-Eppendorf, Hamburg, Germany*, (2010).

Wolbers, Marcel, Alessandro Noci, Paul Delmar, Craig Gower-Page, Sean Yiu, and Jonathan W Bartlett. 2022. "Standard and Reference-Based Conditional Mean Imputation." *Pharmaceutical Statistics* 21 (6): 1246–57.

Fay MP, Brittain EH, Proschan MA. Pointwise confidence intervals for a survival distribution with small samples or heavy censoring, *Biostatistics*. 2013;14:723-36.

Fay MP, Proschan MA, Brittain E. Combining one-sample confidence procedures for inference in the two-sample case. *Biometrics*. 2014;71:146-56

12 References

- 1 Fromberger P, Schröder S, Bauer L, et al. @myTabu—A Placebo Controlled Randomized Trial of a Guided Web-Based Intervention for Individuals Who Sexually Abused Children and Individuals Who Consumed Child Sexual Exploitation Material: A Clinical Study Protocol. *Frontiers in Psychiatry* 2021; **11**: 575464.
- 2 Briken P, Berner W, Flöter A, Jückstock V, von Franqué F. Prävention sexuellen Kindesmissbrauchs im Dunkelfeld – das Hamburger Modell. *PSYCH up2date* 2017; **11**: 243–62.
- 3 Wild TSN, Müller I, Fromberger P, Jordan K, Klein L, Müller JL. Prevention of Sexual Child Abuse: Preliminary Results From an Outpatient Therapy Program. *Frontiers in Psychiatry* 2020; **11**: 88.
- 4 Mann RE, Hanson RK, Thornton D. Assessing Risk for Sexual Recidivism: Some Proposals on the Nature of Psychologically Meaningful Risk Factors. *Sexual Abuse* 2010; **22**: 191–217.
- 5 Seto MC, Augustyn C, Roche KM, Hilkes G. Empirically-based dynamic risk and protective factors for sexual offending. *Clinical Psychology Review* 2023; **106**: 102355.
- 6 Babchishin KM, Merdian HL, Bartels RM, Perkins D. Child Sexual Exploitation Materials Offenders: A Review. *European Psychologist* 2018; **23**: 130–43.
- 7 Reichel R, Daser A, Gnielka FM, Schmidt AF, Blokland A, Lehmann RJB. A review of risk factors for online and mixed child sexual abuse material offending: What is being researched? *Journal of Sexual Aggression* 2024; : 1–44.
- 8 Seto MC. The Motivation-Facilitation Model of Sexual Offending. *Sexual Abuse* 2019; **31**: 3–24.
- 9 Rollnick S, Miller WR, Arkowitz H. Motivational Interviewing in the Treatment of Psychological Problems, 2nd edn. New York, NY: Guilford Press, 2015.
- 10 Ward T, Day A, Howells K, Birgden A. The Multifactor Offender Readiness Model. 2004; **9**.
- 11 Terry KJ, Mitchell EW. Motivation and sex offender treatment efficacy: Leading a horse to water and making it drink? *Int J Offender Ther Compar Criminol*. 2001; **45**: 663–72.
- 12 Postel MG, Haan HA, Ter Huurne ED, Becker ES, Jong CA. Effectiveness of a web-based intervention for problem drinkers and reasons for dropout: Randomized controlled trial. *J Med Internet Res* 2010; : 12 68.
- 13 Melville KM, Casey LM, Kavanagh DJ. Dropout from internet-based treatment for psychological disorders. *Br J Clin Psychol* 2010; **49**: 455–71.
- 14 Andrews DA, Bonta J. The Psychology of Criminal Conduct, 5th edn. New Providence, NJ: Lexis Nexis, 2010.

