

Psychological treatment for clinically significant health anxiety across medical and psychological services: A systematic review with source-linked reanalysis

LienChung Wei

wlc078@berkeley.edu

Kainan University <https://orcid.org/0000-0002-7383-7309>

Systematic Review

Keywords: health anxiety, hypochondriasis, cognitive behavioural therapy, systematic review, behavioral medicine, evidence synthesis

Posted Date: October 5th, 2026

DOI: <https://doi.org/10.21203/rs.3.rs-11237034/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.
[Read Full License](#)

Additional Declarations: The authors declare no competing interests.

Psychological treatment for clinically significant health anxiety across medical and psychological services: A systematic review with source-linked reanalysis

LienChung Wei, MD, MPH^{1,2}

¹ Department of Health Industry Management, Kainan University, Taoyuan, Taiwan

² Department of Addiction Psychiatry, Taoyuan Psychiatric Center, Ministry of Health and Welfare, Taoyuan, Taiwan

Corresponding author

LienChung Wei

Email: wlc078@berkeley.edu

ORCID: 0000-0002-7383-7309

Psychological treatment for clinically significant health anxiety across medical and psychological services

A systematic review with source-linked reanalysis

Abstract

Psychological treatment for health anxiety is delivered across medical and psychological services, but apparent benefit depends on the comparator, outcome and assessment time. This systematic review examined these distinctions using source-linked reanalysis. PubMed and Europe PMC searches through 22 September 2026 identified 8,799 unique records, supplemented by registry and bounded citation searches. Retrieval and adjudication were closed on 1 October 2026. Forty-three primary efficacy reports were organized into 42 conservative synthesis units; uncertain overlap precluded treating these as 42 independent trials. Forty-five numerical results were reproduced in Stata, with repeated assessments and alternative estimands kept distinct. Selected observed-endpoint contrasts for internet cognitive behavioural therapy ranged from a substantial benefit against a discussion forum (Hedges g -1.60 , 95% CI -2.10 to -1.10) to an imprecise difference against behavioural stress management (-0.26 , -0.58 to 0.06). In medical services, the CHAMP trial reported a 12-month model-based Health Anxiety Inventory difference of -2.97 points (-4.37 to -1.57) against usual care. These examples do not constitute a pooled treatment ranking. Eligible studies without usable outcomes remained in the narrative synthesis. Limited database coverage, unavailable reports, single-author AI-assisted screening and result-specific bias constrain completeness and certainty. Psychological treatments should be interpreted against their actual alternatives and clinical settings; symptom improvement alone does not establish improved functioning or reduced healthcare use.

Keywords: health anxiety; hypochondriasis; cognitive behavioural therapy; systematic review; behavioral medicine; evidence synthesis

Introduction

Health anxiety is encountered where physical symptoms, interpretations of illness and care seeking intersect. A psychological treatment may reduce illness-related distress without eliminating a medical condition, and attending medical care may be appropriate even when health anxiety is elevated. These distinctions matter when transferring evidence between specialist psychological services and medical clinics. CHAMP, for example, tested adapted cognitive behavioural therapy (CBT) in medical outpatients, whereas remotely delivered CBT has also been evaluated among repeated users of unscheduled care (Tyrer et al., 2014; Morriss et al., 2019). Neither setting is interchangeable with recruitment through advertisements or psychological services.

The relevant comparison is equally important. Waiting list, discussion forums, supported psychoeducation, usual medical care, active stress management and face-to-face CBT offer different amounts and kinds of help. Labeling all these conditions “control” can make a treatment appear to have a stable effect when the underlying questions differ. Delivery through the internet also does not define a single intervention: full-course exposure-based CBT and a brief interpretation-training task differ in therapeutic content and intensity (Hedman et al., 2014; Antognelli et al., 2020).

This review examines randomized evidence for structured psychological interventions in clinically significant health anxiety, with particular attention to what can be inferred across medical and psychological services. It combines a systematic evidence account with reproducible, source-linked calculations. Its contribution is to preserve comparator, denominator, measurement and time distinctions that can disappear in a single pooled effect.

A separate contextual stream concerns engagement and care delivery; contextual reports are not used to estimate randomized treatment effects.

Methods

The review is reported with reference to PRISMA 2020, PRISMA-S and synthesis-without-meta-analysis guidance (Page et al., 2021; Rethlefsen et al., 2021; Campbell et al., 2020). A dated completion protocol was recorded on 22 September 2026 after exploratory extraction and some results were known; it was not prospectively registered. The protocol, amendments, searches and report-level dispositions accompany the supplementary materials. This retrospective sequence limits protection against outcome-driven analytical choices. The search cutoff was 22 September and the retrieval/adjudication cutoff was 1 October 2026.

