

<https://doi.org/10.1038/s44271-026-00414-x>

# Psychosocial interventions indicate prolonged survival in cancer patients in a systematic review, meta-analysis, and multiverse meta-analysis of randomized controlled trials

Check for updates

A list of authors and their affiliations appears at the end of the paper

Whether psychosocial interventions containing active psychological components prolong survival in cancer patients has been studied for decades, yet findings from primary (RCTs) and secondary research (meta-analyses) remain inconclusive. Our preregistered systematic review, meta-analysis, and multiverse meta-analysis aimed to clarify this research question using contemporary methods of research synthesis. We searched Web of Science, Scopus, MEDLINE, PsycINFO, Embase, and Google Scholar for RCTs of structured psychosocial interventions for cancer patients published until October 17, 2025. We calculated the overall effect; assessed its robustness; estimated a median survival benefit, characterized the psychological components included in the interventions; examined risk of bias, study quality, meta-analytic post hoc power, and sponsorship bias; explored 25 substantive and methodological moderators; and considered publication bias as well as *p*-hacking. Using multiverse meta-analysis, we calculated multiple overall effects based on reasonable specifications employed in prior meta-analyses (descriptive analysis) and compared them with the 95% CI band of 1000 simulated overall effects assuming no true effect (inferential analysis). Psychosocial interventions, provided alongside medical treatment, demonstrated a small, positive and robust overall effect on survival in cancer patients, with an HR of 0.80, 95% CI [0.71, 0.90] across 32 RCTs comprising 5704 participants. Heterogeneity was moderate to substantial with an  $I^2 = 48\%$  and a wide 95% PI (HR 0.49–1.29). Median survival time benefit was estimated at 3.9 months, 95% CI [–0.7, 8.5], based on data from 16 trials. The psychological components most frequently applied were educational, cognitive–behavioral, emotionally expressive, and group-based social support. Low average meta-analytic post hoc power (17%) likely contributed to inconsistent findings among trials. Multiverse meta-analysis confirmed the presence of a general overall survival effect and indicated that previously conflicting meta-analytic conclusions primarily stemmed from differences in effect size metrics and analytic decisions. Psychosocial (psychological) interventions appear to improve survival in cancer patients, with effect sizes comparable in magnitude to effects previously reported in the literature for medical cancer treatments such as chemo-, radio-, and hormone therapy. The certainty of evidence was rated moderate, primarily due to statistical heterogeneity, hence effects might not generalize equally to all populations. Considering survival impact, established psychological benefits,

✉ e-mail: [kenji.asakawa.haas@gmail.com](mailto:kenji.asakawa.haas@gmail.com); [ulrich.tran@univie.ac.at](mailto:ulrich.tran@univie.ac.at)

favorable safety profile, and comparatively low cost, the findings support a paradigm shift toward establishing psychosocial interventions alongside medical therapy as a standard component of comprehensive cancer care; potentially guiding future research and clinical practice.

Each year, ~19.3 million individuals develop cancer, and about 10 million patients die from the disease<sup>1</sup>. Psychological risk factors for cancer development and progression, such as stress, anxiety, depression, anger, bereavement, repression of emotions, helplessness, and hopelessness, are already well known<sup>2-7</sup>.

Psychosocial interventions are provided to help cancer patients overcome the initial shock of diagnosis, difficulties in coping with the illness, associated mental health problems, and irritating side effects of medical treatment through cognitive-behavioral, supportive-expressive, or mindfulness-based techniques<sup>8,9</sup>.

Additionally, they appear to alleviate psychological and physiological risk factors for cancer progression<sup>10</sup> and have been shown to reduce stress in cancer patients<sup>11</sup>, a major factor for the onset, development, and recurrence of cancer<sup>5,6,12-15</sup>, which extends far beyond the benefits of palliative care alone<sup>16</sup>.

Whether psychosocial interventions are similarly able to prolong the survival of cancer patients has been hotly debated since the landmark study by Spiegel et al.<sup>17</sup>, which sparked much research in the field of psycho-oncology. To date, almost half of the respective randomized controlled trials (RCTs) have found a life-extending effect, whereas the remainder have not (see Fig. 1A, B for a timeline with published RCTs and meta-analyses)—including three failed replication attempts of the initial positive finding<sup>18-20</sup>.

Likewise, ten prior meta-analyses (see Table 1) have resulted in similar contradictory conclusions. Four meta-analyses yielded a significant overall effect (at least at one of the examined follow-ups), whereas six did not (see Fig. 1A, B).

Heterogeneity in the results of primary (RCTs) as well as secondary research (reviews and meta-analyses) has fueled controversial debates and led to the formation of two opposing camps. One group argues that the evidence for psychosocial interventions enhancing the survival of cancer patients is convincing<sup>9,21-25</sup>, whereas the other argues it is not<sup>26-36</sup>.

In an attempt to investigate possible causes of the ongoing controversy, we have identified a number of factors that could explain the ambiguous findings of prior meta-analytic research and believe that a modern and comprehensive synthesis of the existing evidence remains necessary.

An overview of the most relevant issues identified in prior meta-analytic research is provided in the following paragraphs addressing the number of included trials, the choice of effect size measures, the investigation of risk of bias factors within studies, the handling of clinical and statistical heterogeneity, the examination of reporting bias across studies, the need for a comprehensible effect size measure for practitioners and patients, and concerns regarding data-analytic decisions in prior meta-analyses (see Table 1).

