Supplementary material has been published as submitted. It has not been copyedited, or typeset by Acta Dermato-Venereologica

Appendix S1. Deviations from Preregistration (PROSPERO)

[Changes - including crossed out sections in case it was not done anymore or added sections in case of new information - are highlighted in yellow]

Citation

Christina Schut, Johanna LM Munz, Jennifer S Schmidt, Andrea WM Evers, Antoinette van Laarhoven, Frederic Maas genannt Bermpohl, Jörg Kupfer. Psychological interventions for chronic itch in children and adults: a systematic review and meta-analysis. PROSPERO 2021 CRD42021245916 Available from: https://www.crd.york.ac.uk/prospero/display_record.php?ID=CRD42021245916

Review question

To assess the effects of psychological interventions on itch and scratching in children and adults with chronic itch.

Searches

Electronic bibliographic databases: The Cochrane Library, PsycINFO, PubMed and Web of Science. Reference searching: Reference lists of identified trials as well as of reviews and meta-analyses will be handsearched for potentially relevant studies. We will include studies on humans published (or electronically pre-published) in English, German or Dutch from inception until the date the searches will be run (~~approximately until end of summer 2021~~ first database search: May 19th, 2021; final database search: August 23rd, 2023).

Types of study to be included

Randomized controlled trials (RCTs) including parallel-group trials and cross-over trials ~~and Non-randomized studies of interventions (NRSIs)~~ will be included. In line with Cuijpers et al. (2017), we will not consider pre-post effect size which indicate the difference between the outcome score at pre-test and the score at post-test within one group in the meta-analysis. ~~NRSIs are only considered in the qualitative review part. In the quantitative review part (meta-analysis) we~~ We will only include RCTs using between-group effect sizes (Cuijpers et al., 2017). In addition, we will only include RCTs if means and standard deviations are reported in the publication, are provided by the study authors or can be recalculated based on other reported statistical parameters.

Condition or domain being studied

Chronic itch (defined as itch lasting longer than 6 weeks (Ständer et al., 2007)) and/ or chronic itch related diseases. Following the International Forum for the Study of Itch (IFSI; Ständer et al., 2007), 'chronic itch related diseases' are defined as diseases with chronic pruritus due to dermatologic (e.g. atopic dermatitis, psoriasis, urticaria), systemic (e.g. chronic kidney disease, Hodgkin's lymphoma), neurological (e.g. postherpetic neuralgia), psychogenic (e.g. depression, obsessive-compulsive disorders), and/ or mixed causes or as diseases with chronic pruritus of undetermined origin (idiopathic pruritus).

Participants/population

Inclusion: Participants of any age (children and adults) who report chronic itch (itch lasting longer than 6 weeks) and/ or have a clinically diagnosed chronic itch related disease; as already mentioned in section 18, the definition of chronic itch related disease' is based on the internationally accepted clinical classification of itch provided in 2007 by the International Forum for the Study of Itch (IFSI; Ständer et al., 2007). There will be no restrictions in terms of age, sex/ gender, ethnicity, education and economic status.

Exclusion: Studies including only healthy individuals with experimentally induced itch; Participants reporting no or acute itch (itch lasting less than 6 weeks).

Intervention(s), exposure(s)

Inclusion: Interventions should be psychological in nature and may vary in content and composition from traditional psychological therapies (e.g. CBT) to interventions of the third wave (e.g. mindfulness-based interventions). We will include 'psychological interventions' defined as "various forms of treatment and psychoeducation—including psychotherapy and behavior modification, among others—aimed at increasing an individual's adaptive and independent mental and behavioral functioning." (American Psychological Association, 2020). The interventions are delivered by one or more certified individual(s) in an individual, couple or group setting or via guided self-help. The guided self-help intervention should give the participant the option to reflect the applied change techniques regarding emotions, cognitions and/or behavior. ~~They~~ All the interventions may have been carried out multiprofessionally with at least one of the team members with an expertise in the change of emotions, cognitions and behavior. The interventions should have been delivered in person (face-to-face), remotely based on interaction between certified individuals and clients (via telephone or digitally, e.g. e-mail, video conference) or remotely via guided self-help (e.g. printed material). The remotely delivered interventions via technology can be web-, computer- or app-based. There will be no limitations in number of sessions.

Exclusion: We exclude studies testing the effects of non-psychological interventions alone, for example (psycho-)pharmacological interventions.

Comparator(s)/control

Studies with any non-psychological active control group ~~(comparison against a non-psychological or another psychological treatment)~~ or with any inactive control group (waiting list, treatment as usual/standard medical care, psychological/attention placebo, no treatment, or other) will be included. ~~Studies with pre-post-comparisons without control group will only be reported in the qualitative review part.~~

Context

There will be no restrictions in terms of settings.

Main outcome(s)

As chronic itch is long-lasting, we will distinguish between studies with and without follow-up (for the latter: at least one month after the end of the intervention). We will only include studies that used validated instruments to measure the primary outcome measures. We will consider outcome measurement at two time points: 1. Immediately following the end of intervention (post-treatment) and 2. the last available follow-up between one and 12 months.

Primary/Main Outcomes:
Itch measured by self-report (e.g., by use of visual analogue or numerical rating scales).
Scratching measured by self-report or by external rating by a health professional or caregiver.
Excoriations as an indicator of scratching measured by self-report or by external rating by a health professional or caregiver.
Severity of disease and impact on daily functioning measures that include measures of itching and/or scratching (e.g., by SCORAD or POEM).

~~Decrease in itch measured by self-report (e.g., by use of visual analogue or numerical rating scales) from before the intervention (baseline) to the last available time point within a one-year-period.~~
~~Decrease in scratching measured by self-report or by external rating by a health professional or caregiver from before the intervention (baseline) to the last available time point within a one-year-period.~~
~~Decrease in severity of disease and impact on daily functioning measures that include measures of itching and/or scratching (e.g., by SCORAD, DLQI, or POEM) from before the intervention (baseline) to the last available time point within a one-year-period.~~

Measures of effect

As all of the primary outcomes of interest are continuous, we will calculate standardized mean differences (SMDs) with 95% confidence intervals (CIs). We will use Hedges' g as the estimate of the effect size.

Additional outcome(s)

~~As chronic itch is long-lasting, we will distinguish between studies with and without follow-up (for the latter: at least one month after the end of the intervention). We will only include studies that used validated instruments to measure the secondary outcome measures.~~
~~Decrease in the severity of disease measured by instruments that do not include measures of itching and/or scratching (e.g., by the Self-Administered Area and Severity Index (SAPASI)) from before the intervention (baseline) to the last available time point within a one-year period. Decrease in impact on functioning in daily life measured by instruments that do not include measures of itching and/ or scratching (e.g. the Spielberger Trait Anxiety Inventory (STAI)) from before the intervention (baseline) to the last available time point within a one-year-period.~~

Secondary outcomes will not be investigated due to incomplete evidence, as secondary outcomes were not part of the search strategy.

Measures of effect

~~As all of the secondary outcomes of interest are continuous, we will calculate standardized mean differences (SMDs) with 95% confidence intervals (CIs). We will use Hedges' g as the estimate of the effect size.~~

Not applicable.

Data extraction (selection and coding)

Study selection: CS, JM and JK will screen titles and abstracts of the records identified through the search against inclusion criteria. CS, JM and JK will check the made decisions independently. Subsequently, ~~JM~~CS will assess all potentially relevant studies in full-text for eligibility ~~using a standardized Cochrane template~~, followed by ~~CS~~JM checking the made decisions ~~entirely~~ independently and document the made decisions in an SPSS file. The screening of the articles that are no original articles (e.g. reviews) will be distributed between CS and JM. Any disagreements between the two review authors will be resolved by discussion or by consultation of a third author (JK). Where required, further information will be requested from authors. We will exclude duplicates, and collate and aggregate all available data from multiple publications or companion documents of the same study. All reasons for exclusion of analyzed full-text studies will be reported. We will not follow a strict hierarchical approach in marking exclusion criteria, but will select criteria based on what is deemed to be the major reason for exclusion. The entire study selection process will be recorded in detail in a PRISMA flow diagram (Page et al., 2021).

Data extraction: JM and CS will check and extract data from included trials independently using ~~the standardized Cochrane template~~ a standardized MS office excel file ~~and CS will check the extracted data~~. Any discrepancies that arise between the two review authors while comparing the data will be resolved through discussion or, if necessary, by JK. If required, authors of included studies will be asked to provide us with relevant information. ~~Thereafter, we will transfer the data into RevMan (The Cochrane Collaboration, 2020).~~ We will extract e.g. general information: authors names, publication title and date, journal, country, funding, COI; Methodology: study design, total duration of trial, number of arms, risk of bias, trial setting; Participants: mean age, sex/gender, inclusion/exclusion criteria, type of disease/diagnosis, ~~mean duration of disease,~~ severity of disease, total sample size, number within each trial arm, number lost to follow-up; Interventions: type and definition, duration, number of sessions, mode of delivery, conducting persons; Control conditions: type and definition, only for active comparisons: duration, number of sessions, mode of delivery; Outcomes: type and definition of all outcomes, measurement tools.

Risk of bias (quality) assessment

The risk of bias (RoB) assessment will be conducted only for RCTs, using the Cochrane's risk-of-bias tool for randomized trials (RoB 2; Sterne et al., 2019) ~~and the tool ROBINS-I (Sterne et al., 2016) for non-randomized studies of interventions (NRSIs)~~. RoB assessments will be conducted separately for post-treatment and follow-up measurements. We will consider the effect of assignment to the interventions at baseline. Two review authors (AvL and FMgB for English studies, FMgB and JS for German studies) will assess RoB independently and any disagreements between review authors' judgements will be resolved by discussion, or, if needed, by involving a third review author (JS for English studies, JM for German studies). If there is a substantial lack of reported information, we will contact the study authors (e.g. regarding pre-specified data analysis plan). First, we will use the signaling questions and supportive information to rate each domain for each study result. These domain-level judgements feed into an overall 'RoB' judgement for each study result. The following RoB domains will be assessed ~~(divided by tools): RoB 2:~~ 1. Randomization process, 2. Deviations from the intended

interventions, 3. Missing outcome data, 4. Measurement of outcome, 5. Selection of reported result.~~; ROBINS-I: 1. Confounding, 2. Selection of participants into the study, 3. Classifications of interventions, 4. Deviations from intended interventions, 5. Missing outcome data, 6. Measurement of outcomes, 7. Selection of reported result~~.