Eligibility and evidence streams

Parallel-group randomized studies were eligible when participants had diagnosed hypochondriasis or illness anxiety, or an explicitly elevated entry threshold on a validated general health-anxiety measure. An incidental health-anxiety outcome did not establish an eligible population. Baseline eligible subgroups within broader trials were retained even when subgroup outcomes could not be separated. Adult-service cohorts permitting ages 16 or 17 were labeled mixed-age; an adult mean age did not establish exclusively adult enrollment. Pediatric-only samples were excluded. Samples of unclear age remained distinguishable. Eligible interventions included CBT, cognitive and exposure therapies, acceptance and commitment therapy (ACT), mindfulness, structured psychoeducation and self-help, and brief therapeutic components. Comparators were classified by their actual content. Cluster allocation and therapy-group dependence were retained as design features. Validated health-anxiety severity was the primary numerical outcome; custom ratings, response, function, healthcare use, acceptability and harms were separate outcomes. Absence of usable outcome

data was not an exclusion criterion. No language, publication-status, open-access or lower-date restriction was specified.

Primary service, observational and qualitative studies were retained in a separate contextual stream. The original first-pass rules excluded individual clinical case reports; later source retrieval added a small historical illustration set. Consequently, case-report coverage is not exhaustive. Contextual and illustrative evidence informs interpretation of delivery and engagement, not comparative efficacy.

Searching and selection

PubMed and Europe PMC were searched with broad population terms, including the Hypochondriasis Medical Subject Heading in PubMed, without requiring intervention or trial terms. Identifier-conservative deduplication retained source membership. Supplementary work comprised ClinicalTrials.gov and ISRCTN searches, backward tracing from three reviews and a bounded forward-citation search from three selected trial reports. Exact executed searches and retrieval dates are supplied in Supplement 1. CENTRAL results were viewed but a complete verified export and screening were not obtained. Completed PsycINFO, Embase and CINAHL searches are not claimed.

Selection used a single-author workflow assisted by OpenAI Codex. The initial title/publication-type pass classified 6,571 records as exclusions; subsequent record assessment considered abstracts, bibliographic evidence and obtainable full reports. Separate AI passes were used for source checks and discrepancy detection but are not independent duplicate human screening. Retrieval depth varied from metadata and archive searches to publisher, repository and institutional-library access. At the cutoff, unavailable reports and retrieved reports with unresolved eligibility remained explicitly labeled rather than being converted into exclusions. Supplementary-search records are reported separately from database records to avoid double counting.

Extraction and synthesis

Reports were linked by registration identifiers, recruitment, allocation, sites, dates and explicit cross-references. The extraction preserved scale version, scoring direction, timing anchor, missing-data method and randomized, observed, imputed or model-analysis denominators. Unreported group sizes were not replaced by equal-allocation assumptions. Estimated marginal means and standard errors were not converted to raw means and standard deviations. Sources unavailable beyond the abstract could establish eligibility when sufficient information was explicit, but did not supply invented outcomes.

The primary synthesis is organized by intervention, comparator and service setting. All retained efficacy units have a report-level account in Supplement 1, including studies without a computable effect. No omnibus psychological-treatment effect or treatment ranking was estimated. Numerical reanalysis was restricted to source-supported inputs; it is therefore a subset of the evidence account, not the full inclusion set. Clinical and methodological differences, and missing usable outcomes, prevented a defensible comprehensive pooled comparison.

The intended main timing target was end of assigned treatment before crossover. A dated amendment distinguished follow-up windows above zero through three months, above three through nine months, and above nine through fifteen months, targeting one, six and twelve months, respectively. Actual time anchors are retained: six months after randomization is not six months after treatment. Multiple times, scales and alternative analyses from the same participants do not provide independent replications. Adult-only and other comprehensive review-wide sensitivity analyses were not performed.

Statistical methods

StataNow/BE 19.5 was used for the numerical analyses and figures. Raw standardized endpoints used meta esize with the exact Hedges small-sample correction and its native

variance estimator. Negative effects favor the first-listed intervention. Raw mean differences used Welch t confidence intervals. Published adjusted/model estimates and their original confidence bounds were retained as reported, including rounding asymmetry; reconstructed approximate standard errors did not replace the reported intervals. Related eligible active arms were combined with within- and between-arm variability before use against a shared control.

Historical CBT–waiting-list analyses were retained only as exploratory methodological sensitivities in Supplement 2. Four overlapping subsets compared the earlier approximate-Hedges convention with native exact-Hedges analyses using DerSimonian–Laird and restricted maximum-likelihood heterogeneity estimators. Pooled intervals used a modified Knapp–Hartung multiplier no smaller than one and $k-1$ degrees of freedom. Study membership, replacement of alternative missing-data summaries and uncorrected partial nesting are explicit. These post hoc subsets do not represent all eligible CBT trials, and no method was selected because it produced statistical significance. No small-study-effect test was applied because no sufficiently comparable independent set contained ten trials.

Independent Python implementations checked calculations against Stata using the same source-linked inputs. The portable Stata package was rerun and compared with the validated outputs. This verifies computational reproduction, not completeness of selection or correctness of every source judgment. Figures 1 and 2 show distinct comparisons without a pooled diamond.

Bias certainty and AI assistance

Result-specific appraisals used RoB 2, with attention to cluster allocation where relevant (Sterne et al., 2019). Signaling responses, source locations and missing-information decisions were preserved. Appraisals were not transferred automatically between observed endpoints, imputed or model estimates, or different follow-up times. Participant awareness was

considered together with outcome assessment and likely influence, rather than treating nonblinding alone as proof of bias.