- 15 Hanson RK, Morton-Bourgon KE. Predictors of Sexual Recidivism: An Updated Meta-Analysis. Ottawa, ON: Public Safety Canada, 2004.
- 16 Hanson RK, Harris AJ, Scott TL, Helmus L. Assessing the Risk of Sexual Offenders on Community Supervision: The Dynamic Supervision Project. Ottawa, ON: Public Safety Canada, 2007.
- 17 Knight RA, Thornton D. Evaluating and Improving Risk Assessment Schemes for Sexual Recidivism: A Long-Term Follow-up of Convicted Sexual Offenders. Washington, DC: U. S. Department of Justice, 2007.
- 18 Linehan MM, Wilks CR. [The course and evolution of dialectical behavior therapy](#). *Am J Psychother* 2015; **69**: 97–110.
- 19 Hayes SC, Strosahl K, Wilson KG. Acceptance and Commitment Therapy: An Experiential Approach to Behavior Change. New York, NY: Guilford Press, 1999.
- 20 Berzins LG, Trestman RL. [The development and implementation of dialectical behavior therapy in forensic settings](#). *Int J Forensic Mental Health* 2004; **3**: 93–103.
- 21 Spijkerman MPJ, Pots WTM, Bohlmeijer ET. [Effectiveness of online mindfulness-based interventions in improving mental health: A review and meta-analysis of randomised controlled trials](#). *Clin Psychol Rev* 2016; **45**: 102–14.
- 22 Jayawardene WP, Lohrmann DK, Erbe RG, Torabi MR. [Effects of preventive online mindfulness interventions on stress and mindfulness: A meta-analysis of randomized controlled trials](#). *Prev Med Rep* 2016; **5**: 150–9.
- 23 Lanza PV, Garcia PF, Lamelas FR, González-Menéndez A. [Acceptance and commitment therapy versus cognitive behavioral therapy in the treatment of substance use disorder with incarcerated women](#). *J Clin Psychol* 2014; **70**: 644–57.
- 24 D'Zurilla TJ, Nezu AM, Maydeu-Olivares A. [Social Problem Solving: Theory and Assessment](#). In: Chang EC, D'Zurilla TJ, Sanna LJ, eds. Social problem solving: Theory, research, and training. Washington: American Psychological Association, 2004: 11–27.
- 25 D'Zurilla TJ, Godfried MR. [Problem solving and behavior modification](#). *J Abnorm Psychol* 1971; **78**: 107–26.
- 26 Nezu AM, Nezu CM, D'Zurilla T. Problem-Solving Therapy: A Treatment Manual. New York, NY: Springer Pub. Co, 2013.
- 27 Wakeling HC. [The psychometric validation of the social problem-solving inventory-revised with UK incarcerated sexual offenders](#). *Sexual Abuse* 2007; **19**: 217–36.
- 28 McGuire J. [A review of effective interventions for reducing aggression and violence](#). *Philosophical Transactions of the Royal Society B: Biological Sciences* 2008; **363**: 2577–97.

- 29 Mann RE, Barnett GD. [Treating cognitive components of sexual offending](#). In: Marshall WL, Marshall LE, eds. The Wiley Handbook on the Theories, Assessment, and Treatment of Sexual Offending. Wiley, 2016: 1385–401.
- 30 Helmus L, Hanson RK, Babchishin KM, Mann RE. [Attitudes supportive of sexual offending predict recidivism: A meta-analysis](#). *Trauma Violence Abuse* 2013; **14**: 34–53.
- 31 Beech AR, Bartels RM, Dixon L. [Assessment and Treatment of Distorted Schemas in Sexual Offenders](#). *Trauma, Violence, & Abuse* 2013; **14**: 54–66.
- 32 Lalumière ML, Chalmers LJ, Quinsey VL, Seto MC. [A test of the mate deprivation hypothesis of sexual coercion](#). *Ethol Sociobiol* 1996; **17**: 299–318.
- 33 Matthes A, Rettenberger M. Die Deutsche Version Des Stable-2007. Wien: Institut für Gewaltforschung und Prävention, 2008.
- 34 Ward T, Beech A. [An integrated theory of sexual offending](#). *Aggress Violent Behav* 2006; **11**: 44–63.
- 35 Ward T. [Dynamic risk factors: Scientific kinds or predictive constructs](#). *Psychology, Crime & Law* 2016; **22**: 2–16.
- 36 Olver M, Stockdale K. Evaluating change in men who have sexually offended: Linkages to risk assessment and management. *Curr Psychiatry Rep* 2020; **22**. DOI:[10.1007/s11920-020-01146-3](https://doi.org/10.1007/s11920-020-01146-3).
- 37 Franqué F, V K, P B. [Which techniques are used in psychotherapeutic interventions for nonparaphilic hypersexual behavior?](#) *Sexual Med Rev* 2015; **3**: 3–10.
- 38 Marshall WL, Marshall LE, Serran GA, O'Brien MD. Washington, DC: American Psychological Association, 2011.
- 39 Guse K, Levine D, Martins S, Lira A, Gaarde J, Westmorland W. [Interventions using new digital media to improve adolescent sexual health: A systematic review](#). *J Adolesc Health* 2012; **51**: 535–43.
- 40 Bailey JM, Hsu KJ, Bernhard PA. [An Internet study of men sexually attracted to children: Sexual attraction patterns](#). *Journal of Abnormal Psychology* 2016; **125**: 976–88.
- 41 Löwe B, Spitzer R, Zipfel S, Herzog W. Gesundheitsfragebogen Für Patienten (PHQ-D. *Komplett Kurzform Testmappe Manual Fragebögen Schablonen* 2002; **2**: 5–7.
- 42 Gräfe K, Zipfel S, Herzog W, Löwe B. [Screening psychischer störungen mit dem ‘Gesundheitsfragebogen für Patienten \(PHQ-D\)’](#). *Diagnostica* 2004; **50**: 171–81.
- 43 Seto MC, Stephens S, Lalumière ML, Cantor JM. [The Revised Screening Scale for Pedophilic Interests \(SSPI-2\): Development and Criterion-Related Validation](#). *Sexual Abuse* 2017; **29**: 619–35.