Strategy for data synthesis

~~We will first provide a narrative synthesis. Thereafter, we will conduct a quantitative synthesis using only the results from the RCTs.~~ For our data synthesis, we~~We~~ will calculate Hedges' g as the estimate of effect sizes, 95% CIs and one-sided p-values for each outcome. Missing standard deviations (SD) will be identified using: 1) standard error or range, 2) CI's, t-value or p-value, or 3) contacting the original authors ~~4) by SDs from other studies using a validated imputation method as described in the Cochrane Handbook (Higgins et al., 2021)~~. We will pool the results using a random effects model. Heterogeneity will be investigated by visual inspection of the forest plots and the I^2-statistic. Following the recommendations of Higgins et al. (2003), we will define I^2-value of 25% as low, 50% as moderate and 75% as high heterogeneity. Data not reported in the included articles but provided by the authors were analyzed per protocol. CS and JM ~~AvL and FMgB~~ will assess the quality of evidence independently using the five criteria of the GRADE approach to rate a body of evidence for every outcome that is important for decision-making (Guyatt et al., 2008): Study limitations, Inconsistency of results, Indirectness of evidence, Imprecision, and Publication bias. The assessment will be discussed with the remaining authors. ~~Disagreements will be resolved by discussion with JS if necessary.~~ We will assess publication bias using funnel plots if 10 trials are available (Sterne et al., 2011). We aim to perform sensitivity analyses by removing studies rated as having 'high risk of bias'. If studies have one intervention condition and more than one comparison condition, the most inactive one - thus the condition with least amount of participant contact - will be chosen (Portnoy et al., 2008). If studies have more than one intervention condition and one comparison condition, we will choose the intervention with the largest amount of psychological content in the review. ~~the data will be split into substudies to compare all intervention groups with the comparison group~~. Where studies include more than one intervention and control group the data will be separated into substudies.

Analysis of subgroups or subsets

If there are at least two studies, subgroup or subset analyses will be conducted in order to differ between the effects of different psychological interventions (nature of intervention) and between effects in different participant groups (e.g., children vs. adults, comparison between different types of diseases). We also aim to look at different effects which arose due to the mode of delivery and the duration of intervention.

Contact details for further information

PD Dr. Christina Schut
Christina.Schut@mp.med.uni-giessen.de

Organisational affiliation of the review

University of Giessen, Institute of Medical Psychology
https://www.uni-giessen.de/fbz/fb11/institute/med_psych

Review team members and their organisational affiliations

PD Dr Christina Schut. University of Giessen, Institute of Medical Psychology
Miss Johanna LM Munz. University of Giessen, Institute of Medical Psychology
Professor Jennifer S Schmidt. University of Applied Science Münster
Professor Andrea WM Evers. University of Leiden, Health, Medical and Neuropsychology Unit
Assistant/Associate Professor Antoinette van Laarhoven. University of Leiden, Health, Medical and Neuropsychology Unit
Mr Frederic Maas genannt Bermpohl. University of Wuppertal, Department of Clinical Psychology and Psychotherapy
PD Dr Jörg Kupfer. University of Giessen, Institute of Medical Psychology

Type and method of review

Intervention, Meta-analysis, Systematic review

Anticipated or actual start date

15 May 2021

Anticipated completion date

30 October 2022

Funding sources/sponsors

None

Grant number(s)

State the funder, grant or award number and the date of award

Conflicts of interest

/

Language

English

Country

Germany, Netherlands

Stage of review

Review Ongoing

Subject index terms status

Subject indexing assigned by CRD

Subject index terms

Adult; Child; Humans; Pruritus; Psychosocial Intervention; Psychotherapy

Date of registration in PROSPERO

27 May 2021

Date of first submission

03 May 2021

Details of any existing review of the same topic by the same authors

/

Stage of review at time of this submission

Stage	Started	Completed
Preliminary searches	Yes	No
Piloting of the study selection process	Yes	No
Formal screening of search results against eligibility criteria	No	No
Data extraction	No	No
Risk of bias (quality) assessment	No	No
Data analysis	No	No

The record owner confirms that the information they have supplied for this submission is accurate and complete and they understand that deliberate provision of inaccurate information or omission of data may be construed as scientific misconduct.
The record owner confirms that they will update the status of the review when it is completed and will add publication details in due course.

Appendix S2. Search strategies

S2a. Search strategy Cochrane Library

#1	MeSH descriptor: [Pruritus] explode all trees
#2	pruri* OR itch* OR scratch*
#3	#1 OR #2
#4	MeSH descriptor: [Psychotherapy] explode all trees
#5	MeSH descriptor: [Counseling] explode all trees
#6	(psych* AND (interven* OR treat* OR therap* OR technique* OR approach* OR manag* OR support* OR program* OR train*)) OR psychotherap* OR ((behavior* OR behaviour* OR cognitive OR psychodynamic*) AND (therap* OR interven* OR treat* OR technique* OR approach* OR psychotherap* OR modif*)) OR "habit reversal" OR biofeedback OR mindful* OR meditat* OR "Acceptance and Commitment" OR compassion OR arousal reduc* OR relax* OR autogenic* OR hypno* OR stress* OR coping OR problem solving OR imagery OR distraction* OR education* OR counseling OR counselling OR telemedicine OR tele-medicine OR telehealth OR tele-health OR ehealth OR e-health OR mobile health OR mhealth OR m-health OR self-help OR self-administ* OR self-manag*
#7	#4 OR #5 OR #6
#8	#3 AND #7

S2b. Search strategy PsycINFO

S1	DE "Pruritus" OR DE "Scratching"
S2	TX (pruri* OR itch* OR scratch*)
S3	S1 OR S2
S4	DE "Psychotherapy" OR DE "Adlerian Psychotherapy" OR DE "Adolescent Psychotherapy" OR DE "Multisystemic Therapy" OR DE "Affirmative Therapy" OR DE "Analytical Psychotherapy" OR DE "Autogenic Training" OR DE "Brief Psychotherapy" OR DE "Brief Relational Therapy" OR DE "Child Psychotherapy" OR DE "Play Therapy"

	OR DE "Client Centered Therapy" OR DE "Conversion Therapy" OR DE "Couples Therapy" OR DE "Eclectic Psychotherapy" OR DE "Emotion Focused Therapy" OR DE "Existential Therapy" OR DE "Experiential Psychotherapy" OR DE "Expressive Psychotherapy" OR DE "Eye Movement Desensitization Therapy" OR DE "Feminist Therapy" OR DE "Geriatric Psychotherapy" OR DE "Gestalt Therapy" OR DE "Empty Chair Technique" OR DE "Group Psychotherapy" OR DE "Encounter Group Therapy" OR DE "Marathon Group Therapy" OR DE "Therapeutic Community" OR DE "Guided Imagery" OR DE "Humanistic Psychotherapy" OR DE "Hypnotherapy" OR DE "Age Regression (Hypnotic)" OR DE "Ericksonian Psychotherapy" OR DE "Posthypnotic Suggestions" OR DE "Individual Psychotherapy" OR DE "Insight Therapy" OR DE "Integrative Psychotherapy" OR DE "Schema Therapy" OR DE "Interpersonal Psychotherapy" OR DE "Logotherapy" OR DE "Narrative Therapy" OR DE "Network Therapy" OR DE "Persuasion Therapy" OR DE "Primal Therapy" OR DE "Psychoanalysis" OR DE "Dream Analysis" OR DE "Self-Analysis" OR DE "Psychodrama" OR DE "Psychodynamic Psychotherapy" OR DE "Psychotherapeutic Counseling" OR DE "Family Therapy" OR DE "Conjoint Therapy" OR DE "Strategic Family Therapy" OR DE "Structural Family Therapy" OR DE "Psychotherapeutic Techniques" OR DE "Active Listening" OR DE "Animal Assisted Therapy" OR DE "Centering" OR DE "Cotherapy" OR DE "Free Association" OR DE "Life Review" OR DE "Mirroring" OR DE "Morita Therapy" OR DE "Motivational Interviewing" OR DE "Mutual Storytelling Technique" OR DE "Paradoxical Techniques" OR DE "Rational Emotive Behavior Therapy" OR DE "Reality Therapy" OR DE "Relationship Therapy" OR DE "Solution Focused Therapy" OR DE "Strategic Therapy" OR DE "Supportive Psychotherapy" OR DE "Transactional Analysis" OR DE "Counseling" OR DE "Community Counseling" OR DE "Cross Cultural Counseling" OR DE "Educational Counseling" OR DE "Genetic Counseling" OR DE "Gerontological Counseling" OR DE "Grief Counseling" OR DE "Group Counseling" OR DE "Marriage Counseling" OR DE "Microcounseling" OR DE "Multicultural Counseling" OR DE "Occupational Guidance" OR DE "Pastoral Counseling" OR DE "Peer Counseling" OR DE "Premarital Counseling" OR DE "Rehabilitation Counseling" OR DE "School Counseling"
S5	TX (psych* AND (interven* OR treat* OR therap* OR technique* OR approach* OR manag* OR support* OR program* OR train*)) OR TX psychotherap* OR TX ((behavior* OR behaviour* OR cognitive OR psychodynamic*) AND (therap* OR interven* OR treat* OR technique* OR approach* OR psychotherap* OR modif*)) OR TX biofeedback OR TX "habit reversal" OR TX "Acceptance and Commitment" OR TX mindful* OR TX meditat* OR TX compassion OR TX (arousal reduc* OR relax* OR

	autogenic*) OR TX hypno* OR TX stress* OR TX coping OR TX problem solving OR TX imagery OR TX distraction* OR TX education* OR TX (counseling OR counselling) OR TX (telemedicine OR tele-medicine OR telehealth OR tele-health OR ehealth OR e-health OR mobile health OR mhealth OR m-health) OR TX (self-help OR self-administ* OR self-manag*)
S6	S4 OR S5
S7	S3 AND S6
S8	PO Animal
S9	PO Human OR PO Transgender OR PO Inpatient OR PO Outpatient
S10	S8 NOT S9
S11	S7 NOT S10

S2c. Search strategy PubMed

#1	Pruritus [MeSH Terms]
#2	pruri* OR itch* OR scratch*
#3	#1 OR #2
#4	Psychotherapy [MeSH Terms] OR Counseling [MeSH Terms]
#5	(psych* AND (interven* OR treat* OR therap* OR technique* OR approach* OR manag* OR support* OR program* OR train*)) OR psychotherap* OR ((behavior* OR behaviour* OR cognitive OR psychodynamic*) AND (therap* OR interven* OR treat* OR technique* OR approach* OR psychotherap* OR modif*)) OR "habit reversal" OR biofeedback OR mindful* OR meditat* OR "Acceptance and Commitment" OR compassion OR arousal reduc* OR relax* OR autogenic* OR hypno* OR stress* OR coping OR problem solving OR imagery OR distraction* OR education* OR counseling OR counselling OR telemedicine OR tele-medicine OR telehealth OR tele-health OR ehealth OR e-health OR mobile health OR mhealth OR m-health OR self-help OR self-administ* OR self-manag*
#6	#4 OR #5

#7	#3 AND #6
#8	Animals [MeSH Terms] NOT Humans [MeSH Terms]
#9	#7 NOT #8

S2d. Search strategy Web of Science

#1	TS=(pruri* OR itch* OR scratch*)
#2	TS=((psych* AND (interven* OR treat* OR therap* OR technique* OR approach* OR manag* OR support* OR program* OR train*)) OR psychotherap* OR ((behavior* OR behaviour* OR cognitive OR psychodynamic*) AND (therap* OR interven* OR treat* OR technique* OR approach* OR psychotherap* OR modif*)) OR "habit reversal" OR biofeedback OR mindful* OR meditat* OR "Acceptance and Commitment" OR compassion OR arousal reduc* OR relax* OR autogenic* OR hypno* OR stress* OR coping OR problem solving OR imagery OR distraction* OR education* OR counseling OR counselling OR telemedicine OR tele-medicine OR telehealth OR tele-health OR ehealth OR e-health OR mobile health OR mhealth OR m-health OR self-help OR self-administ* OR self-manag*)
#3	#1 AND #2
#4	TI=("veterinary" OR "rabbit" OR "rabbits" OR "animal" OR "animals" OR "mouse" OR "mice" OR "rodent" OR "rodents" OR "rat" OR "rats" OR "pig" OR "pigs" OR "porcine" OR "horse" OR "horses" OR "equine" OR "cow" OR "cows" OR "bovine" OR "goat" OR "goats" OR "sheep" OR "ovine" OR "canine" OR "dog" OR "dogs" OR "feline" OR "cat" OR "cats")
#5	#3 NOT #4

Appendix S3. List of extracted data

The following variables, grouped into five categories, were extracted.