Ten finite comparison profiles used GRADE considerations for specified outcomes and time points (Guyatt et al., 2008). They do not grade entire treatment classes or compensate for unavailable evidence. Inconsistency was not estimable for a single-study profile. No automatic publication-bias downgrade was made solely for a small study count, but missing reports and restricted source coverage limit every profile. A review-wide certainty grade is not assigned to exploratory pooled subsets.

OpenAI Codex assisted searching, source organization and inspection, extraction checks, preliminary appraisal, Stata programming, independent numerical checks, drafting and reference formatting during September–October 2026. This substantive use extended beyond language editing. The exact underlying model identifier was not consistently recorded across sessions. Multiple AI passes are not human inter-rater reliability, and the single author retains responsibility for the submitted work.

Results

Evidence account and availability

The database searches retrieved 8,152 PubMed and 7,673 Europe PMC records, with 8,799 unique records after deduplication; 647 were unique to Europe PMC. The record ledger comprises 6,571 first-pass exclusions, 1,698 subsequent exclusions, 36 primary efficacy reports, 65 supporting reports, 75 contextual reports, three protocols and 351 records not definitively classified. Of the last group, 349 target reports were not retrieved through the documented routes and two retrieved reports had unresolved population eligibility. These are record dispositions, not independent-study or full-text-exclusion counts. The supplementary flow account retains the additional-source searches and unresolved citation leads separately.

After integrating additional sources and report linkage, 43 primary efficacy reports were organized into 42 conservative synthesis units. This is not a claim of 42 verified independent trials: two Shabahang reports have unresolved participant overlap and are conservatively grouped, and allocation at course level is uncertain in Buwalda's trial. Weck's earlier mechanistic report is linked to the later trial and does not add another trial. The 45 numerical rows comprise 27 named units plus the 2021 Shabahang report, including repeated times and alternative estimands. Fourteen other named units remain without a current Stata result row; author-reported findings may still be described. Supplement 1 maps all units to sources, eligibility and analytical status.

Unavailable evidence is consequential rather than merely administrative. Some retained trials could establish eligibility from an abstract but lacked verifiable outcome data, as in explanatory therapy for hypochondriasis (Fava et al., 2000). Other potentially relevant records remained unclassified. The available synthesis therefore cannot establish the magnitude or direction of effects in all eligible but unretrieved evidence.

Comparator differences in internet treatment

Four trials illustrate why internet treatment should not be treated as a uniform comparison. Guided internet CBT against a moderated discussion forum had a reconstructed observed-endpoint Hedges g of -1.60 (95% CI -2.10 to -1.10), whereas internet exposure-based CBT against internet behavioural stress management had $g=-0.26$ (-0.58 to 0.06) (Hedman et al., 2011; Hedman et al., 2014). The latter is a raw endpoint comparison and does not replace the investigators' longitudinal analysis. Internet CBT against supported psychoeducation and monitoring had a published standardized contrast of -1.39 (-1.93 to -0.87), retained with its original interval and model context (Newby et al., 2018). These results are not a statistical test of effect modification by comparator.

The internet-versus-face-to-face CBT observed endpoint was close to zero, $g=0.07$ (-0.21 to 0.35), with 97 observed participants per arm from 102 randomized per arm (Axelsson et al., 2020). This superiority-style endpoint reconstruction contains no noninferiority margin; it cannot reproduce the original trial's noninferiority conclusion or establish equivalence.

Figure 1 places distinct endpoint comparisons alongside a brief interpretation-training trial to make their differences visible. It is an illustrative display, not a complete network or treatment ranking.

Treatment in medical services

CHAMP randomized 444 medical outpatients and reported HAI model differences of -4.83 (95% CI -6.21 to -3.45) at six months and -2.97 (-4.37 to -1.57) at twelve months against usual care (Tyrer et al., 2014; Tyrer, Salkovskis, et al., 2017). The cohort permitted ages 16–75 and is labeled mixed-age adult-service evidence. Remote CBT in repeat users of unscheduled care yielded source-reported six- and twelve-month HAI contrasts of -2.81 (-5.11 to -0.50) (Morriss et al., 2019). Figure 2 preserves the two trial families and repeated randomization/baseline-anchored times. Neither comparison alone establishes reduced service use, cost equivalence or improved quality of life.

A smaller genitourinary-clinic study provides raw available-case effects at three, six and twelve months; the six-month estimate was $g=-1.07$ (-1.70 to -0.44), based on 20 and 23 observed participants from 23 and 26 randomized (Seivewright et al., 2008). Missing data and unmasked self-ratings limit interpretation. Variable treatment duration and boosters mean the last assessment is not a uniform treatment-free durability test.

Other medical-service evidence could not be reduced to a verified raw standardized endpoint. Practice-suite allocation, adjusted means, an unavailable primary Whiteley result and a control-mean inconsistency require preservation of the original design and prevent simple individual-level reconstruction in the Barsky reports (Barsky & Ahern, 2004; Barsky et al.,

2013). The severe-COVID-anxiety remote-CBT pilot has follow-up dispersion information but unverified per-measure outcome denominators (Crawford et al., 2024; King et al., 2025). Baseline eligible subgroups in IDRIS, a Halifax somatic-symptom trial and COPIC remain included without substituting the whole-trial result for an unavailable health-anxiety subgroup contrast.