### Possible causes for the contradictory findings in prior meta-analyses

**Number and selection of included trials.** A meta-analysis is considered to produce the highest form of scientific evidence<sup>37</sup> and its execution therefore requires quite delicate handling<sup>38</sup>. Prior meta-analyses on the topic (see Table 1) followed the current gold standard in evidence-based research and analyzed RCTs only<sup>39</sup>, whereas the most recent attempt<sup>40</sup> applied even stricter criteria and focused exclusively on RCTs that used passive controls and an intention-to-treat (ITT) analysis.

The application of selective inclusion criteria might serve the aim of increasing the validity of the meta-analytic overall effect<sup>41</sup>. However, in our view, the feasibility of such an approach depends on the amount of available evidence and, on top of that, may address only indirectly some of the risks of bias within studies (e.g., study quality) and reporting bias across studies (e.g., publication bias) that could affect the meta-analytic result. Further, selective

inclusion criteria might even allow investigators to exploit researcher degrees of freedom resulting in the mass production of redundant meta-analyses that often contain only fragments of the totality of evidence<sup>42</sup>.

Moreover, the first meta-analysis on this topic was published almost two decades ago<sup>43</sup>, when only a few RCTs were available. As a consequence, prior meta-analyses (see Table 1) frequently involved only a handful of trials and none of them more than 15 RCTs.

Hence, the statistical power of the prior meta-analyses to identify small-to-moderate overall effects might have been insufficient<sup>38</sup>. In addition, a trend can be observed that overall effects tended to be not significant in earlier meta-analyses, but significant in more recent meta-analyses (see Fig. 1A, B). Thus, updating this line of meta-analytic research, without additional restrictions, still appears to be required to adequately address the pending research question of whether psychosocial interventions may prolong the survival of cancer patients or not.

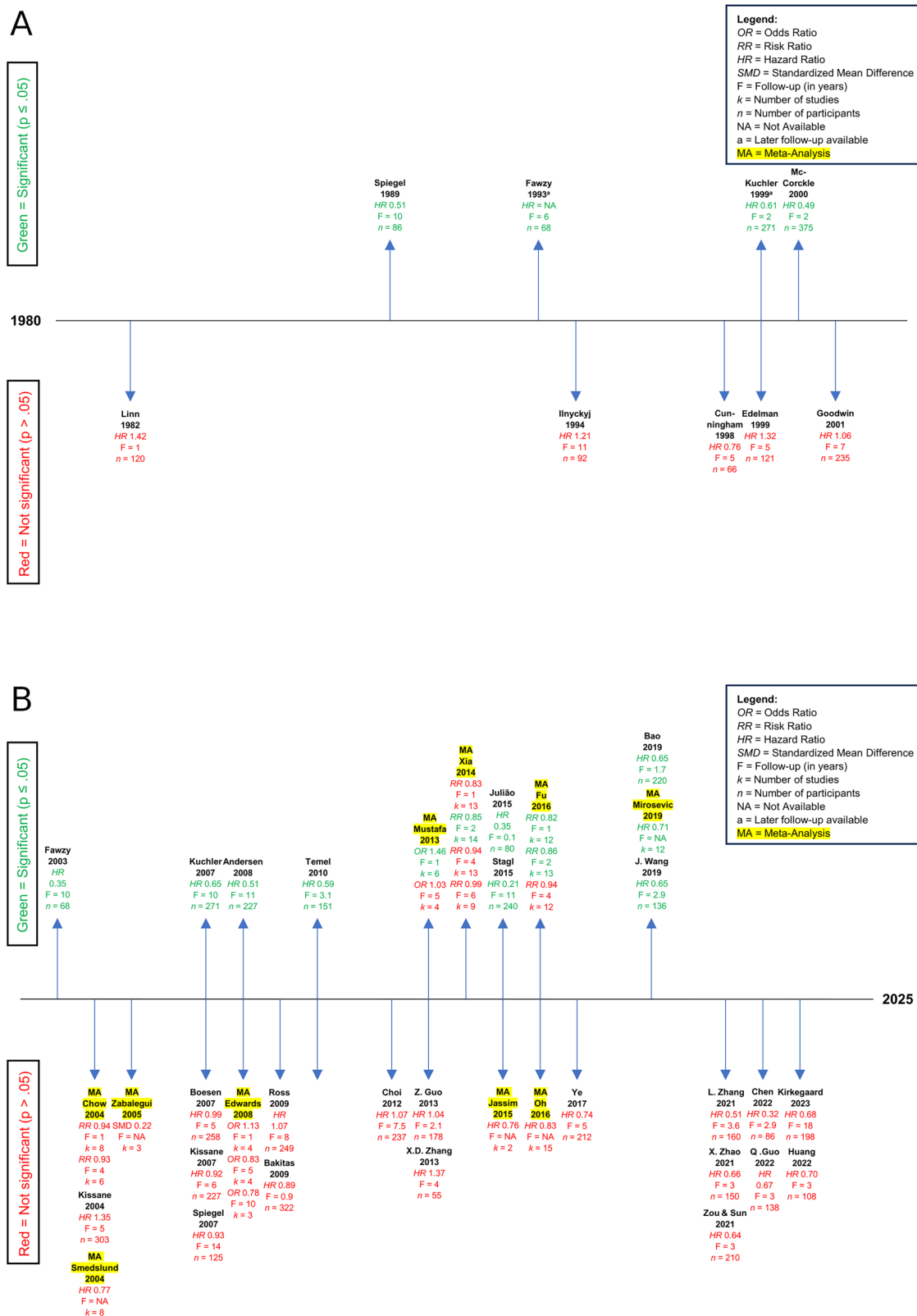
**Choice of the effect size measure.** Prior meta-analyses differed in their use of effect size measures: odds ratio (OR) vs. risk ratio (RR) vs. hazard ratio (HR). Yet, in the context of survival data, the use of the HR is preferable over the OR and the RR, as it considers information of time-to-event data from all participants, including those who were lost during the trial<sup>44</sup>. Moreover, the HR considers the whole study period and does not allow study authors or meta-analysts to selectively pick one (or multiple) of many possible follow-up intervals as endpoints (e.g., at 1, 5, or 10 years), which can be readily done with the OR and the RR (see Table 1).