***General information*:**

- Author(s); Year of publication
- Citation (according to APA)
- Protocol: register + registration code/number
- Source through which the study was identified
- Data request need: need for data request & if so, which data and email of corresponding author
- Funding
- Conflict of interest
- Country the trial was conducted in

***Methodology*:**

- Study Design: parallel-group or cross-over
- Form of randomization
- Blinding
- Number of trial arms
- Post-assessment: Duration from baseline to post-treatment assessment
- Follow-Up: Duration from baseline to follow-up assessment
- Recruitment method
- Incentive provided for completing the treatment/control group & if so, which one

Participants:

- Target group: study population and age group
- Inclusion and exclusion criteria (especially comorbid disorders)
- Mean age and SD (all groups)
- Number and proportion of female participants (all groups)
- Number of people assessed for eligibility
- Number of participants who declined to participate
- Number of participants randomized (for all groups)
- Number of persons allocated that completed at least one module in the treatment or participated in the control group (for all groups, if possible)
- Number of participants completing all modules as advised (for all groups, if possible)
- Number of participants included in the analysis of the corresponding outcome (all groups)

Interventions:

- Form of treatment(s)
- Guidance: guided vs. self-guided; if guided, person(s) who delivered treatment
- Setting: individual vs. group
- Delivery mode: face-to-face vs. internet
- Number of sessions: in general & per week
- Duration of treatment (in weeks)
- Duration of (extra) sessions (in minutes)
- Total time spent practitioning (in minutes): direct contact; not applicable for self-guided interventions
- Form of control group(s)
- Reasons for dropping out of treatment/control

***Outcome variables*:**

- Form of outcome variable
- Instrument used for measurement of outcome variable
- Form of measurement: self-rated vs. externally rated (if applicable); if externally rated, external person who conducted measurement
- Means and standard deviations at baseline, post-treatment and follow-up assessments (if available) for each outcome variable (all groups)
- Other measures that allow to estimate the effect, if standard deviations were not available (see Appendix S4a)

Appendix S4. Preparation of data

S4a. Estimation of missing standard deviations

1. Study of Melin et al. (1986)

- Estimation method based on Cochrane recommendation (see: https://training.cochrane.org/handbook/current/chapter-06#section-6-5-2-2)
- Estimation:

outcome	N_E	N_C	p value*	degrees of freedom	t statistic	mean difference	SE	SD
itch intensity	15	18	0,001	31	3,63	4,03	1,11	3,18
scratch intensity	15	18	0,001	31	3,63	4,13	1,14	3,26

Note. N_C = number of participants assessed in the control group; N_E = number of participants assessed in the experimental/intervention group; SD = standard deviation; SE = standard error.

* conservative estimation for both outcomes as $p = 0.00$ or $p = 0.000$ was reported in the publication

2. Study of Senser et al. (2004)

- Estimation method based on paper of Walter, S. D. & Yao, X. (2007). Effect sizes can be calculated for studies reporting ranges for outcome variables in systematic reviews. *Journal of Clinical Epidemiology, 60,* 849–852. (see: https://doi.org/10.1016/j.jclinepi.2006.11.003)
- Estimation:

			Conversion factor *f* (see table 1, Walter & Yao, 2007)									
outcome	N_E	N_C	f_E	f_C	range (pre_E)	range (pre_C)	range ($post_E$)	range ($post_C$)	SD (pre_E)	SD (pre_C)	SD ($post_E$)	SD ($post_C$)
itch intensity	4	6	0,486	0,395	2	2	3,2	3,7	0,97	0,79	1,56	1,46
scratch frequency	7	9	0,37	0,337	212	187	112,5	89	78,44	63,02	41,63	29,99

Note. f_C = conversion factor for control group; f_E = conversion factor for experimental/intervention group; N_C = number of participants assessed in control group; N_E = number of participants assessed in experimental/ intervention group; $post_C$ = post-treatment measurement in control group; $post_E$ = post-treatment measurement in experimental/ intervention group; pre_C = baseline measurement in control group; pre_E = baseline measurement in experimental/ intervention group; SD = standard deviation.

S4b. Prioritization of study arms

study	no. arms	intervention group_1	intervention group_2	intervention group_3	intervention group_4	control group	decision
Ehlers et al., 1995	5	DE	AT	CBT	DE + CBT	TAU	We chose DE + CBT as intervention group, as it included the largest amount of psychological intervention elements. We did not consider TAU as control group since this group was not included in the randomization process. We chose DE as control group because it did not contain any psychological treatment.
Niebel et al., 2000	3	PE_{par}_direct	PE_{par}_indirect	-	-	TAU + DE	We chose PE_{par}_direct as intervention group, as it included more active content than PE_{par}_indirect (e.g. group discussion). We chose TAU + DE as control group because this was the only condition that did not contain any psychological treatment.
Rotter et al., 2023	4	HYP	IF	EXE	-	TAU	We chose HYP as intervention group, as this was the only condition that included psychological content. We chose TAU as control group because this was the only condition that did not contain any psychological treatment.
Sokel et al., 1993	3	HYP	BF	-	-	GD	We chose HYP as intervention group, as this was the only condition that focused on itch reduction specifically. We chose GD as control group because this was the only condition that did not contain any psychological treatment.

Note. AT = autogenic training; CBT = cognitive behavioural therapy; DE = dermatological education/ information; EXE = exercise group program; GD = patient group discussion; HYP = hypnotherapy; no. = number; IF = Intermittent fasting with diet adjustments group program; PE_{par}_direct = parental education, delivered in person; PE_{par}_indirect = parental education, delivered indirect via 100-minute-video and booklet; TAU = treatment as usual.

S4c. Prioritization of outcome measures

study	multiple outcome	outcome assessment_1	outcome assessment_2	outcome assessment_3	decision
Ehlers et al., 1995	Itch Int, SR Scratch Int, SR	diary (scale 0-10)	dermatological interview (scale 0-10)	-	We decided to include the assessment by diary. We assume that the assessment by an external person could be more biased due to the mere presence of the rater and/or the way the question was asked.
Hedman-Lagerlöf et al., 2021	Itch Int, SR	VAS	5-D-itch scale	-	We decided to include the assessment by VAS, as it is more comparable to other measures of Itch Int (SR) included in the meta-analysis.
Heratizadeh et al., 2017	Itch Int, SR	PO-SCORAD	SCORAD	-	We decided to include the assessment by PO-SCORAD. We assume that the assessment by an external person could be more biased due to the mere presence of the rater and/or the way the question was asked.
Melin et al., 1989	Scratch Freq, SR	golf counter to monitor scratching for daily frequency of scratching	frequency of scratching in 'worst situation' (scale 0-6)	-	We decided to include the assessment by golf counter, as it is more comparable to other measures of Scratch Freq (SR) included in the meta-analysis.
Rotter et al., 2023	Itch Int, SR	VAS	SCORAD	-	We decided to include the assessment by VAS. We assume that the assessment by an external person could be more biased due to the mere presence of the rater and/or the way the question was asked.

Schut et al., 2013	Itch Int, SR	PO-SCORAD	SCORAD	-	We decided to include the assessment by PO-SCORAD. We assume that the assessment by an external person could be more biased due to the mere presence of the rater and/or the way the question was asked.
Sokel et al., 1993	Exco, ER	severity of surface damage (scale 0-3)	percentage of surface damage coverage	-	We decided to include the assessment by severity of surface damage, as it is more comparable to other measures of Exco (ER) included in the meta-analysis.
Staab et al., 2006	Exco, ER	PO-SCORAD/ Skin Detective (assessed by caregivers)	SCORAD (assessed by dermatologists)	-	We decided to include the assessment by SCORAD. We assume that the external assessment by trained professionals could be less biased and more independent than the external assessment of caregivers.
Van Beugen et al., 2016	Itch Freq, SR	ISDL: Item 14 "My skin disease was accompanied by itching during the past 4 weeks." (scale 1-4)	ISDL: Item 15 "I had attacks of itching during the past 4 weeks." (scale 1-4)	ISDL: Item 16 "I suffered from itching continuously during the past 4 weeks." (scale 1-4)	We decided to include the assessment by ISDL Item 14, as it is more comparable to other measures of Itch Freq (SR) included in the meta-analysis.

Note. ER = externally rated outcome; Exco = excoriations; ISDL = The Impact of Chronic Skin Disease on Daily Life; Itch Freq = itch frequency; Itch Int = itch intensity; PO-SCORAD = Patient-Oriented Scoring for Atopic Dermatitis; SCORAD = Scoring for Atopic Dermatitis; Scratch Freq = scratching frequency; Scratch Int = scratching intensity; SR = self-reported outcome; VAS = visual analogue scale; WL = wait-list.