Other therapies components and unfavorable findings

Evidence extends beyond CBT. Group ACT against waiting list plus usual care and internet ACT against a discussion forum reported six-month post-treatment Whiteley-7 differences of -20.5 (-29.4 to -11.7) and -19.0 (-27.2 to -10.8), respectively (Eilenberg et al., 2016; Hoffmann et al., 2021). Their comparators, treatment durations and randomization-relative assessment times differ, so they remain separate. Mindfulness-based cognitive therapy, combined mindfulness/ACT in caregivers and Benson relaxation were also retained; scale components, observed versus randomized denominators, and endpoint versus change-score estimands are specified in the supplement. A favorable interval from one estimand and an interval crossing zero from another do not themselves demonstrate a difference between estimands.

Brief component trials address narrower questions. Kerstner's two-week attribution intervention had a raw modified-SHAI effect of $g=-0.60$ (-1.03 to -0.17), before optional subsequent CBT (Kerstner et al., 2015). Interpretation training in Antognelli's study had $g=-0.25$ (-0.69 to 0.18) at treatment end (Antognelli et al., 2020). Other component studies include attention modification, text reminders, single-session anxiety-sensitivity training and worry postponement. Generic worry measures do not replace a health-anxiety outcome; conditional dispersion assumptions remain sensitivity analyses. Shared-arm comparisons and later assessments are not independent trials.

Findings that fail to support the intervention are retained. The one-month anxiety-sensitivity-training model contrast was -1.89 points (-4.65 to 0.87) (O'Bryan, 2019; O'Bryan et al., 2021). Schwind's external-versus-body-attention endpoint difference largely mirrored a baseline gap, while the source health-anxiety interaction was nonsignificant (Schwind et al., 2016). A preprint comparing prescribed health-related with leisure videos alongside identical medication reported worse week-four SHAI values with health-related content; undefined dispersion and completer selection prevented a verified uncertainty estimate. Its high-risk results do not establish medication efficacy (Liao et al., 2026).

The small metacognitive-therapy pilot produced a very large waiting-list endpoint effect, $g=-3.21$ (-4.51 to -1.91), in 20 participants (Bailey & Wells, 2024). Its size, unmasked ratings and lack of an active comparator limit comparative interpretation; later treated-only follow-up cannot establish controlled maintenance. CBT comparisons against medication, combined treatment, psychodynamic therapy, problem-solving psychoeducation and other active therapies likewise remain distinct. Fallon's trial did not establish CBT superiority over placebo on the secondary continuous Whiteley result, while primary responder outcomes answer a different question (Fallon et al., 2017). Supplement 1 accounts for these findings and all numerically unavailable studies without selecting only favorable outcomes.

Exploratory analyses appraisal and engagement

In the historical three-unit CBT–waiting-list subset, exact-Hedges restricted maximum-likelihood pooling gave $g=-1.02$ (95% modified Knapp–Hartung CI -1.89 to -0.15), whereas DerSimonian–Laird gave -1.07 (-2.17 to 0.03). These overlapping analyses are methodological sensitivities, not competing confirmatory answers. The supplement identifies all memberships, including a combined active-arm contrast containing printed bibliotherapy and a broader sensitivity with uncorrected partial nesting. Statistical significance is not a sound basis for selecting among these outputs.

The appraisal archive contains 57 result-specific assessments. Forty of the 45 numerical rows have exact result mappings; five alternate-estimand or follow-up rows have only limited mapping and are explicitly marked descriptive/sensitivity results. Common concerns concern missing outcomes, subjective ratings with allocation awareness and incomplete prespecification. The ten finite certainty profiles range from very low to moderate; their scope is the specified comparison and result, not a therapy-wide endorsement. The CHAMP twelve-month symptom result was rated moderate, whereas the displayed internet comparisons were rated low. Unavailable evidence and incomplete specialist-database coverage remain additional constraints on the body of evidence.

Contextual observations illustrate why access and engagement need their own denominators. In Avia's nonrandomized educational intervention, later treatment attendance differed from initial allocation and some candidates were referred for individual treatment (Avia et al., 1996). Townend's nurse-delivered pilot had three treated patients after one dropout; it cannot establish component equivalence (Townend, 1994). Interviews with six self-selected, previously treated women described feeling unheard, mismatch with treatment focus and the importance of a warm therapeutic relationship (Papis, 2015). These illustrative sources neither estimate prevalence nor show that a treatment generally fails. The full contextual inventory is separate from the efficacy synthesis.

Discussion

The principal finding is that clinical interpretation changes when the actual comparator and estimand are preserved. Available evidence supports symptom benefit in several specific psychological-treatment comparisons, including medical-service CBT and some internet interventions. It does not support treating every control condition as interchangeable, ranking treatments from isolated effect sizes, or inferring service savings from a symptom score. This

distinction is especially important where psychological care is introduced within ongoing medical treatment.