Hence, the HR helps to measure survival more precisely and safeguards against possible questionable research practices (i.e., the exploitation of researcher degrees of freedom to achieve a desired result<sup>45</sup>). However, only four of the prior meta-analyses have used the HR, whereas the other six have used either the OR or RR (see Table 1).

**Risk of bias within studies.** Risk of bias in primary research represents a threat to the validity of science such as selection bias (sample not representative of the target population, due to non-randomization or high dropout rates), design bias (inherent flaws in the study design, such as additional interventions or inadequate control groups), data analysis bias (manipulation of data analysis until statistically significant results are found), and financial or other conflicts of interest (researchers' personal interests influencing the outcome)<sup>46</sup>.

Common threats to the validity of meta-analytic research that might have a profound influence on the overall effect are: (1) the methodological and statistical quality of the included studies, which is a crucial factor for the overall quality of a meta-analysis<sup>47</sup>; (2) the meta-analytic post hoc power of included trials (i.e., the probability that a study can detect the observed overall effect of a meta-analysis), which is critical for the validity of meta-analytic hypothesis testing<sup>46,48,49</sup>; (3) the risk of sponsorship bias (i.e., the distortion of study effects by financial interests), which is important for a precise and unbiased overall effect estimate<sup>50,51</sup>, as well as (4) the multiplicity of analyses (i.e., conducting numerous analyses) and (5) the flexibility in analyses (i.e., adjusting analytical methods based on results), which both increase the risk of finding significant results by chance<sup>52,53</sup>.

Five out of ten prior meta-analyses collected data with the Risk-of-Bias Tool 1<sup>54</sup>, seven out of the ten prior meta-analyses presented descriptive data on study quality, but only two have evaluated its impact on survival by excluding low-quality studies in the form of a sensitivity analysis (see Table 1). However, none of the meta-analyses reviewed here have investigated (meta-analytic) post hoc power or risk of sponsorship bias. It follows that these risk of bias factors still need to be examined systematically in this field of research.



**Fig. 1 | Significance-based timeline of published RCTs and meta-analyses on psychosocial interventions and survival in cancer patients. A** Timeline covering studies published from 1980 to 2002. **B** Timeline covering studies published from 2002 to 2025. Note that Fawzy et al., 2003 and Kuchler et al., 2007 were follow-up publications.

**Clinical and statistical heterogeneity.** The meta-analytic overall effect is meaningful if the included studies are sufficiently homogenous, whereas it can be misleading in the face of high clinical (substantive and methodological) or high statistical heterogeneity<sup>55,56</sup>. In this sense, substantive moderators (e.g., age, sex, cancer type, tumor stage, or

hormone receptor status<sup>20,57,58</sup>) and methodological moderators (e.g., intervention type, nature of control group, type of statistical analysis, or length of follow-up interval<sup>57,58</sup>) might exert an effect on overall survival time and should therefore be investigated in a thorough moderator analysis.

**Table 1 | List of the ten prior meta-analyses that examined the effects of psychosocial interventions on survival in cancer patients (data refers to survival analysis)**

First author and year of publication	Number of included RCTs	Overall effect at last follow-up [95% CI]	p value	Follow-up intervals (years)	Statistical heterogeneity (I <sup>2</sup> )	Substantive moderators analyzed	Methodological moderators analyzed	Statistical tests missing	Risk of bias (RoB) and study quality assessment	Reporting bias assessment
Chow et al. 2004 <sup>127</sup>	8, 6	RR = 0.93 [0.77, 1.13]	0.50	1, 4	I <sup>2</sup> = NA	None	None	None	None	None
Smedslund et al. 2004 <sup>69</sup>	8	HR = 0.77 [0.56, 1.06]	0.10	Pooled	I <sup>2</sup> = 74.90%	None	Study design, intervention format	TfSD	None	Funnel plot
Zabalegui et al. 2005 <sup>174</sup>	3	SMD = 0.22 [-0.24, 0.68]	0.40	Pooled	I <sup>2</sup> = NA	None	None	None	None	None
Edwards et al. 2008 <sup>75</sup>	4, 4, 3	OR = 0.78 [0.37, 1.64]	0.51	1, 5, 10	I <sup>2</sup> = 35.00%	None	Intervention type	TfSD	Study quality <sup>c</sup>	None
Mustafa et al. 2013 <sup>76</sup>	6, 4	OR = 1.03 [0.42, 2.52]	0.95	1, 5	I <sup>2</sup> = 70.00%	None	Intervention type <sup>a</sup>	None	RoB, study quality <sup>c</sup>	None
Xia et al. 2014 <sup>177</sup>	13, 14, 13, 9	RR = 0.99 [0.86, 1.15]	0.90	1, 2, 4, 6	I <sup>2</sup> = 70.00%	None	None	None	Study quality <sup>c</sup> , sensitivity analysis	Funnel plot
Jassim et al. 2015 <sup>78</sup>	2	HR = 0.76 [0.25, 2.32]	0.63	Pooled	I <sup>2</sup> = 84.00%	None	None	None	RoB, study quality <sup>c</sup>	None
Fu et al. 2016 <sup>56</sup>	12, 13, 12	RR = 0.94 [0.85, 1.04]	0.24	1, 2, 4	I <sup>2</sup> = 34.00%	None	Intervention format	TfSD	RoB, study quality <sup>c</sup>	None
Oh et al. 2016 <sup>68</sup>	15	HR = 0.83 [0.68, 1.01]	0.06	Pooled	I <sup>2</sup> = 64.00%	Metastasis <sup>a</sup> , cancer type <sup>a</sup>	Intervention type/format <sup>a</sup> /provider <sup>a</sup> /dose <sup>a</sup> , follow-up <sup>a</sup>	TfSD (except intervention type)	RoB, study quality <sup>c</sup> , sensitivity analysis	Funnel plot, trim-and-fill, fail-safe N, Egger's regression test
Mirosevic et al. 2019 <sup>9</sup>	12	HR = 0.71 [0.58, 0.88]	0.001	Pooled	I <sup>2</sup> = 62.60%	Marital status, tumor stage, age	Cancer type <sup>b</sup> , intervention type <sup>b</sup> /format <sup>b</sup> /dose <sup>b</sup> /period <sup>b</sup>	None	RoB, study quality <sup>c</sup>	Funnel plot