Appendix S5. Descriptive statistics of outcomes

Table S5a. Means and standard deviations of study outcomes

Study	Outcome measures	Conditions	Baseline		Post-Treatment[a]		Follow-Up[b]		Clinically meaningful?[c] (Yes/No)	
			n	*M* ± *SD*	*n*	*M* ± *SD*	*n*	*M* ± *SD*	BL - PT	BL - FU
Bosecker et al., 2011	Itch Int, SR (VAS; scale 1-10)	PE_{pat}	91	7.4 ± 2.1	91	3.5 ± 2.5	–	–	yes	–
		TAU	175	7.1 ± 2.1	175	3.2 ± 2.3	–	–	yes	–
Ehlers et al., 1995	Itch Int, SR (scale 0-10)	CBT	15	3.2 ± 1.5	–	–	15	1.5 ± 0.9	–	yes
		DE	6	3.2 ± 1	–	–	6	2.9 ± 1.3	–	no
	Itch Freq, SR (no. incidents)	CBT	15	2.4 ± 1.3	–	–	15	1.1 ± 0.9	–	–
		DE	6	2.6 ± 1.3	–	–	6	3 ± 2	–	–
	Scratch Int, SR (scale 0-10)	CBT	15	2.6 ± 1.4	–	–	15	1.3 ± 1	–	–
		DE	6	3.3 ± 1.2	–	–	6	2.7 ± 1.4	–	–
	Scratch Freq, SR (no. incidents)	CBT	15	2.2 ± 1.3	–	–	15	1 ± 0.6	–	–
		DE	6	2.5 ± 1	–	–	6	3 ± 2.1	–	–
Farahani et al., 2013	Itch Int, SR *	RELAX	55	2.6 ± 3.6	55	2.1 ± 1.7	–	–	–	–
		TAU	55	3.5 ± 3.8	55	3.6 ± 2.6	–	–	–	–
Futamura et al., 2013	Exco, ER (SCORAD; scale 0-3)	PE_{par}	29	0.9 ± 0.6	–	–	28	0.3 ± 0.4	–	–
		TAU + DE	30	0.9 ± 0.6	–	–	28	0.6 ± 0.6	–	–
Hedman-Lagerlöf et al., 2021	Itch Int, SR (VAS; scale 0-10)	CBT	51	4.8 ± 2.6	51	2.4 ± 2.1	–	–	yes	–
		TAU + DE	51	4.8 ± 2.4	51	3.4 ± 2.6	–	–	yes	–
	Itch Freq, SR *	CBT	51	5.6 ± 2.6	43	2.7 ± 2.2	–	–	–	–
		TAU + DE	51	6 ± 2.6	48	3.8 ± 2.7	–	–	–	–

Study	Outcome measures	Conditions	Baseline		Post-Treatment[a]		Follow-Up[b]		Clinical meaningfulness[c] (Yes/No)	
			n	*M* ± *SD*	*n*	*M* ± *SD*	*n*	*M* ± *SD*	BL - PT	BL - FU
Heratizadeh et al., 2017	Itch Int, SR (PO-SCORAD; scale 0-10)	PE_{pat}	164	5.7 ± 2.6	141	4 ± 2.6	129	3.8 ± 2.6	yes	yes
		WL	138	6.2 ± 2.3	113	5.4 ± 2.5	104	5 ± 2.8	no	no
	Exco, SR (PO-SCORAD; scale 0-3)	PE_{pat}	164	1.7 ± 1	141	1.3 ± 1	129	1.2 ± 0.9	–	–
		WL	138	1.7 ± 1	113	1.5 ± 1	104	1.4 ± 0.9	–	–
	Exco, ER (SCORAD; scale 0-3)	PE_{pat}	163	1.3 ± 0.8	140	0.9 ± 0.9	125	0.7 ± 0.8	–	–
		WL	142	1.3 ± 0.8	112	1.2 ± 0.9	102	1.1 ± 0.8	–	–
Kishimoto et al., 2023	Itch Int, SR (scale 0-10)	MIND	56	5.3 ± 2.5	54	3 ± 2.5	54	2.5 ± 2.4	yes	yes
		WL	51	5.5 ± 2.5	51	4.6 ± 2.5	51	4.7 ± 2.4	no	no
	Scratch Int, SR (scale 0-10)	MIND	56	5.7 ± 2.3	54	3.2 ± 2.3	54	3.1 ± 2.5	–	–
		WL	51	5.9 ± 2.3	51	5.3 ± 2.3	51	5.1 ± 2.5	–	–
Melin et al., 1986	Itch Int, SR (scale 0-6)	HR	4	3 ± 1	4	2 ± 1.6	–	–	–	–
		TAU	6	3 ± 0.8	6	2 ± 1.5	–	–	–	–
	Scratch Freq, SR (no. incidents)	HR	7	92 ± 78.4	7	24 ± 41.6	–	–	–	–
		TAU	9	90 ± 63	9	45 ± 30	–	–	–	–
Niebel et al., 2000	Exco, ER (scale 0-3)	PE_{par}	18	1.6 ± 1	18	0.7 ± 0.9	–	–	–	–
		TAU + DE	14	1.9 ± 1	14	1.1 ± 1.2	–	–	–	–
Norén et al., 2018	Scratch Freq, ER (no. incidents)	HR	18	19.3 ± 17.8	15	4.7 ± 9.7	–	–	–	–
		TAU	21	18.5 ± 16.7	18	3.4 ± 5.5	–	–	–	–
	Exco, ER (SCORAD; scale 0-3)	HR	18	10.6 ± 6.1	15	1.9 ± 3.7	–	–	–	–
		TAU	21	12.9 ± 6.5	17	4 ± 4.7	–	–	–	–

Study	Outcome measures	Conditions	Baseline		Post-Treatment[a]		Follow-Up[b]		Clinical meaningfulness[c] (Yes/No)	
			n	*M* ± *SD*	*n*	*M* ± *SD*	*n*	*M* ± *SD*	BL - PT	BL - FU
Rotter et al., 2023	Itch Int, SR (VAS; scale 0-100)	HYP	6	63.2 ± 18	6	26 ± 16.4	4	29.5 ± 20.1	–	–
		TAU	9	62.1 ± 17.3	8	39.3 ± 27	9	48.1 ± 24.5	–	–
Santer et al., 2022	Itch Int, SR (scale 0-10)	PE_{pat}	168	NI	130	4.5 ± 2.6	–	–	–	–
		TAU + DE	169	NI	144	4.7 ± 2.7	–	–	–	–
	Itch Freq, SR (POEM; scale 0-4)	PE_{pat}	168	3.2 ± 1.1	135	2.2 ± 1.3	–	–	–	–
		TAU + DE	169	3.2 ± 1.2	148	2.5 ± 1.4	–	–	–	–
Schut et al., 2013	Itch Int, SR (PO-SCORAD; scale 0-10)	CBT	14	3.9 ± 2.4	14	2.8 ± 1.8	–	–	no	–
		WL	14	3 ± 2.1	14	2.2 ± 1.6	–	–	no	–
	Exco, SR (PO-SCORAD; scale 0-3)	CBT	14	1.1 ± 0.9	13	0.7 ± 1	–	–	–	–
		WL	14	0.6 ± 0.6	14	0.6 ± 0.6	–	–	–	–
	Exco, ER (SCORAD; scale 0-3)	CBT	14	0.9 ± 0.7	14	0.9 ± 0.5	–	–	–	–
		WL	14	0.9 ± 0.6	14	0.9 ± 0.8	–	–	–	–
Senser et al., 2004	Itch Int, SR (VAS; scale 0-10)	HYP	15	4.8 ± 1.9	15	2.4 ± 3.2	–	–	yes	–
		WL	18	4.8 ± 2.5	18	6.5 ± 3.2	–	–	no	–
	Scratch Int, SR (VAS; scale 0-10)	HYP	15	4.3 ± 2.1	15	2.2 ± 3.3	–	–	–	–
		WL	18	4.9 ± 2.1	18	6.4 ± 3.3	–	–	–	–
Sokel et al., 1993	Exco, ER (scale 0-3)	HYP	18	20.8 ± 10.2	12	17.5 ± 10.9	12	15.9 ± 8.1	–	–
		GD	16	18.4 ± 13.5	10	16.9 ± 13.6	10	22.5 ± 12.5	–	–

Study	Outcome measures	Conditions	Baseline		Post-Treatment[a]		Follow-Up[b]		Clinical meaningfulness[c] (Yes/No)	
			n	*M* ± *SD*	*n*	*M* ± *SD*	*n*	*M* ± *SD*	BL – PT	BL – FU
Staab et al., 2006_substudy1	Exco, ER (SCORAD; scale 0-3)	PE_{par}	274	1.2 ± 0.9	–	–	274	0.7 ± 0.8	–	–
		WL	244	1.1 ± 0.8	–	–	244	0.8 ± 0.8	–	–
Staab et al., 2006_substudy2	Exco, ER (SCORAD; scale 0-3)	PE_{par}/PE_{pat}	102	1.3 ± 0.9	–	–	102	0.8 ± 0.8	–	–
		WL	83	1.5 ± 0.9	–	–	83	1 ± 0.8	–	–
Staab et al., 2006_substudy3	Itch Int, SR (SCORAD; scale 0-10)	PE_{pat}	70	5.5 ± 2.2	–	–	70	2.8 ± 1.8	–	yes
		WL	50	4.9 ± 2.3	–	–	50	4.4 ± 2.5	–	no
	Exco, ER (SCORAD; scale 0-3)	PE_{pat}	70	1.3 ± 0.8	–	–	69	0.8 ± 0.7	–	–
		WL	50	1.2 ± 1	–	–	50	1.1 ± 0.8	–	–
van Beugen et al., 2016	Itch Int, SR (ISDL; scale 0-10)	CBT	59	5.3 ± 2.6	45	3.4 ± 2.4	36	3.9 ± 2.7	yes	yes
		TAU	58	4.8 ± 2.9	50	4.1 ± 3	44	3.6 ± 2.3	no	no
	Itch Freq, SR (ISDL; scale 1-4)	CBT	60	2.8 ± 0.9	44	2.2 ± 1	37	2.3 ± 0.9	–	–
		TAU	59	2.7 ± 1	49	2.2 ± 1	43	2.1 ± 1	–	–
Zhai et al., 2023	Itch Int, SR (scale 0-10)	CBT	11	6.8 ± 2.4	11	2.4 ± 1.4	–	–	yes	–
		TAU	9	6.7 ± 1.9	9	4 ± 2.2	–	–	yes	–
	Exco, ER (EASI; scale 0-3)	CBT	11	6.3 ± 2.3	11	2.3 ± 1.5	–	–	–	–
		TAU	9	5.6 ± 1.6	9	3.1 ± 1.4	–	–	–	–

Note. BL = Baseline; CBT = cognitive behavioural therapy; DE = dermatological education/ information; EASI = Eczema Area and Severity Index; ER = externally rated outcome; Exco = excoriations; FU = last available follow-up assessment; GD = patient group discussion; HR = Habit Reversal Training; HYP = hypnotherapy; Itch Freq = itch frequency; Itch Int = itch intensity; ISDL = The Impact of Chronic Skin Disease on Daily Life; MIND = Mindfulness and Self-Compassion Training; *n* = sample size; NI = no information; NRS = Numerical Rating Scale; PE_{par} = parental education; PE_{pat} = patient education; POEM = Patient-Oriented Eczema Measure; PO-SCORAD = Patient-Oriented Scoring for Atopic Dermatitis; post-treatment assessment (immediately after intervention); RELAX = relaxation therapy; SCORAD = Scoring for Atopic Dermatitis; Scratch Freq = scratching frequency; Scratch Int = scratching intensity; SR = self-reported outcome; TAU = treatment as usual; VAS = visual analogue scale; WL = wait-list.