The reanalysis also shows why an apparently simple effect-size calculation requires clinical and design information. Observed endpoints, imputed summaries, change scores and longitudinal models differ in participants and assumptions. A model standard error does not become a raw standard deviation merely by multiplying by a sample-size term. Likewise, delayed treatment, shared controls, repeated follow-up and therapy-group nesting can inflate apparent information if treated as independent data. Source-linked reporting makes these decisions inspectable without claiming that more computation repairs an inadequate source. For practice, the question is what additional help a defined intervention offers over the care a patient can actually receive. Discussion forums, supported information, medication management and face-to-face psychotherapy are meaningful alternatives, not empty labels. Treatment selection should also consider patient goals, functioning, tolerability, access and ongoing medical responsibilities. The present review cannot identify a universal best format or an optimal service pathway; the contextual observations suggest issues to examine rather than tested implementation solutions.

Future trials would improve synthesis by reporting randomized and analyzed denominators for each outcome, scale versions and anchors, prespecified estimands, and outcomes for eligible baseline subgroups. Function, healthcare use, acceptability and adverse outcomes need explicit definitions rather than inference from symptom change. For mixed-age services, age-specific reporting would improve applicability. Shared intervention groups, clustered delivery and transitions after waiting-list periods should be clear enough to prevent accidental duplication.

Limitations

This evidence account has important limitations. PubMed and Europe PMC overlap heavily, and specialist psychological databases were not comprehensively searched. Registry and citation work was bounded; some citation leads and 351 database records remained without a definitive eligibility decision at the retrieval cutoff. Unavailable potentially relevant studies may alter the findings. These limitations preclude describing the review as an exhaustive census or interpreting an absence of extracted effects as an absence of evidence.

AI assistance increased the scale of source inspection but did not provide independent human duplicate screening, extraction or appraisal. The first-pass exclusions were not all independently re-adjudicated. A retrospectively dated protocol and post hoc analytical subsets introduce risks of selection and analytical flexibility. Some findings depend on abstracts, indexed publisher tables, a thesis or a preprint, and source access was unequal. Report overlap is not fully resolved for the Shabahang pair. These uncertainties are documented rather than resolved through assumptions.

The numerical set is selective by data usability. Raw endpoint reanalysis may discard valid covariate adjustment and cannot reproduce every original analysis. Five alternate results lack a fully matched appraisal, and certainty profiles address only ten finite comparisons. Broad generalizations about durability, harms, remission, functioning or healthcare use are therefore unwarranted. Further retrieval, specialist-database searching and independent human verification would materially strengthen a subsequent update.

Conclusions

Psychological treatment for clinically significant health anxiety should be interpreted in relation to the actual alternative, population, measurement and assessment time. Source-linked reanalysis supports several specific symptom comparisons while retaining unfavorable, unavailable and analytically uncertain findings. Reproducible Stata calculations make the

numerical decisions inspectable; they do not eliminate limitations in retrieval, study design or review conduct. Claims about sustained benefit and service impact require outcomes that directly measure those questions.

Declarations

Funding: No specific funding was received for this review. Competing interests: The author declares no competing interests.

Ethics: This review used published and publicly accessible aggregate research reports and did not enroll participants or generate individual-level clinical data; ethics approval and participant consent were not applicable.

Data availability: Extracted aggregate inputs, report-level dispositions, result-specific assessments, supplementary analyses and Stata programs are supplied with this article. The cited source publications remain subject to their publishers' access conditions. No individual participant dataset was obtained and no public repository deposition is claimed.

Author contribution: The single author defined the review question and interpretation, directed the AI-assisted workflow, and takes responsibility for the submitted manuscript. The substantive uses and limitations of AI assistance are described in Methods.

References

- Antognelli, S. L., Sharrock, M. J., & Newby, J. M. (2020). A randomised controlled trial of computerised interpretation bias modification for health anxiety. *Journal of Behavior Therapy and Experimental Psychiatry*, 66, Article 101518.
<https://doi.org/10.1016/j.jbtep.2019.101518>
- Avia, M. D., Ruiz, M. A., Olivares, M. E., Crespo, M., Guisado, A. B., Sánchez, A., & Varela, A. (1996). The meaning of psychological symptoms: Effectiveness of a group

- intervention with hypochondriacal patients. *Behaviour Research and Therapy*, 34(1), 23–31. [https://doi.org/10.1016/0005-7967\(95\)00052-y](https://doi.org/10.1016/0005-7967(95)00052-y)
- Axelsson, E., Andersson, E., Ljótsson, B., Björkander, D., Hedman-Lagerlöf, M., & Hedman-Lagerlöf, E. (2020). Effect of internet vs face-to-face cognitive behavior therapy for health anxiety: A randomized noninferiority clinical trial. *JAMA Psychiatry*, 77(9), 915–924. <https://doi.org/10.1001/jamapsychiatry.2020.0940>
- Bailey, R., & Wells, A. (2024). Feasibility and preliminary efficacy of metacognitive therapy for health anxiety: A pilot RCT. *Journal of Affective Disorders Reports*, 16, Article 100751. <https://doi.org/10.1016/j.jadr.2024.100751>
- Barsky, A. J., & Ahern, D. K. (2004). Cognitive behavior therapy for hypochondriasis: A randomized controlled trial. *JAMA*, 291(12), 1464–1470. <https://doi.org/10.1001/jama.291.12.1464>
- Barsky, A. J., Ahern, D. K., Bauer, M. R., Nolido, N., & Orav, E. J. (2013). A randomized trial of treatments for high-utilizing somatizing patients. *Journal of General Internal Medicine*, 28(11), 1396–1404. <https://doi.org/10.1007/s11606-013-2392-6>
- Campbell, M., McKenzie, J. E., Sowden, A., Katikireddi, S. V., Brennan, S. E., Ellis, S., Hartmann-Boyce, J., Ryan, R., Shepperd, S., Thomas, J., Welch, V., & Thomson, H. (2020). Synthesis without meta-analysis (SWiM) in systematic reviews: Reporting guideline. *BMJ*, 368, Article l6890. <https://doi.org/10.1136/bmj.l6890>
- Crawford, M. J., King, J. D., McQuaid, A., Bassett, P., Leeson, V. C., Tella, O., Di Simplicio, M., Tyrer, P., Tyrer, H., Watt, R. G., & Barnicot, K. (2024). Severe COVID anxiety among adults in the United Kingdom: Cohort study and nested feasibility trial. *BMC Psychiatry*, 24(1), Article 27. <https://doi.org/10.1186/s12888-023-05446-9>