HR/OR/RR < 1 and a positive SMD indicate a survival benefit of the intervention; HR/OR/RR > 1 and a negative SMD indicate a survival disadvantage; HR/OR/RR = 1 or SMD = 0 indicates no effect. HR hazard ratio, OR odds ratio, RR risk ratio, SMD standardized mean difference, RCTs randomized controlled trials, 95% CI 95% confidence interval, I<sup>2</sup> proportion of true variance to total variance, NA not available, TfSD test for subgroup differences.

<sup>a</sup>Subgroups with overlapping 95% CIs.

<sup>b</sup>Subgroup summary effects not reported.

<sup>c</sup>Only descriptive information, not statistically analyzed.

**Table 2 | Classification and components of each type of psychosocial intervention**

Type of intervention	Specific components
Cognitive-behavioral	Target psychoeducation, goal setting, problem-solving, cognitive reappraisal, relaxation techniques, mindfulness training, coping skills training, stress management, and activity planning
Meaning-enhancing	Emphasize meaning-making and/or expression of disease-related emotions and feelings
Dignity therapy, life review, and narrative-driven	Give patients the opportunity to reflect on their own life and its personal meaning
Other counseling	Counseling that does not fit into the other categories (e.g., individually tailored approaches to meet specific needs)
Education-only	Aim to transmit knowledge about illness and symptoms, to improve communication between patients and physicians, and aid with decision-making and symptom management
Musical therapy, writing therapy, and others	Interventions that are mainly based on music therapy or expressive writing and other types of social support (e.g., web-based)

Categories and components were drawn from ref. 59.

Note that for a long time the term “psychosocial intervention” was rather vaguely defined, but finally has been described and grouped into six distinct categories by Teo et al.<sup>59</sup> with respect to cancer patients: (1) cognitive-behavioral; (2) meaning-enhancing; (3) dignity therapy, life review, and narrative-driven; (4) other counseling; (5) education-only; and (6) musical therapy, writing therapy, and others (see Table 2 for a detailed description of these categories). Examining the previous literature regarding the topic, the impression may arise that the goal of the ongoing debate was to (generally) declare all types of psychosocial interventions as either effective or ineffective in prolonging cancer survival. However, one should consider that specific types of interventions may work well in certain populations, whereas in others they may not<sup>21</sup>. Therefore, further investigation in this area could prove beneficial.

Possible causes of heterogeneity should be investigated not only carefully but also in a systematic manner that examines all potentially relevant moderators, along with those interactions between moderators that are deemed important and make the most sense<sup>60</sup>.

In addition, machine-learning and bootstrap-based approaches may help to identify the most relevant moderators in an exploratory fashion, especially when the number of moderators is larger than the number of included studies, thus providing results which then can be matched to (or contrasted against) the results of conventional moderator analysis<sup>61</sup>.

Clinical heterogeneity was clearly evident in the corpus of primary studies (RCTs), and statistical heterogeneity was identified in eight of the 10 prior meta-analyses (see Table 1). Yet, prior meta-analyses often did not sufficiently investigate sources of heterogeneity. Six meta-analyses investigated categorical moderators in subgroup analyses. However, only two meta-analyses inspected more than two moderators and only two applied statistical tests to explore actual subgroup differences (see Table 1).

In some cases, continuous moderators were transformed into categories by dividing studies into groups with low, middle, and high values of a given characteristic. While this approach may aid interpretability, it can also introduce issues concerning the justification of cutoff values and may increase false-positive rates<sup>62,63</sup>. Furthermore, none of the prior meta-analyses controlled for possible type-1 errors in the moderator analysis. In sum, this casts doubt on the evidentiary value of some of the reported moderators regarding survival. Consequently, a broad range of possible moderators still needs to be investigated systematically.

**Reporting bias across studies.** The ten prior meta-analyses did not apply proper (or any) contemporary methods to examine publication bias and *p*-hacking (see Table 1). Publication bias may result in the distortion of the meta-analytic overall effect, due to a higher likelihood of primary studies being published if they report significant effects<sup>64,65</sup>, whereas *p*-hacking denotes the manipulation of data analysis with the aim to produce a significant result<sup>66,67</sup>.