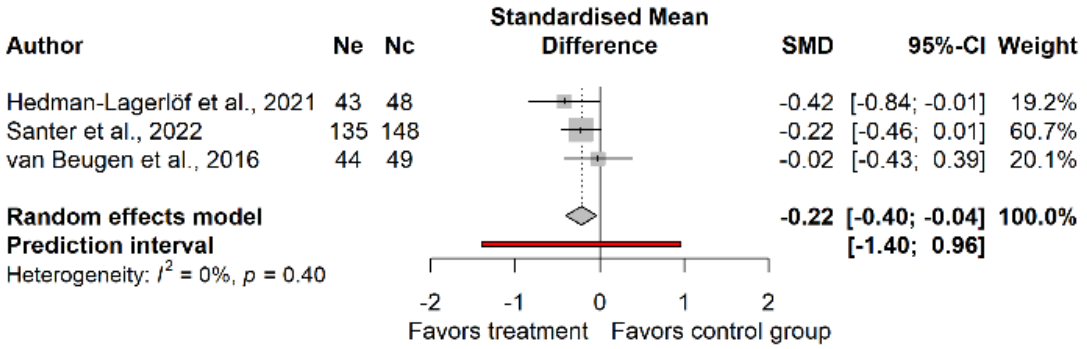
* outcome scale not reported

[a] immediately after intervention; [b] last available follow-up assessment; [c] Minimal Clinically Important Difference between BL and PT assessments (BL – PT) and between BL and FU assessments (BL – FU), evaluated according to Riepe C, Osada N, Reich A et al. Minimal Clinically Important Difference in Chronic Pruritus Appears to be Dependent on Baseline Itch Severity. Acta Derm Venereol. 2019;99:1288-1290; could only be evaluated for studies assessing itch intensity on scales of 0 to 10.

Appendix S6. Forest Plots

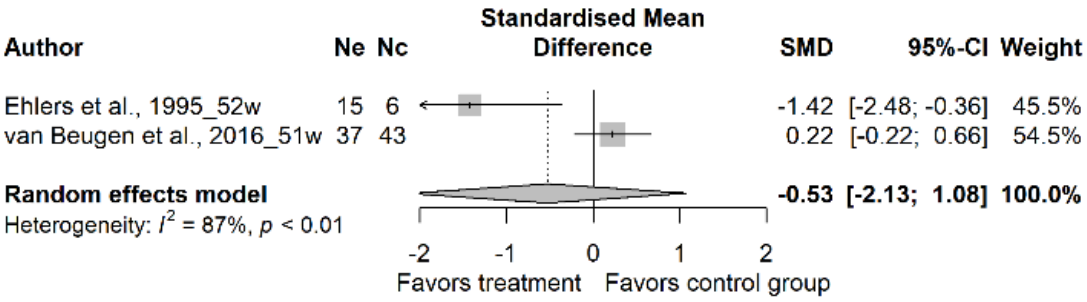
S6a. Forest Plot – Itch frequency

Post-Treatment



Illustrated are post-treatment effects of psychological interventions on itch frequency. Abbreviations: Nc = Number of persons in the control group; Ne = Number of persons in the experimental group; SMD = Standardized mean difference (Hedge's *g*); CI = Confidence Interval.

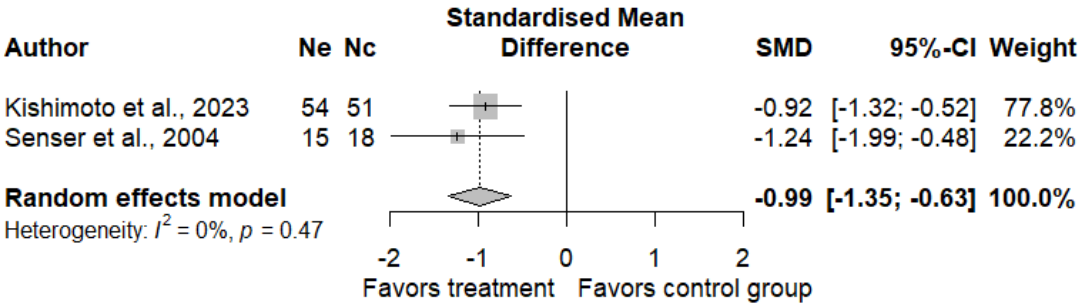
Follow-Up



Illustrated are follow-up effects of psychological interventions on itch frequency. Abbreviations: Nc = Number of persons in the control group; Ne = Number of persons in the experimental group; SMD = Standardized mean difference (Hedge's g); CI = Confidence Interval.

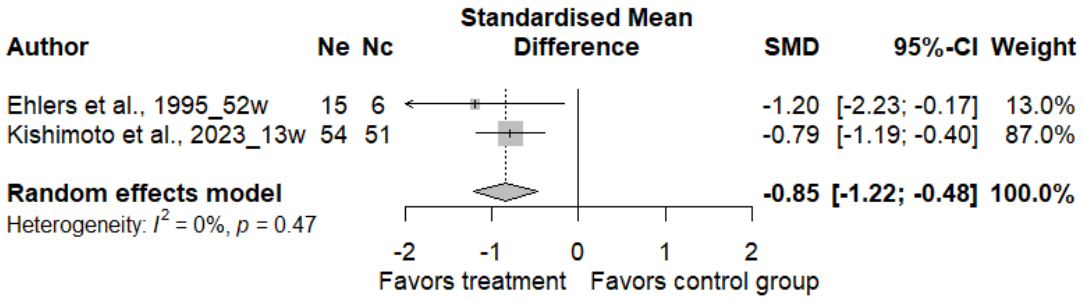
S6b. Forest Plot – Scratching intensity

Post-Treatment



Illustrated are post-treatment effects of psychological interventions on scratching intensity. Abbreviations: Nc = Number of persons in the control group; Ne = Number of persons in the experimental group; SMD = Standardized mean difference (Hedge's g); CI = Confidence Interval.

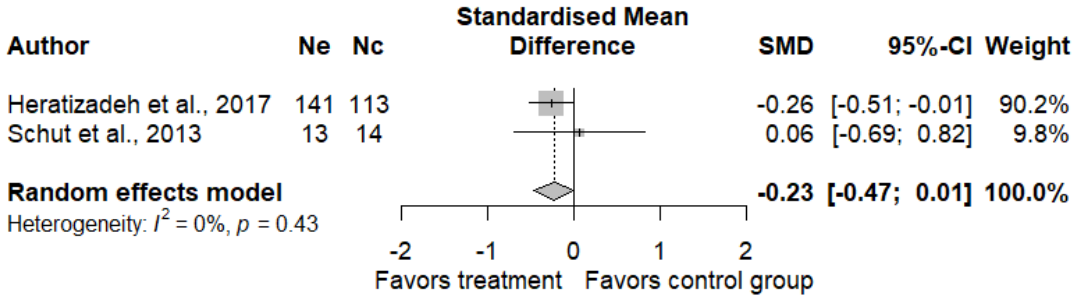
Follow-Up



Illustrated are follow-up effects of psychological interventions on scratching intensity. Abbreviations: Nc = Number of persons in the control group; Ne = Number of persons in the experimental group; SMD = Standardized mean difference (Hedge's g); CI = Confidence Interval.

S6c. Forest Plot – Self-rated excoriations

Post-Treatment



Illustrated are post-treatment effects of psychological interventions on self-rated excoriations. Abbreviations: Nc = Number of persons in the control group; Ne = Number of persons in the experimental group; SMD = Standardized mean difference (Hedge's g); CI = Confidence Interval.

Appendix S7. GRADE assessment

Table S7a. Results of GRADE rating for **Post-Treatment measurement**

Quality assessment						Summary of findings				
						No. of participants		Anticipated absolute effects (95% CI)		
Outcome (No. studies)	Limitations (Risk of Bias)[a]	Publication Bias[b]	Indirectness[c]	Imprecision[d]	Inconsistency[e]	Intervention	Control	Effect for no psychological interventions (CG)	Effect for psychological interventions (IG)	Certainty of the evidence (GRADE)
Itch intensity, SR (12 RCTs)	Serious limitations	Likely	No serious indirectness	Serious imprecision	No serious inconsistency	617	694	***SMD* 0.37 *SD* lower** (0.60 lower to 0.15 lower) **in IG than CG**		⨁◯◯◯ **Very low**
Itch frequency, SR (3 RCTs)	Serious limitations	Very likely	No serious indirectness	Serious imprecision	No serious inconsistency	222	245	***SMD* 0.22 *SD* lower** (0.40 lower to 0.04 lower) **in IG than CG**		⨁◯◯◯ **Very low**
Scratching intensity, SR (2 RCTs)	Serious limitations	Undetected	No serious indirectness	Very serious imprecision	No serious inconsistency	69	69	***SMD* 0.99 *SD* lower** (1.35 lower to 0.63 lower) **in IG than CG**		⨁◯◯◯ **Very low**
Excoriations, SR (2 RCTs)	Very serious limitations	Undetected	No serious indirectness	Serious imprecision	No serious inconsistency	154	127	***SMD* 0.23 *SD* lower** (0.47 lower to 0.01 higher) **in IG than CG**		⨁◯◯◯ **Very low**
Excoriations, ER (6 RCTs)	Very serious limitations	Likely	No serious indirectness	Serious imprecision	No serious inconsistency	210	176	***SMD* 0.29 *SD* lower** (0.49 lower to 0.09 lower) **in IG than CG**		⨁◯◯◯ **Very low**

Note. CG = control groups; CI = confidence interval; ER = externally rated outcome; GRADE = Grading of Recommendations Assessment, Development, and Evaluation; IG = intervention groups; No. = Number; *SD* = standard deviation; *SMD* = standardized mean difference (Hedge's g); SR = self-reported outcome; RCTs = randomized controlled trials; RoB = Risk of Bias.

[a] **Limitations**: *SR itch intensity*: More than half of the studies (Hedman-Lagerlöf et al., 2021; Heratizadeh et al., 2017; Kishimoto et al., 2023; van Beugen et al., 2016) which substantially contributed to the pooled effect size due to sample size ($n \geq 45$ per group), revealed only some concerns or low RoB in the RoB assessment; this is why we only downgraded one level. *SR itch frequency*: As two of the three studies only revealed some concerns in the RoB assessment (Hedman-Lagerlöf et al., 2021; van Beugen et al., 2016), we only downgraded one level. *SR scratching intensity*: Limitations were downgraded only one level as both included studies (Kishimoto et al., 2023; Senser et al., 2004) revealed an overall RoB of only some concerns. *SR and ER excoriations*: We downgraded two levels since all included studies had an overall high RoB.