- Eilenberg, T., Fink, P., Jensen, J. S., Rief, W., & Frosthalm, L. (2016). Acceptance and commitment group therapy (ACT-G) for health anxiety: A randomized controlled trial. *Psychological Medicine*, 46(1), 103–115. <https://doi.org/10.1017/s0033291715001579>
- Fallon, B. A., Ahern, D. K., Pavlicova, M., Slavov, I., Skritskya, N., & Barsky, A. J. (2017). A randomized controlled trial of medication and cognitive-behavioral therapy for hypochondriasis. *American Journal of Psychiatry*, 174(8), 756–764. <https://doi.org/10.1176/appi.ajp.2017.16020189>
- Fava, G. A., Grandi, S., Rafanelli, C., Fabbri, S., & Cazzaro, M. (2000). Explanatory therapy in hypochondriasis. *The Journal of Clinical Psychiatry*, 61(4), 317–322. <https://doi.org/10.4088/jcp.v61n0414>
- Guyatt, G. H., Oxman, A. D., Vist, G. E., Kunz, R., Falck-Ytter, Y., Alonso-Coello, P., Schünemann, H. J., & GRADE Working Group. (2008). GRADE: An emerging consensus on rating quality of evidence and strength of recommendations. *BMJ*, 336(7650), 924–926. <https://doi.org/10.1136/bmj.39489.470347.ad>
- Hedman, E., Andersson, G., Andersson, E., Ljótsson, B., Rück, C., Asmundson, G. J. G., & Lindefors, N. (2011). Internet-based cognitive-behavioural therapy for severe health anxiety: Randomised controlled trial. *The British Journal of Psychiatry*, 198(3), 230–236. <https://doi.org/10.1192/bjp.bp.110.086843>
- Hedman, E., Axelsson, E., Görling, A., Ritzman, C., Ronnheden, M., El Alaoui, S., Andersson, E., Lekander, M., & Ljótsson, B. (2014). Internet-delivered exposure-based cognitive-behavioural therapy and behavioural stress management for severe health anxiety: Randomised controlled trial. *British Journal of Psychiatry*, 205(4), 307–314. <https://doi.org/10.1192/bjp.bp.113.140913>
- Hoffmann, D., Rask, C. U., Hedman-Lagerlöf, E., Jensen, J. S., & Frosthalm, L. (2021). Efficacy of internet-delivered acceptance and commitment therapy for severe health

- anxiety: Results from a randomized, controlled trial. *Psychological Medicine*, 51(15), 2685–2695. <https://doi.org/10.1017/s0033291720001312>
- Kerstner, T., Witthöft, M., Mier, D., Diener, C., Rist, F., & Bailer, J. (2015). A diary-based modification of symptom attributions in pathological health anxiety: Effects on symptom report and cognitive biases. *Journal of Consulting and Clinical Psychology*, 83(3), 578–589. <https://doi.org/10.1037/a0039056>
- King, J. D., McQuaid, A., Barnicot, K., Bassett, P., Leeson, V. C., Di Simplicio, M., Tyrer, P., Tyrer, H., Watt, R. G., & Crawford, M. J. (2025). Trajectory of severe COVID anxiety and predictors for recovery in an 18-month UK cohort. *BJPsych Open*, 11(6), Article e263. <https://doi.org/10.1192/bjo.2025.10884>
- Liao, Y., Wang, W., Fang, Y., Zhang, W., Song, S., & Zhang, K. (2026). Impact of short videos on agomelatine and pregabalin efficacy for illness anxiety disorder: A randomized controlled trial [Preprint]. Research Square. <https://doi.org/10.21203/rs.3.rs-9566804/v1>
- Morriss, R., Patel, S., Malins, S., Guo, B., Higton, F., James, M., Wu, M., Brown, P., Boycott, N., Kaylor-Hughes, C., Morris, M., Rowley, E., Simpson, J., Smart, D., Stubley, M., Kai, J., & Tyrer, H. (2019). Clinical and economic outcomes of remotely delivered cognitive behaviour therapy versus treatment as usual for repeat unscheduled care users with severe health anxiety: A multicentre randomised controlled trial. *BMC Medicine*, 17(1), Article 16. <https://doi.org/10.1186/s12916-019-1253-5>
- Newby, J. M., Smith, J., Uppal, S., Mason, E., Mahoney, A. E. J., & Andrews, G. (2018). Internet-based cognitive behavioral therapy versus psychoeducation control for illness anxiety disorder and somatic symptom disorder: A randomized controlled trial. *Journal of Consulting and Clinical Psychology*, 86(1), 89–98. <https://doi.org/10.1037/ccp0000248>