Publication bias was investigated in five of the 10 prior meta-analyses. Four of them used graphical methods (funnel plots), while only one<sup>68</sup>

applied statistical tests to search for evidence of publication bias (see Table 1); however, using a method (Egger’s regression test<sup>69</sup>) for which an enhanced extension (FAT-PET-PEESE<sup>70</sup>) is available and a procedure (Rosenthal’s fail-safe  $N^{71}$ ), which is not considered valid anymore<sup>64,65</sup>. This meta-analysis was also the only one that tried to adjust the overall effect for publication bias, but yet again, with a method (Trim-and-Fill<sup>72</sup>) that is not deemed appropriate anymore<sup>64,66</sup>. None of the prior meta-analyses checked for the evidence of *p*-hacking (see Table 1). It follows that the totality of evidence still needs to be tested for the presence of reporting bias with contemporary methods.

**Data-analytic decisions.** The visualization of study overlap for all meta-analyses published so far (including the present meta-analysis) with the GROOVE Tool<sup>73</sup> resulted in a high overlap of evidence (see Fig. 2 for a matrix of evidence of GROOVE), as the overall corrected covered area adjusted by structural zeros was 43% (see Supplementary Fig. A1, A2 for overall results of GROOVE). From these figures, it can be concluded that some meta-analyses are redundant, whereas others are very selective. Hence, it seemed self-evident to examine the complexity of prior meta-analytic decisions and their effects on the outcomes of these meta-analyses.

In general, there are many ways to analyze data (various researcher degrees of freedom), which in some cases may yield broadly similar results, whereas in other cases vastly different ones. Accordingly, the question of *which* data should be analyzed (the data universe) and *how* these data should be analyzed (the model universe) have become more and more relevant in recent years, for primary studies<sup>41,74,75</sup> as well as meta-analysis<sup>76</sup>. Four of the ten prior meta-analyses reported a significant overall effect, whereas six did not (see Table 1). These meta-analyses differed in numerous ways concerning the data that were analyzed (e.g., at times a focus on metastatic or non-metastatic female breast cancer patients) and how these data were analyzed (e.g., survival analysis after 1, 2, 4, 5, 6, or 10 years of follow-up).

The specification curve and multiverse-analysis approach to meta-analysis<sup>76,77</sup> can provide insights into the causes of heterogeneity in the results of past meta-analyses and specifically enables evaluation of the generalizability (vs. narrowness) and robustness (vs. fragility) of findings across all different meta-analytic specifications. It does so by: (1) identifying reasonable specifications of (past) meta-analyses concerning which data were meta-analyzed (WHICH factors; data universe) and how these data were meta-analyzed (HOW factors; model universe); (2) calculating the overall effects of all of these specifications and their combinations; (3) presenting the results graphically with descriptive (meta-analytic) specification curve plots (which display the meta-analytic overall effect of each specification, in an ascending order, together with the number of included trials) and histograms with the number of corresponding *p* values; and (4) applying parametric bootstrap techniques and presenting inferential (meta-analytic) specification curve plots, which allow one to infer whether observed overall effects from all of these specifications deviate from the 95% confidence interval (CI) bands of simulated true overall effects of zero. Therefore, conducting a specification curve and multiverse meta-analysis is

Primary Studies		Systematic Reviews										
Study ID	Reference	Chow 2004	Smedslund 2004	Zabalegui 2005	Edwards 2008	Mustafa 2013	Xia 2014	Jassim 2015	Fu 2016	Oh 2016	Mirosevic 2019	Asakawa-Haas 2021
1	Linn 1982	1					1					1
2	Spiegel 1989	1	1	1	1	1	1			1	1	1
3	Fawzy 1993	1	X				X			X		X
4	Ilnyckyj 1994	1	1				1			1		1
5	Cunningham 1998	1	1	1	1	1	1			1	1	1
6	Edelman 1999	1	1	1	1	1	1			1	1	1
7	Kuchler 1999	1	1				X		X	X	X	X
8	McCorkle 2000		1				1			1	1	1
9	Goodwin 2001	1	1		1	1	1			1	1	1
10	Fawzy 2003		1				1			1		1
11	Kissane 2004						1	1	1	1		1
12	Boesen 2007	X	X	X			1		1	1		1
13	Kissane 2007	X	X	X		1	1		1	1		1
14	Kuchler 2007	X	X	X			1		1	1	1	1
15	Spiegel 2007	X	X	X			1			1	1	1
16	Andersen 2008	X	X	X			1	1	1	1	1	1
17	Bakitas 2009	X	X	X	X						1	1
18	Ross 2009	X	X	X	X		1		1	1	1	1
19	Temel 2010	X	X	X	X						1	1
20	Choi 2012	X	X	X	X				1	1		1
21	Guo, Z. 2013	X	X	X	X				1			1
22	Zhang, X. D. 2013	X	X	X	X				1			1
23	Juliao 2015	X	X	X	X	X	X					1
24	Stagl 2015	X	X	X	X	X	X				1	1
25	Ye 2017	X	X	X	X	X	X	X	X	X		1
26	Bao 2019	X	X	X	X	X	X	X	X	X		1
27	Wang, J. 2019	X	X	X	X	X	X	X	X	X		1
28	Zhang, L. 2021	X	X	X	X	X	X	X	X	X	X	1
29	Zhao, X. 2021	X	X	X	X	X	X	X	X	X	X	1
30	Zhou & Sun 2021	X	X	X	X	X	X	X	X	X	X	1
31	Chen 2022	X	X	X	X	X	X	X	X	X	X	1
32	Guo, Q. 2022	X	X	X	X	X	X	X	X	X	X	1
33	Huang 2022	X	X	X	X	X	X	X	X	X	X	1
34	Kirkegaard 2023	X	X	X	X	X	X	X	X	X	X	1

**Fig. 2 | Matrix of evidence of the GROOVE tool for all meta-analyses and all included RCTs (note that refs. 182,185 are follow-up studies).** The green cells marked with 1 indicate a primary study which was included in the respective meta-analysis, the white cells indicate a primary study which was not included, and the

black cells marked with X indicate a structural zero (a study that could not have been included either because it was published after the meta-analysis or because a follow-up publication has been incorporated instead).

deemed necessary to resolve the existing ambiguities regarding this research question.