[b] **Publication Bias:** Regarding all outcomes, visual inspection of funnel plots did not indicate publication bias. *SR scratching intensity and SR excoriations*: We did not detect a publication bias ("undetected") as there was only one study assessing SR scratching intensity and no study assessing SR excoriations that could not be included in the analysis due to missing data; *SR itch intensity*: Egger's test was only conducted for SR itch intensity since it was the only outcome with $k \geq 10$ studies; it also did not provide any evidence for publication bias. We downgraded one level because one industry-sponsored study with a small sample size was identified (Melin et al., 1986) and we detected eight studies that could not be included in the meta-analysis due to missing data. *SR itch frequency*: We downgraded two levels as we identified $k = 5$ studies that could not be included in the meta-analysis due to missing data. *ER excoriations*: We downgraded one level as we identified $k = 6$ studies that could not be included in the meta-analysis due to missing data.

[c] **Indirectness:** Our aim was to investigate effects of all kinds of interventions with a psychological component delivered by internet as well as in face-to-face settings in patients of all age groups with all kinds of chronic itch due to different origin. This aim was mainly reached as effects of different kinds of psychological interventions were tested in different patient and age groups in the included studies. Moreover, there were no concerns regarding the directness of outcome measurement as no surrogate parameters were used and the test criteria of the used instruments were sufficient.

[d] **Imprecision:** We defined $g = -0.2$ as our minimally important effect. Even though the CI of pooled effect size of *SR scratching* intensity did only include effects of $g \leq -0.2$ and the point estimate was $g = -0.99$, we downgraded two levels as the sample size was small (30% of optimal information size could not be reached). We also downgraded one level for all other outcomes as CI of pooled effect size included $g = -0.2$, but not $g = 0.2$.

[e] **Inconsistency:** For all outcome parameters, no serious inconsistency was detected. Even though I^2 was 69% for itch intensity, two authors (CS and JM) found plausible explanations for this heterogeneity of results (e.g. differences in severity of skin disease at baseline, in type of intervention and delivery mode, and in outcome measures). Subgroup analyses were conducted if a sufficient number of studies was available. These did not reveal any significant group differences.

Table S7b. Results of GRADE rating for **Follow-Up measurement**

Quality assessment						Summary of findings				
						No. of participants		Anticipated absolute effects (95% CI)		
Outcome (No. studies)	Limitations (Risk of Bias)[a]	Publication Bias[b]	Indirectness[c]	Imprecision[d]	Inconsistency[e]	Intervention	Control	Effect for no psychological interventions (CG)	Effect for psychological interventions (IG)	Certainty of the evidence (GRADE)
Itch intensity, SR (6 RCTs)	Serious limitations	Very likely	No serious indirectness	No serious imprecision	No serious inconsistency	308	264	***SMD* 0.59 *SD* lower** (0.97 lower to 0.21 lower) **in IG than CG**		⨁◯◯◯ **Very low**
Itch frequency, SR (2 RCTs)	Serious limitations	Very Likely	No serious indirectness	Very serious imprecision	No serious inconsistency	52	49	***SMD* 0.53 *SD* lower** (2.13 lower to 1.08 higher) **in IG than CG**		⨁◯◯◯ **Very low**
Scratching intensity, SR (2 RCTs)	Serious limitations	Undetected	No serious indirectness	Very serious imprecision	No serious inconsistency	69	57	***SMD* 0.85 *SD* lower** (1.22 lower to 0.48 lower) **in IG than CG**		⨁◯◯◯ **Very low**
Excoriations, ER (6 RCTs)	Very serious limitations	Very likely	No serious indirectness	Serious imprecision	No serious inconsistency	610	517	***SMD* 0.34 *SD* lower** (0.53 lower to 0.15 lower) **in IG than CG**		⨁◯◯◯ **Very low**

Note. CG = control groups; CI = confidence interval; ER = externally rated outcome; GRADE = Grading of Recommendations Assessment, Development, and Evaluation; IG = intervention groups; No. = Number; *SD* = standard deviation; *SMD* = standardized mean difference (Hedge's g); SR = self-reported outcome; RCTs = randomized controlled trials; RoB = Risk of Bias.

[a] **Limitations**: *SR itch intensity*: Half of the studies (Heratizadeh et al., 2017; Kishimoto et al., 2023; van Beugen et al., 2016) which substantially contributed to the pooled effect size due to sample size ($n \geq 36$ per group), revealed only some concerns or low RoB in the RoB assessment; this is why we only downgraded one level. *SR itch frequency and SR scratching intensity*: As one of the two studies each revealed only some concerns in the RoB assessment (Kishimoto et al., 2023; van Beugen et al., 2016), we only downgraded one level. *ER excoriations*: We downgraded two levels since the majority (5/6) of included studies had an overall high RoB.

[b] **Publication Bias:** Regarding all outcomes, visual inspection of funnel plots did not indicate publication bias. *SR itch intensity, SR itch frequency and ER excoriations*: We downgraded two levels as we identified more studies not included in the analysis due to missing data than included in the analyses (SR itch intensity: $k = 8$ vs. $k = 6$; SR itch frequency: $k = 6$ vs. $k = 2$; ER excoriations: $k = 9$ vs. $k = 6$); *SR scratching intensity*: We did not detect publication bias ("undetected") as we only found one study not included in the analysis due to missing outcome data.

[c] **Indirectness:** Our aim was to investigate effects of all kinds of interventions with a psychological component delivered by internet as well as in face-to-face settings in patients of all age groups with all kinds of chronic itch due to different origin. This aim was mainly reached as effects of different kinds of psychological interventions were tested in different patient and age groups in the included studies. Moreover, there were no concerns regarding the directness of outcome measurement as no surrogate parameters were used and the test criteria of the used instruments were sufficient.

[d] **Imprecision:** We defined $g = -0.2$ as our minimally important effect. *SR itch intensity*: We did not downgrade as the CI of the pooled effect size did only include effects of $g < -0.2$ with medium effect size. *SR itch frequency*: We downgraded two levels since the CI of the pooled effect size did include effects of $0.2 < g < -0.2$. *SR scratching intensity*: Even though the CI of pooled effect size did only include effects of $g \leq -0.2$ and the point estimate was $g = -0.85$, we downgraded two levels as the sample size was small (30% of optimal information size could not be reached). *ER excoriations*: We downgraded one level as CI of pooled effect size crossed $g = -0.2$, but not $g = 0.2$.

[e] **Inconsistency:** For all outcome parameters, no serious inconsistency was detected. Even though I^2 was 69%, 87% and 49% for itch intensity, itch frequency, and externally rated excoriations, respectively, two authors (CS and JM) found plausible explanations for this heterogeneity of results (e.g. differences in severity of skin disease at baseline, in type of intervention and delivery mode, and in outcome measures). Subgroup analyses were conducted if a sufficient number of studies was available. These did not reveal any significant group differences.

S7c. Risk of bias assessment

(1) Post-Treatment

(a) Itch intensity (self-rated)

Study	D1	D2	D3	D4	D5	Overall
Bosecker et al., 2011	-	-	-	-	+	-
Farahani et al., 2013	-	-	-	-	!	-
Hedman-Lagerlöf et al., 2021	+	!	+	-	+	!
Heratizadeh et al., 2017	+	+	+	-	+	+
Kishimoto et al., 2023	+	!	+	-	+	!
Melin et al., 1986	-	-	-	-	!	-
Rotter et al., 2023	-	-	-	-	!	-
Santer et al., 2022	+	!	-	-	+	-
Schut et al., 2013	+	-	-	-	!	-
Senser et al., 2004	!	!	+	-	!	!
van Beugen et al., 2016	!	!	+	-	+	!
Zhai et al., 2023	-	!	+	-	!	-

(b) Itch frequency (self-rated)

Study	D1	D2	D3	D4	D5	Overall
Hedman-Lagerlöf et al., 2021	+	!	+	-	+	!
Santer et al., 2022	+	!	-	-	+	-
van Beugen et al., 2016	!	!	+	-	+	!

(c) Scratching intensity (self-rated)

Study	D1	D2	D3	D4	D5	Overall
Kishimoto et al., 2023	+	!	+	-	+	!
Senser et al., 2004	!	!	+	-	!	!

(d) Excoriations (self-rated)

Study	D1	D2	D3	D4	D5	Overall
Heratizadeh et al., 2017	+	+	+	-	+	-
Schut et al., 2013	+	-	-	-	!	-

(e) Excoriations (externally rated)

Study	D1	D2	D3	D4	D5	Overall
Heratizadeh et al., 2017	+	+	+	-	+	-
Niebel et al., 2000	-	-	-	-	!	-
Norén et al., 2018	+	-	-	+	!	-
Schut et al., 2013	+	-	-	+	!	-
Sokel et al., 1993	!	-	-	+	!	-
Zhai et al., 2023	-	!	+	-	!	-

Domains:
D1: Randomisation process
D2: Deviations from the intended interventions
D3: Missing outcome data
D4: Measurement of the outcome
D5: Selection of the reported result

Judgement:

+ Low risk
! Some concerns
- High risk

2) Follow-Up

(a) Itch intensity (self-rated)

Study	D1	D2	D3	D4	D5	Overall
Ehlers et al., 1995	!	-	-	-	!	-
Heratizadeh et al., 2017	+	+	+	-	+	+
Kishimoto et al., 2023	+	!	+	-	+	!
Rotter et al., 2023	-	-	-	-	!	-
Staab et al., 2006_substudy3	+	-	-	-	!	-
van Beugen et al., 2016	!	!	+	-	+	!

(b) Itch frequency (self-rated)

Study	D1	D2	D3	D4	D5	Overall
Ehlers et al., 1995	!	-	-	-	!	-
van Beugen et al., 2016	!	!	+	-	+	!

(c) Scratching intensity (self-rated)

Study	D1	D2	D3	D4	D5	Overall
Ehlers et al., 1995	!	-	-	-	!	-
Kishimoto et al., 2023	+	!	+	-	+	!