- 44 Screener BPICDS. Institute for Sex Research, Sexual Medicine, and Forensic Psychiatry. Hamburg: University Medical Center Hamburg-Eppendorf, 2019.
- 45 Sonja Etzler, Katharina Nitsche, Ann-Sophie Tröger, Peter Fromberger, Martin Rettenberger. Development and Initial Validation of the Child Sexual Abuse Risk Evaluation Self-Report (CARES) Scales: Assessment of Acute- and Stable-Dynamic Risk Factors via Self-Report in Individuals Convicted of Sexual Offenses Against Children. submitted; published online submitted.
- 46 Casey S, Day A, Howells K, Ward T. [Assessing Suitability for Offender Rehabilitation: Development and Validation of the Treatment Readiness Questionnaire](#). *Criminal Justice and Behavior* 2007; **34**: 1427–40.
- 47 Rollnick S, Heather N, Gold R, Hall W. [Development of a short ‘readiness to change’ questionnaire for use in brief, opportunistic interventions among excessive drinkers](#). *Br J Addict* 1992; **87**: 743–54.
- 48 Hannöver W, Thyrian H JR, U R, HJ M, C J, U. [The readiness to change questionnaire in subjects with hazardous alcohol consumption, alcohol misuse and dependence in a general population survey](#). *Alcohol Alcohol* 2002; **37**: 362–9.
- 49 Merz J. A Questionnaire for the measurement of psychological reactance. *Diagnostica* 1983; **29**: 75–82.
- 50 Herzberg PY. [Zur psychometrischen Optimierung einer Reaktanzskala mittels klassischer und IRT-basierter Analysemethoden](#). *Diagnostica* 2002; **48**: 163–71.
- 51 Dunkel D, Antretter E, Fröhlich-Walser S, Haring C. [Evaluation of the short-form social support questionnaire \(SOZU-K-22\) in clinical and non-clinical samples](#). *Psychother Psychosom Med Psychol* 2005; **55**: 266–77.
- 52 Fydrich T, Sommer G, Tydecks S, Brähler E. Fragebogen zur sozialen Unterstützung (F-SozU): Normierung der Kurzform (K-14). ; : 6.
- 53 Russell DW. [UCLA loneliness scale \(version 3\): Reliability, validity, and factor structure](#). *J Pers Assess* 1996; **66**: 20–40.
- 54 Döring N, Bortz J. Psychometrische Einsamkeitsforschung: Deutsche Neukonstruktion der UCLA Loneliness Scale. *Diagnostica* 1993; **39**: 224–39.
- 55 Neutze J, Grundmann D, Scherner G, Beier KM. [Undetected and detected child sexual abuse and child pornography offenders](#). *International Journal of Law and Psychiatry* 2012; **35**: 168–75.
- 56 Gratz KL, Roemer L. [Multidimensional assessment of emotion regulation and dysregulation: Development, factor structure, and initial validation of the difficulties in emotion regulation scale](#). *J Psychopathol Behav Assess* 2004; **26**: 41–54.