**A tangible effect size for practitioners, patients, and policymakers**

Practitioners, patients, and policymakers may have difficulties interpreting effect sizes such as the HR, OR, or RR, which are frequently used in meta-analyses about survival. Some statistical knowledge about specific cutoff values that allow classification of an overall effect ratio as small, medium, or

large is necessary, while the question remains about what a small, medium, or large effect actually means. Therefore, we also provide a more tangible effect size, the overall median survival (time) difference<sup>78,79</sup>. This is a weighted average of the difference in median survival time between the intervention and control group.

**Objectives of the present analysis**

The purpose of our preregistered systematic review, meta-analysis, and multiverse meta-analysis was to investigate whether psychosocial

**Table 3 | List of general eligibility criteria according to the PICOTS framework**

Categories	Description
General	Articles in English with no restriction on the year of publication
Patient	Male and female patients ( $\geq 18$ years) with a primary (tumor) or secondary cancer (metastasis) diagnosis of any type and any stage, but without a diagnosis of a mental disorder
Intervention	Some form of psychosocial or psychotherapeutic intervention
Comparison	Active (placebo intervention) or passive controls (no intervention)
Outcome	HRs of all-cause survival or all-cause mortality from primary studies, prior meta-analyses, or Kaplan–Meier survival curves from which an effect size can be estimated
Timing	Any length of intervention or follow-up interval
Study Type	RCTs analyzed prospectively or retrospectively

Eligibility criteria listed according to the PICOTS framework (ref. 80).

interventions prolong the survival of cancer patients, addressing all of the above issues. Therefore, we (1) systematically searched for eligible RCTs, calculated HRs where necessary, and estimated an overall effect HR based on the currently available corpus of primary studies; (2) estimated the overall median survival difference between the intervention and control groups to provide a more tangible effect size; (3) applied sensitivity analyses for both effect size measures; (4) investigated risk of bias, study quality, meta-analytic post hoc power, and risk of sponsorship bias within primary studies to assess the validity of included RCTs; (5) inspected clinical and statistical heterogeneity with a systematic approach to moderator analysis that examined all previously explored moderators (identified from included trials, prior meta-analyses, and other relevant literature), while controlling for type-1 errors and also drawing on a machine-learning approach; (6) examined publication bias and *p*-hacking across primary studies in order to assess the evidential value of the meta-analytic findings; (7) planned to estimate a meta-analytic overall effect adjusted for publication bias (in case the appropriate conditions would be met); (8) explored possible causes for heterogeneity in the results of prior meta-analyses; (9) checked for the existence of true (nonzero) overall effects across reasonable specifications by specification curve and multiverse meta-analysis, and (10) evaluated the certainty of all the gathered evidence.

Consequently, we present a comprehensive analysis of the effects of psychosocial interventions on survival in cancer patients, designed to clarify sources of ambiguity in both primary and prior secondary research findings and contribute to a more coherent scientific understanding.

## Methods

### Study eligibility criteria

General eligibility criteria of primary studies are listed according to the patient, intervention, comparison, outcome, timing, and study type (PICOTS<sup>80</sup>) framework (see Table 3 for a list of general eligibility criteria).

### Specific inclusion criteria

Studies were required to focus on psychosocial interventions designed to improve the mental, emotional, and social well-being of cancer patients. Eligible interventions comprised structured approaches containing active psychological components, such as psychoeducation, mindfulness-based interventions, cognitive-behavioral therapy (CBT), emotional-expressive therapy, or peer support. These interventions primarily targeted mental health outcomes, including anxiety, depression, and distress. Included trials could involve palliative care, provided that psychosocial interventions constituted a central element of the study and were administered to both the intervention as well as the control group. General recommendations on lifestyle factors (e.g., nutrition and exercise) were acceptable, but systematic lifestyle interventions were not allowed to be a core component, particularly when they were provided exclusively to the intervention group.

### Specific exclusion criteria

Trials were excluded if they focused primarily on palliative care, symptom control (e.g., pain or nausea management), or end-of-life care without a structured psychosocial intervention. Studies that lacked defined psychological components (e.g., general discussions without therapeutic techniques or unstructured emotional support) were also excluded. Research integrating systematic lifestyle interventions (e.g., structured dietary changes or exercise programs) was excluded to ensure that psychosocial effects could be isolated from other factors, especially when administration differed between the intervention and the control groups.

### Registration and protocol

This meta-analysis and multiverse meta-analysis was preregistered on October 18, 2020 at the Open Science Framework (<https://osf.io/4knzu>) after an initial literature search but before any analysis of the current dataset took place (registration number 10.17605/OSF.IO/4KNZU). The article follows the PRISMA 2020 reporting guidelines<sup>81</sup>. The methodology for our analysis was refined throughout the course of the project as a more suitable approach was identified in line with the guidelines of good scientific practice. Hence, the final form deviates from the preregistration in the following ways:

First, exporting citations from the database CENTRAL is no longer free (since 2019), so instead, Web of Science, Scopus, and Embase were added to the literature search. Second, the database OpenGrey closed in 2020, so we used ProQuest Dissertations & Theses Global as an alternative. Third, effect sizes (HRs) were calculated using the approach of ref. 82, which is a more precise method than that of ref. 83—especially when some information on censoring is available. Fourth, standard forest plots were used instead of rainforest plots to present the results, as the R package *Metaviz*<sup>84</sup> did not allow the presentation of subgroup summary effects calculated with pooled  $\tau^2$  estimates.