(d) Excoriations (externally rated)

Study	D1	D2	D3	D4	D5	Overall
Futamura et al., 2013	+	!	+	+	+	!
Heratizadeh et al., 2017	+	+	+	-	+	-
Sokel et al., 1993	!	-	-	+	!	-
Staab et al., 2006_substudy1	+	-	-	-	!	-
Staab et al., 2006_substudy2	+	-	-	-	!	-
Staab et al., 2006_substudy3	+	-	-	-	!	-

Domains:
D1: Randomisation process
D2: Deviations from the intended interventions
D3: Missing outcome data
D4: Measurement of the outcome
D5: Selection of the reported result

Judgement:

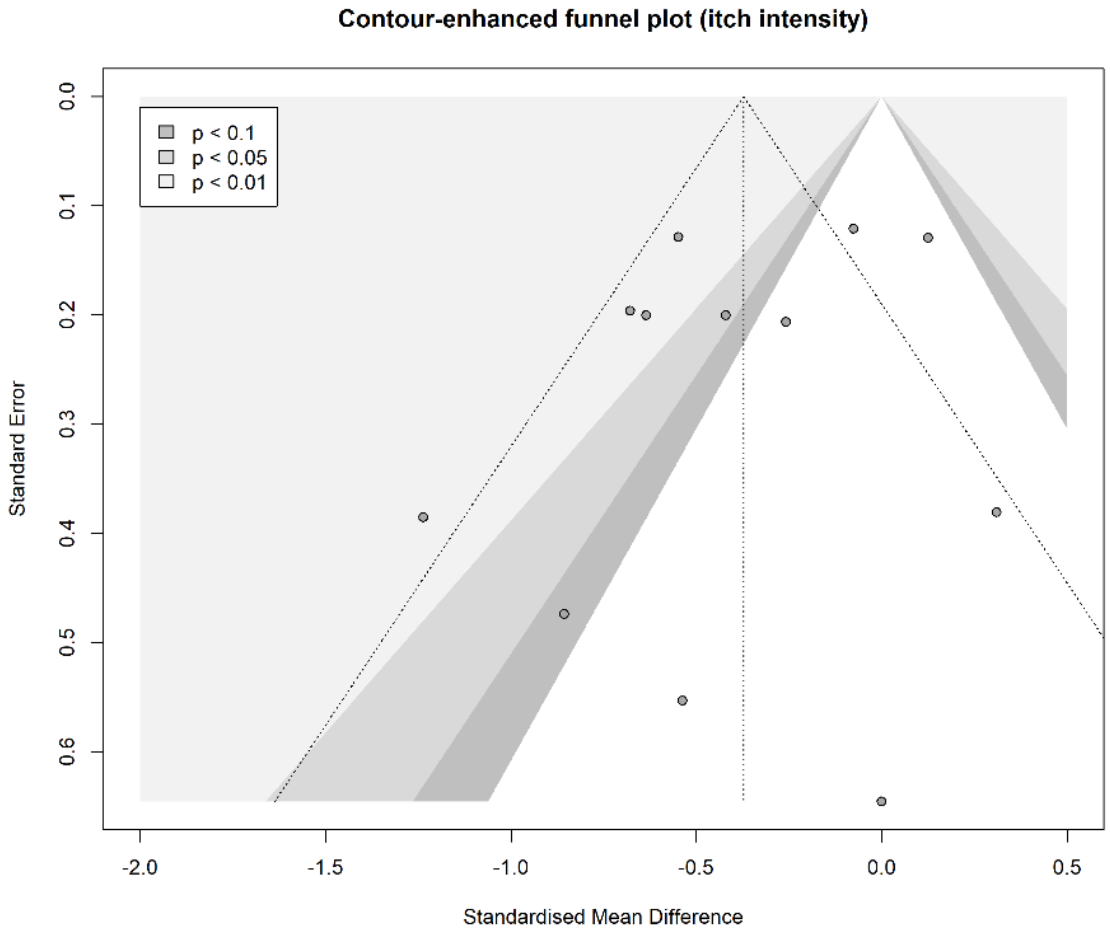
+ Low risk
! Some concerns
- High risk

S7d. Publication bias

(1) Post-Treatment

(a) Itch intensity

Funnel plot



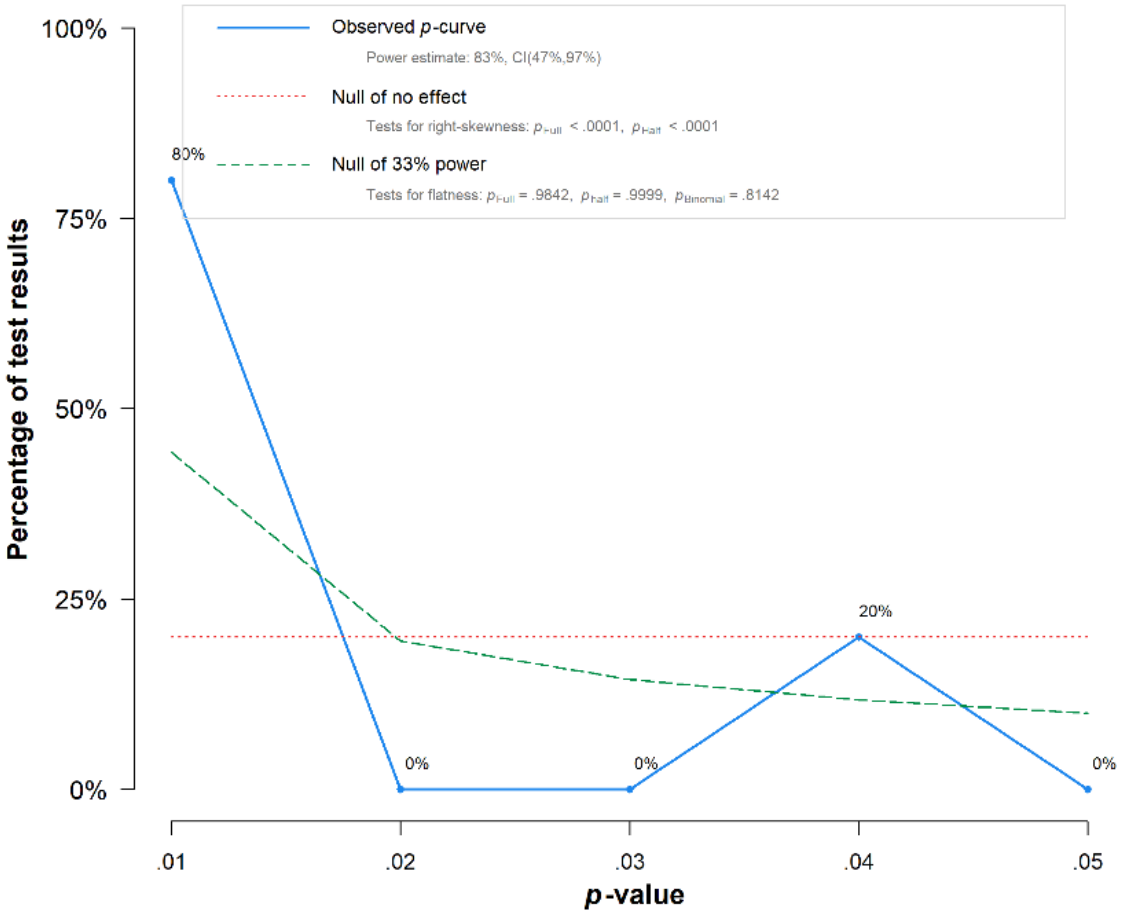
Egger's test

```
> eggers.test(m.cont.post.II)
Eggers' test of the intercept
=============================

 intercept       95% CI      t         p
    -1.164 -3.37 - 1.04 -1.033 0.3259882

Eggers' test does not indicate the presence of funnel plot asymmetry.
```

p-curve

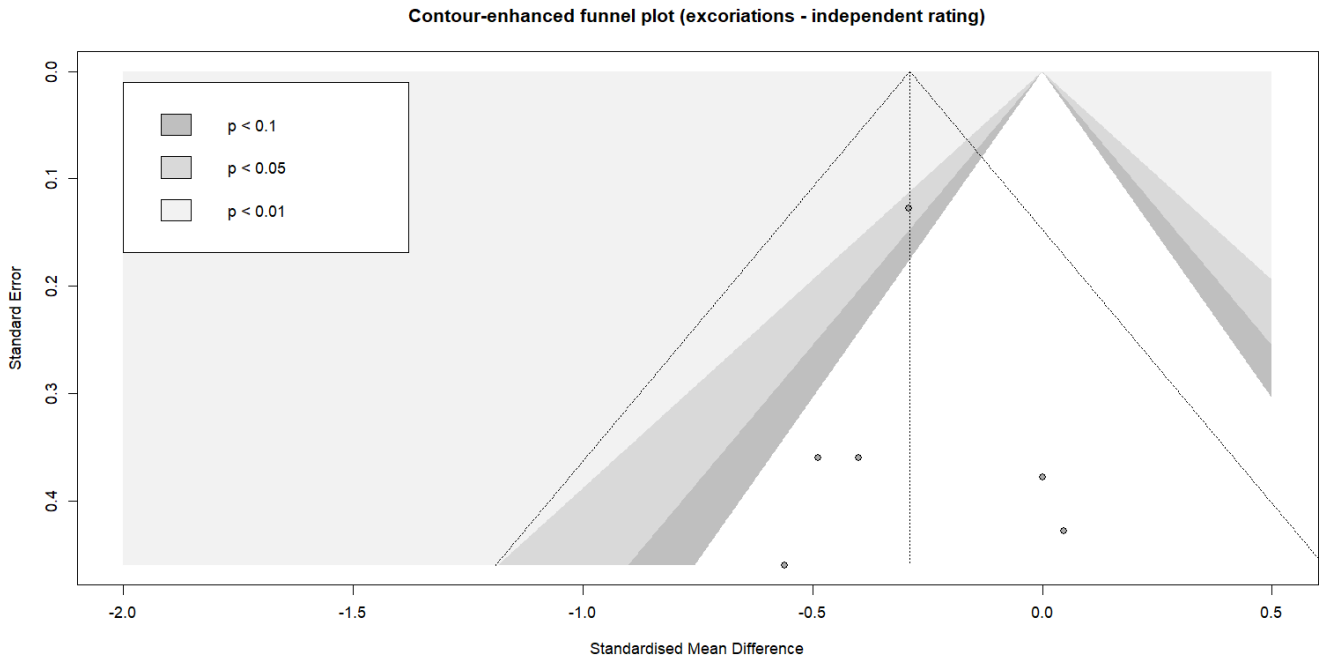


Note: The observed *p*-curve includes 5 statistically significant ($p < .05$) results, of which 4 are $p < .025$.
There were 7 additional results entered but excluded from *p*-curve because they were $p > .05$.