- O'Bryan, E. M. (2019). A randomized controlled trial evaluating the efficacy of a brief computerized anxiety sensitivity reduction intervention for health anxiety [Doctoral dissertation, University of Cincinnati]. OhioLINK.
https://etd.ohiolink.edu/acprod/odb_etd/ws/send_file/send?accession=ucin1572879397495098&disposition=inline
- O'Bryan, E. M., McLeish, A. C., Norr, A. M., Ely, S., Bass, Z., Davies, C. D., Capron, D. W., Schmidt, N. B., & Mano, K. E. J. (2021). A randomized controlled trial evaluating the efficacy of a brief computerized anxiety sensitivity reduction intervention for health anxiety. *Journal of Anxiety Disorders*, 82, Article 102425.
<https://doi.org/10.1016/j.janxdis.2021.102425>
- Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., Shamseer, L., Tetzlaff, J. M., Akl, E. A., Brennan, S. E., Chou, R., Glanville, J., Grimshaw, J. M., Hróbjartsson, A., Lalu, M. M., Li, T., Loder, E. W., Mayo-Wilson, E., McDonald, S., ... Moher, D. (2021). The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *BMJ*, 372, Article n71. <https://doi.org/10.1136/bmj.n71>
- Papis, K. G. (2015). Emotional intelligence and sociotropy-autonomy in young women with DSM-IV-TR hypochondriasis: A mixed-method study [Doctoral dissertation, University of Wolverhampton]. Wolverhampton Intellectual Repository and E-Theses.
<https://hdl.handle.net/2436/582290>
- Rethlefsen, M. L., Kirtley, S., Waffenschmidt, S., Ayala, A. P., Moher, D., Page, M. J., Koffel, J. B., & PRISMA-S Group. (2021). PRISMA-S: An extension to the PRISMA statement for reporting literature searches in systematic reviews. *Systematic Reviews*, 10(1), Article 39. <https://doi.org/10.1186/s13643-020-01542-z>

- Schwind, J., Gropalis, M., Witthöft, M., & Weck, F. (2016). The effects of attention training on health anxiety: An experimental investigation. *Cognitive Therapy and Research*, 40(2), 245–255. <https://doi.org/10.1007/s10608-015-9745-x>
- Seivewright, H., Green, J., Salkovskis, P., Barrett, B., Nur, U., & Tyrer, P. (2008). Cognitive-behavioural therapy for health anxiety in a genitourinary medicine clinic: Randomised controlled trial. *British Journal of Psychiatry*, 193(4), 332–337. <https://doi.org/10.1192/bjp.bp.108.052936>
- Sterne, J. A. C., Savović, J., Page, M. J., Elbers, R. G., Blencowe, N. S., Boutron, I., Cates, C. J., Cheng, H.-Y., Corbett, M. S., Eldridge, S. M., Emberson, J. R., Hernán, M. A., Hopewell, S., Hróbjartsson, A., Junqueira, D. R., Jüni, P., Kirkham, J. J., Lasserson, T., Li, T., ... Higgins, J. P. T. (2019). RoB 2: A revised tool for assessing risk of bias in randomised trials. *BMJ*, 366, Article 14898. <https://doi.org/10.1136/bmj.14898>
- Townend, M. (1994). The efficacy of cognitive-behavioural psychotherapy for hypochondriasis: A pilot investigation. *Journal of Psychiatric and Mental Health Nursing*, 1(3), 185–186. <https://doi.org/10.1111/j.1365-2850.1994.tb00043.x>
- Tyrer, P., Cooper, S., Salkovskis, P., Tyrer, H., Crawford, M., Byford, S., Dupont, S., Finnis, S., Green, J., McLaren, E., Murphy, D., Reid, S., Smith, G., Wang, D., Warwick, H., Petkova, H., & Barrett, B. (2014). Clinical and cost-effectiveness of cognitive behaviour therapy for health anxiety in medical patients: A multicentre randomised controlled trial. *The Lancet*, 383(9913), 219–225. [https://doi.org/10.1016/s0140-6736\(13\)61905-4](https://doi.org/10.1016/s0140-6736(13)61905-4)
- Tyrer, P., Salkovskis, P., Tyrer, H., Wang, D., Crawford, M. J., Dupont, S., Cooper, S., Green, J., Murphy, D., Smith, G., Bhogal, S., Nourmand, S., Lazarevic, V., Loeberberg, G., Evered, R., Kings, S., McNulty, A., Lisseman-Stones, Y., McAllister, S., ... Barrett, B. (2017). Cognitive-behaviour therapy for health anxiety in medical patients (CHAMP): A

randomised controlled trial with outcomes to 5 years. *Health Technology Assessment*, 21(50), 1–58. <https://doi.org/10.3310/hta21500>

Figure 1. Selected comparator-specific raw endpoint reanalyses. Negative Hedges g favors the first-listed intervention. These source-linked examples are not a pooled treatment ranking or a complete evidence set. The active-comparator endpoint analyses do not replace original longitudinal or noninferiority analyses. Bars show 95% confidence intervals.