Fifth, the R package *MetaForest* was used to explore moderators instead of *Meta-CART*<sup>85</sup>, as it had a higher power to detect effects, offered better protection against overfitting, and captured interactions more effectively. Sixth, the additional FE moderator analysis was dropped, as *MetaForest* proved to be a far more powerful tool to identify possible moderators in an exploratory manner. Seventh, *FAT-PET-PEESE* was added to test for publication bias in an exploratory manner. Eighth, the overall effect was planned to be adjusted using the best appropriate method according to the *Meta Showdown Explorer* instead of predetermining it beforehand.

### Ethical considerations

Ethics approval was not required, as this study analyzed only previously published trials with existing ethics approval and informed consent, in accordance with the University of Vienna Ethics Committee policy. No individual patient data were collected.

### Information sources

Electronic databases such as Web of Science, Scopus, MEDLINE (via PubMed), PsycINFO (via EBSCO), Embase, and Google Scholar (first 500 results for RCTs, first 100 results for meta-analyses) were searched for eligible RCTs and prior meta-analyses on October 18, 2020 and were regularly updated until October 17, 2025. A similar combination of these databases has been proposed to be the most promising for systematic literature searches<sup>86</sup>. In addition, the reference lists of all eligible RCTs and prior meta-analyses that dealt with the same topic were hand-searched. On top of that, databases *ClinicalTrials.gov* and *ProQuest Dissertations & Theses Global* were searched for gray literature.

### Literature search

The main search terms were (psychotherapy OR psychosocial) AND (cancer OR neoplasm) AND (survival OR mortality) AND “randomized controlled trial” (for the search of RCTs) or meta-analysis (for the search of meta-analyses; see Supplementary Tables A1–A3 for full lists of search strings).

### Selection process of primary studies

The literature search and study screening were conducted independently by two authors (KDAH, LB), with disagreements resolved through discussion and consensus. At first, all identified articles were imported into a reference management program (EndNote 21), and duplicates were removed. Next, titles and abstracts were screened for eligibility. Finally, full-text versions of seemingly eligible trials were obtained and evaluated for final inclusion.

### Data collection process

A data extraction sheet was created that contained all variables of interest. It was pilot-tested on ten randomly selected RCTs and refined accordingly. Data were collected independently by two authors (KDAH, LB) and subsequently checked for consistency. Any disagreements were again resolved by discussion and consensus. Seven of 18 contacted authors of eligible RCTs provided more detailed information about the articles of Guo, Z. et al.<sup>87</sup>, Julião et al.<sup>88</sup>, Kirkegaard et al.<sup>89</sup>, Kissane et al.<sup>19,90</sup>, Ross et al.<sup>91</sup>, Spiegel et al.<sup>17,20</sup>, and Temel et al.<sup>92</sup>.

### Data items

Data extracted from eligible RCTs included (1) general study information: first author, year of publication, sample size, number of deaths, name of interventions, and intervention components (e.g., psychoeducation, cognitive restructuring, and expression of emotions); (2) effect sizes: HR (plus 95% CIs) and median survival time (plus minimum, maximum, first and third quartile of survival time in months); (3) moderator variables: age (years), sex (% male), marital status (% married), cancer type (breast, colorectal, lung, melanoma, other, and mixed samples), tumor stage (stage I–IV, mixed samples), tumor phase (early [stage I & II], late [stage III & IV], mixed samples), (distant) metastasis (metastatic [stage IV], non-metastatic [stage I–III], mixed samples), (baseline) estrogen receptor status positive (%), (baseline) progesterone receptor status positive (%), (baseline) lymph node status positive (%), (baseline) surgery before study entry (%), (baseline) chemotherapy before study entry (%), (baseline) radiotherapy before study entry (%), (baseline) hormone therapy before study entry (%), intervention type (CBT-based, meaning-enhancing, narrative-driven, other counseling, education-only, others), intervention format (individual, group), intervention dose (hours), intervention period (weeks), intervention onset (during hospital stay, after hospital stay), background of intervention provider (medical, psychological, social, mixed teams), (significant) differences between groups in attendance of psychosocial interventions outside of study (yes, no), control group type (active: minimal psychosocial input such as health education, relaxation training, or guidance material; passive: treatment as usual without any structured psychosocial input), length of follow-up (months), analysis type (ITT, not ITT), patients lost to intervention (%), and continent of origin (Africa, Asia, Australia, Europe, North America, South America); and (4) risk of bias assessment: risk of bias (low, some concerns, high), study quality (score), meta-analytic post hoc power (%), funding sources (no funding, government funding only, [also] public charity funding involved, [also] private foundation funding involved, [also] for-profit organization funding involved, for-profit organization funding only), and risk of sponsorship bias (low, moderate, high). Note that some categories could not be analyzed due to an insufficient number of studies.

In addition, intervention components were systematically extracted from each included study based on intervention descriptions in the main article and supplementary materials. Each intervention was coded for the presence of psychological and supportive components reported in the articles. Multiple components could be coded per study, reflecting the multicomponent nature of the interventions. Component frequencies were summarized descriptively by absolute and relative occurrence (in significant and non-significant studies). This analysis was exploratory and descriptive in nature, and did not assess causal effects.