(1) Post-Treatment

(b) Externally rated excoriations

Funnel plot



Egger's test

```
> eggers.test(m.cont.post.EXCO.extern)
Eggers' test of the intercept
============================

 intercept       95% CI     t         p
     0.048 -1.06 - 1.16 0.084 0.9371489

Eggers' test does not indicate the presence of funnel plot asymmetry.
Warning message:
In eggers.test(m.cont.post.EXCO.extern) :
  Your meta-analysis contains k = 6 studies. Egger's test may lack the statistical power to detect bias when th
e number of studies is small (i.e., k<10).
```

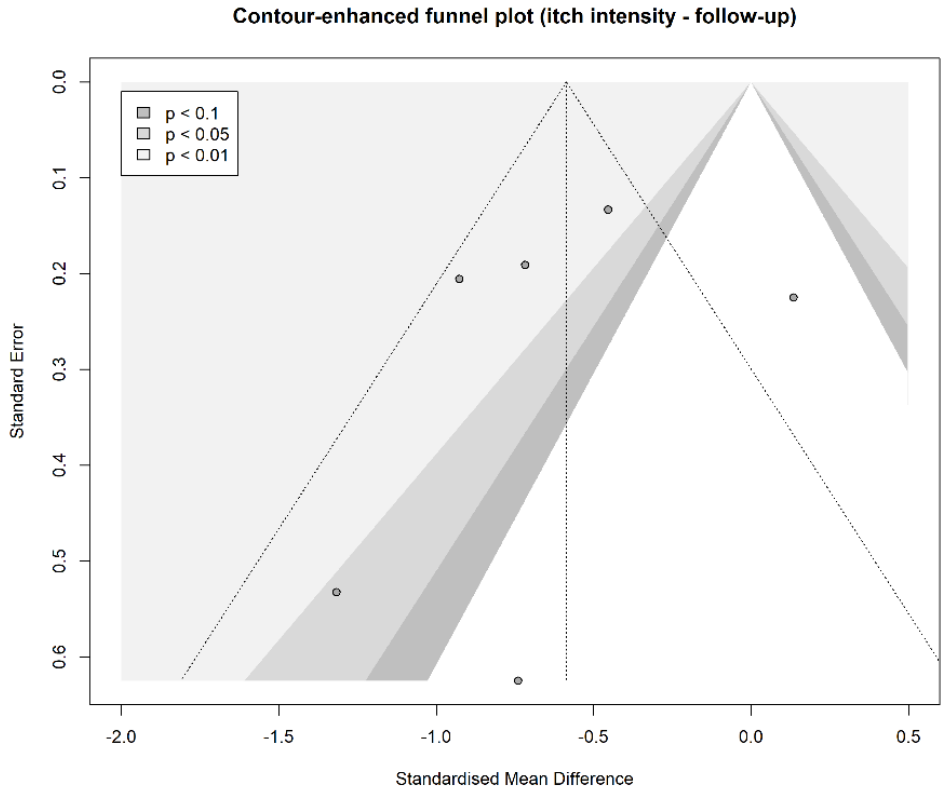
p-curve

```
Error in pcurve(m.cont.post.EXCO.extern) :
  Two or less significant (p<0.05) effect sizes were detected, so p-curve analysis cannot be conducted.
```

(2) Follow-Up

(a) Itch intensity

Funnel Plot



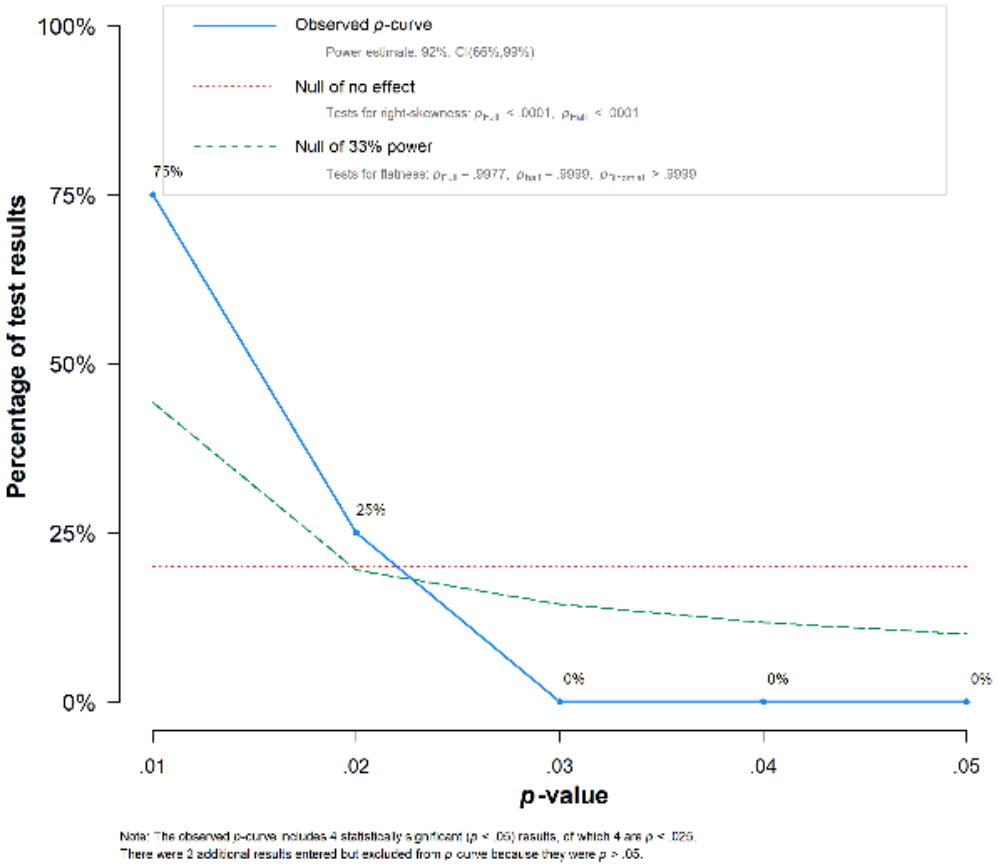
Egger's test

```
> eggers.test(m.cont.FU.II)
Eggers' test of the intercept
============================

 intercept       95% CI      t         p
    -1.012 -4.61 - 2.58 -0.552 0.6105178

Eggers' test does not indicate the presence of funnel plot asymmetry.
Warning message:
In eggers.test(m.cont.FU.II) :
  Your meta-analysis contains k = 6 studies. Egger's test may lack the statistical power to detect bias when th
e number of studies is small (i.e., k<10).
```

p-curve

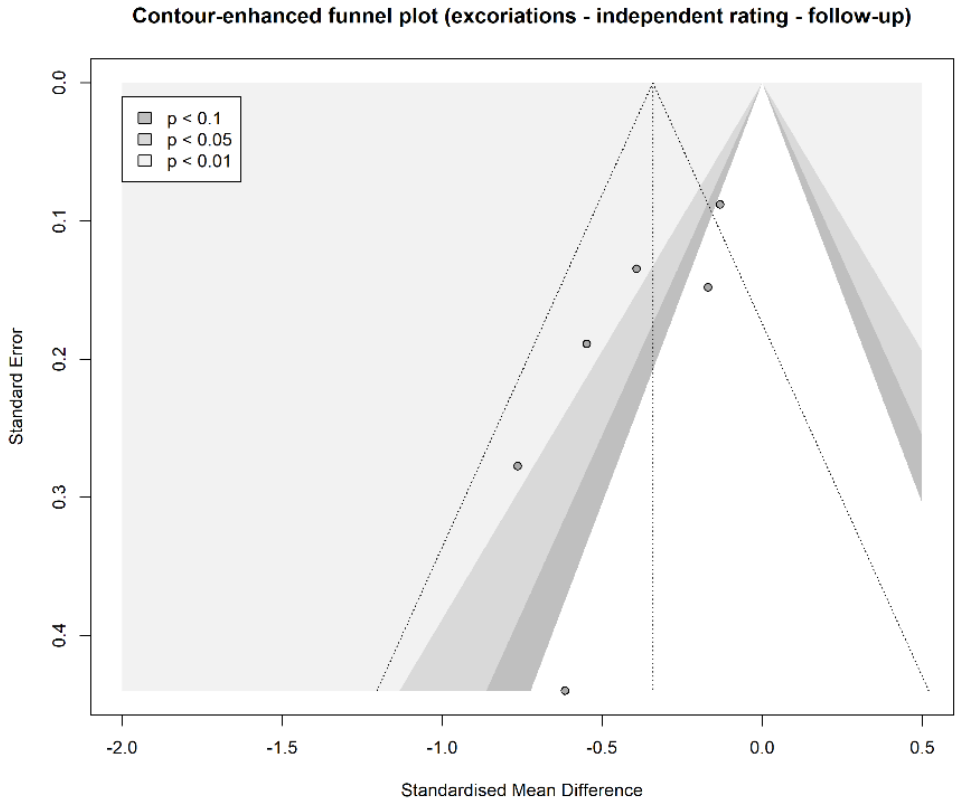


Note: The observed p-curve includes 4 statistically significant ($p < .05$) results, of which 4 are $p < .025$.
There were 2 additional results entered but excluded from p curve because they were $p > .05$.

(2) Follow-Up

(b) Externally rated excoriations

Funnel Plot



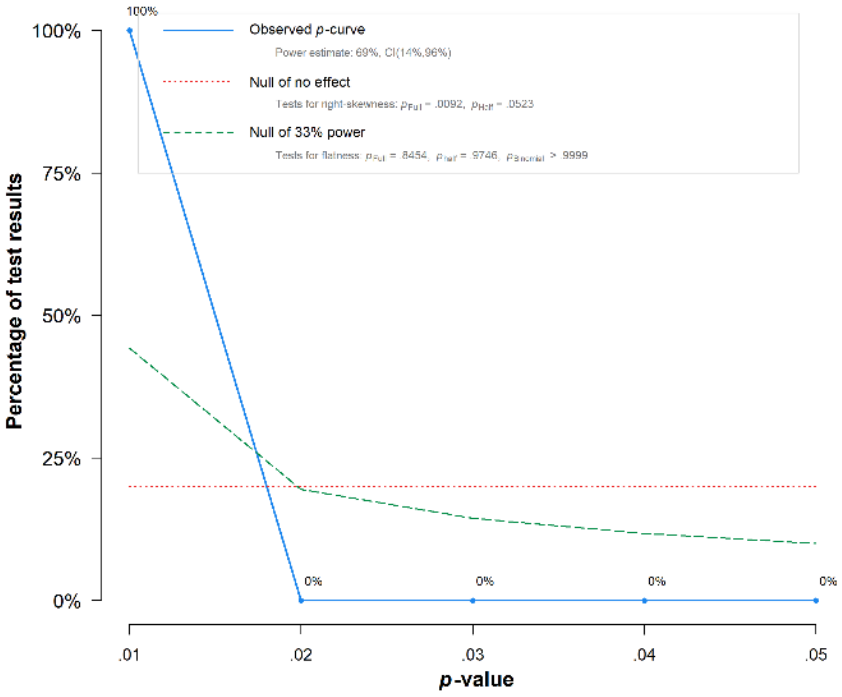
Egger's Test

```
> eggers.test(m.cont.FU.EXCO.extern)
Eggers' test of the intercept
=============================

 intercept        95% CI     t          p
    -2.381 -4.09 - -0.67 -2.73 0.05243358

Eggers' test does not indicate the presence of funnel plot asymmetry.
Warning message:
In eggers.test(m.cont.FU.EXCO.extern) :
  Your meta-analysis contains k = 6 studies. Egger's test may lack the statistical power to detect bias when th
e number of studies is small (i.e., k<10).
```

p-curve



Note: The observed *p*-curve includes 3 statistically significant ($p < .05$) results, of which 3 are $p < .025$.
There were 3 additional results entered but excluded from *p*-curve because they were $p > .05$.