- 57 Ehring T, Fischer S, Schnülle J, Bösterling A, Tuschen-Caffier B. [Characteristics of emotion regulation in recovered depressed versus never depressed individuals](#). *Pers Indiv Differ* 2008; **44**: 1574–84.
- 58 Scherer A, Eberle N, Boecker M, Vögele C, Gauggel S, Forkmann T. [The negative affect repair questionnaire: Factor analysis and psychometric evaluation in three samples](#). *BMC Psychiatry* 2013; **13**: 16.
- 59 Patton JH, Stanford MS, Barratt ES. Factor structure of the Barratt impulsiveness scale. *J Clin Psychol* 1995; **51**: 768–74.
- 60 Meule A, Vögele C, Kübler A. [Psychometrische Evaluation Der Deutschen Barratt Impulsiveness Scale–Kurzversion \(BIS-15\)](#). *Diagnostica*. 2011; **57**: 126–33.
- 61 Cortoni F, Marshall WL. [Sex As a Coping Strategy and Its Relationship to Juvenile Sexual History and Intimacy in Sexual Offenders](#). *Sexual Abuse* 2001; **13**: 27–43.
- 62 Graf A. [Psychometrische Überprüfung einer deutschsprachigen Übersetzung des SPSI-R](#). *Z Differ Diagn Psychol*. 2003; **24**: 277–91.
- 63 Bumby KM. Assessing the cognitive distortions of child molesters and rapists: Development and validation of the MOLEST and RAPE scales. ; : 18.
- 64 Feelgood S, Schaefer GA, Hoyer J. KV-M-Skala zur Erfassung kognitiver Verzerrungen bei Missbrauchern. 2009.
- 65 Klein V, Rettenberger M, Boom K-D, Briken P. [Eine Validierungsstudie der deutschen Version des Hypersexual Behavior Inventory \(HBI\)](#). *PPmP - Psychotherapie · Psychosomatik · Medizinische Psychologie* 2013; **64**: 136–40.
- 66 Tozdan S, Jakob C, Schuhmann P, Budde M, Briken P. [Spezifische Selbstwirksamkeit zur Beeinflussung des sexuellen Interesses an Kindern \(SSIK\): Konstruktion und Validierung eines Messinstruments](#). *PPmP - Psychotherapie · Psychosomatik · Medizinische Psychologie* 2015; **65**: 345–52.
- 67 Briken P. Sexual Outlet Inventory Revised. Institute for Sex Research and Forensic Psychiatry. Hamburg, Germany: University Medical Center Hamburg-Eppendorf, 2010.
- 68 Mack C, Yundina E. Emotionale Kongruenz mit der Kinderwelt als mögliches diagnostisches Merkmal von Pädophilie? In: Müller JM, Rösler M, Briken P, Fromberger P, Jordan K, eds. EFPP Jahrbuch 2012 – Empirische Forschung in der Forensischen Psychiatrie, Psychologie und Psychotherapie. 2012: 71–81.
- 69 Banse R, Schmidt AF, Clarbour J. [Indirect measures of sexual interest in child sex offenders: A multimethod approach](#). *Criminal Justice Behav* 2010; **37**: 319–35.
- 70 Bech P. Measuring the Dimension of Psychological General Well-Being by the WHO-5. *Quality of Life Newsletter* 2004; **32**: 15–6.

- 71 Topp CW, Østergaard SD, Søndergaard S, Bech P. The WHO-5 well-being index: A systematic review of the literature. *Psychother Psychosom* 2015; **84**: 167–76.
- 72 Wilmers F, Munder T, Leonhart R, Herzog T, Plassmann R, Barth J. Die deutschsprachige Version des Working Alliance Inventory-Short Revised (WAI-SR)-Ein schulenübergreifendes, ökonomisches und empirisch validiertes Instrument zur Erfassung der therapeutischen Allianz. *Klin Diagn Eval* 2008; **1**: 343–58.
- 73 Horvath A. The development of the working alliance inventory. New York, NY: Guilford Press, 1986.
- 74 Hopewell S, Chan A-W, Collins GS, *et al.* CONSORT 2025 statement: Updated guideline for reporting randomised trials. *The Lancet* 2025; **405**: 1633–40.