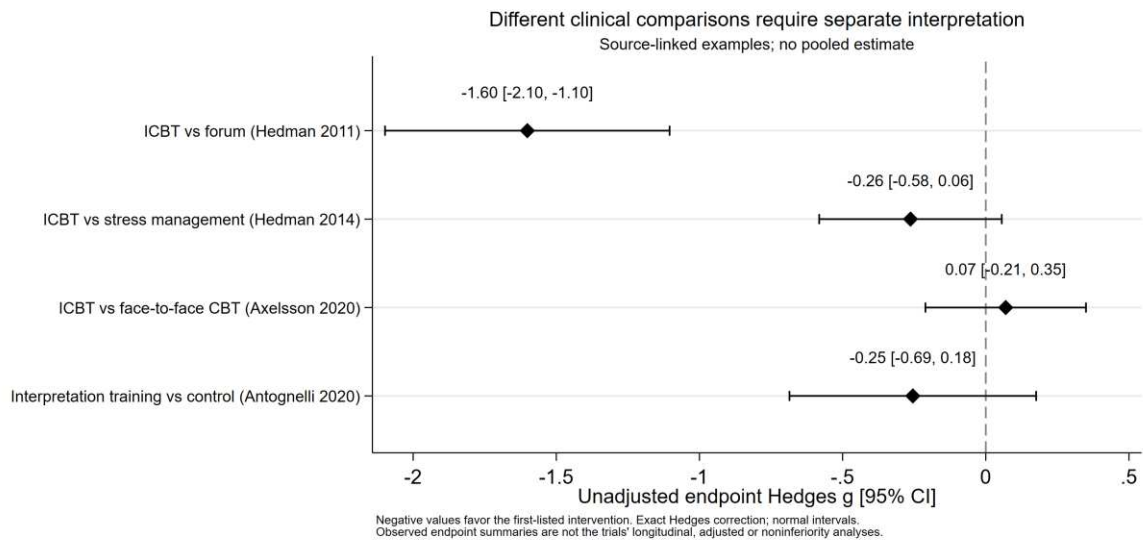
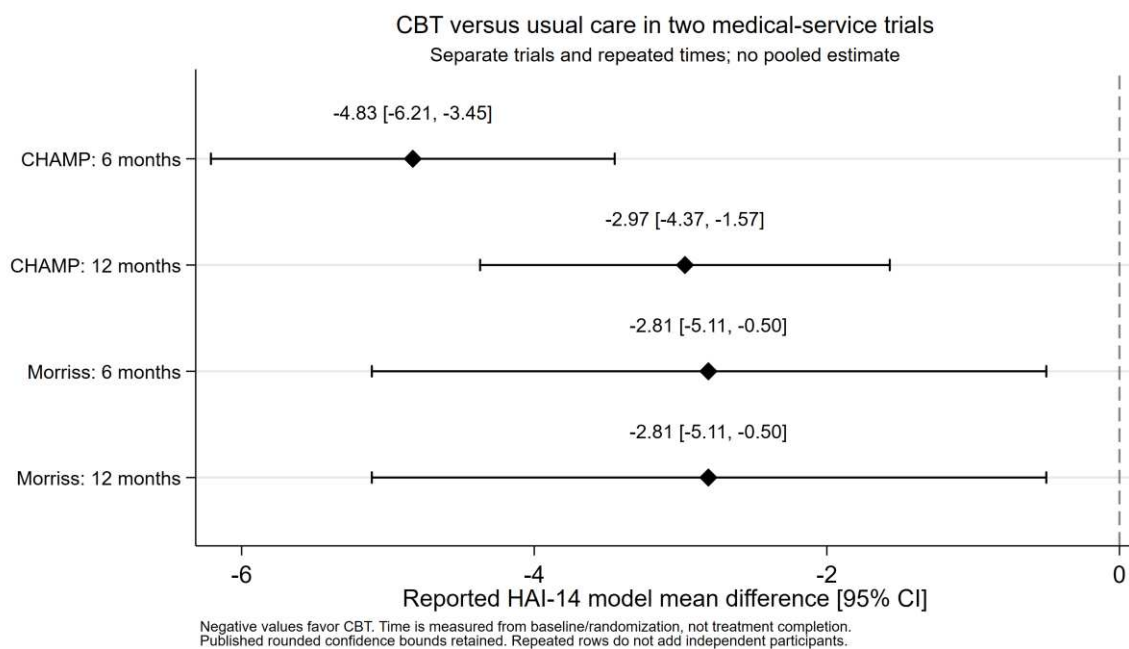


Figure 2. Source-reported HAI model contrasts in two medical-service trials. The six- and twelve-month rows are repeated results from the same respective trial, not four independent studies. Time is anchored to baseline/randomization rather than treatment completion. Negative values favor CBT. Original confidence bounds are retained. Morriss reports the same model contrast at six and twelve months; the repeated rows are not independent confirmations.



Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [Supplement1Evidence.docx](#)
- [Supplement2DataandStata.zip](#)
- [PRISMAChecklist.docx](#)