### Effect size measures

The proportional hazards assumption in survival analysis posits that the ratio of hazard rates between two groups remains constant over time, which

is important for the calculation of the Cox proportional hazards regression model. Both primary studies (see Table 4) and prior meta-analyses (see Table 1) have adopted this assumption. According to established guidelines, the assumption is likely met if curves do not cross and the distance between them remains relatively constant over time<sup>93</sup>. We assessed its validity by visually inspecting published Kaplan–Meier survival curves from all incorporated RCTs, as individual patient data were unavailable. Our inspection revealed no major violations and minor deviations, observed in a few cases, were not considered sufficient to bias effect size (HR) estimates or the overall conclusions.

Survival in RCTs was measured from trial entry or randomization until death, censoring, or the end of the study. Effect sizes such as HRs (preferably univariate; i.e., without any control for covariates) and median survival time were taken from published articles, communication with the authors, or prior meta-analyses.

However, in case these effect sizes were unavailable, survival probabilities were extracted from published Kaplan–Meier survival curves with GetData Graph Digitizer software 2.26<sup>94</sup>. Data for the HR calculations were reconstructed by dividing published Kaplan–Meier survival curves into appropriate time intervals (e.g., 1-, 2-, or 3-month segments) and extracting information regarding the number of events (i.e., deaths) from cumulative survival probability as well as the number of censored patients (i.e., drop-outs) from tick marks in the survival curves and from information in text, figures, or tables in the articles.

The numbers at risk for each time interval were then calculated using Excel sheets, following a method developed by ref. 82, which makes it possible to incorporate data on censoring while implementing the approach of ref. 83—a widely used method for reconstructing data from Kaplan–Meier survival curves. For transparency and reproducibility, all calculation spreadsheets are publicly accessible (see Supplemental File S1 for HR and death number calculation sheets at <https://osf.io/wnxdb/files>).

The HR for each trial in question was then computed from the number of events and the numbers at risk for each time interval by an online HR calculator<sup>95</sup>. HRs of 0.77, 0.53, and 0.36 represented small, medium, and large effects, respectively (the inverse of HR 1.3, 1.9, and 2.8<sup>96</sup>). HR <1 indicated a survival benefit for the intervention group, HR >1 a survival benefit for the control group, and HR = 1 no effect.

Data extraction and calculation of effect sizes were executed by one author (KDAH) and subsequently checked by another author (UST), who is an active researcher in the fields of quantitative methods and statistics. This independent verification ensured methodological rigor and minimized error risk.

As the analysis is based on reconstructed individual patient data from Kaplan–Meier curves, the estimates are subject to the usual limitations of this approach; however, all steps were performed using validated, peer-reviewed methods to ensure the highest possible accuracy.

### Meta-analytic synthesis methods

All eligible RCTs were included in a single synthesis without further grouping. Statistical analysis was carried out using R 4.3.0 and the R package Metafor 4.0.0<sup>97</sup>. The R code of the meta-analysis and multiverse meta-analysis (see Supplemental File S2 for the data sheet and the R script of the meta-analysis at <https://osf.io/wnxdb/files>) was written and validated by the first author (KDAH) while subsequently cross-checked by a colleague, Johannes K. Vilsmeier, at our institution. HRs were transformed to a log scale (logHR) for calculations and retransformed for presentation<sup>38</sup>.

Clinical (substantive and methodological) heterogeneity among primary studies was accounted for with a random-effects (RE) meta-analytic model using the restricted maximum-likelihood (REML) estimator, as it has the most reasonable properties to estimate the overall effect<sup>38,98,99</sup>.

Statistical heterogeneity (i.e., variation in true population effects) was analyzed with the *Q* test. The relative amount of true (between-study) variance to total (observed) variance  $I^2$ , the absolute amount of true variance  $\tau^2$  (log scale), and the true standard deviation  $\tau$  (log scale) are reported (a simple retransformation on the original scale was not possible with the latter

**Table 4 | Important study characteristics of the 32 included RCTs that investigated psychosocial interventions and survival in cancer patients**

Name and year of publication	Sample size (int./con.)	Mean age (years)	Cancer type (and stage)	Intervention type	Intervention components	Intervention format	Intervention dose (hours)	Intervention provider	Nature of control group	Follow-up (years)	HR and 95% CI
Linn 1992 <sup>125a</sup>	120 (62/58)	58.0	Mixed (Stage IV)	Narrative-driven	ADM, TS, RL, MBC	Individual	NA	Psychological	Passive	1	1.417; 95% CI [0.938, 2.139]
Spiegel 1989 <sup>17</sup>	86 (50/36)	54.7	Breast (Stage IV)	Meaning-enhancing	CT, EE, IC, MBC, GS, HY	Group	78	Mixed	Passive	10	0.512; 95% CI [0.310, 0.835]
Iljnyckij 1994 <sup>179</sup>	92 (61/31)	NA	Mixed (Stage I-IV)	Other counseling	OSS, GS	Group	24	Social	Passive	11	1.213; 95% CI [0.722, 2.039]
Cunningham 1998 <sup>83</sup>	66 (30/36)	50.6	Breast (Stage IV)	CBT-based	GS, CT, RT, GOS	Group	70	Psychological	Active	5	0.764; 95% CI [0.432, 1.353]
Edeiman 1999 <sup>180</sup>	121 (60/61)	49.9	Breast (Stage IV)	CBT-based	CT, CR, IC, GOS, RT, PE, GS	Group	24	Psychological	Passive	5	1.323; 95% CI [0.841, 2.080]
McCorkle 2000 <sup>126</sup>	375 (190/185)	65.0	Mixed (Stage I-IV)	Other counseling	ME, PS, CO	Individual	12	Medical	Passive	2	0.490; 95% CI [0.321, 0.752]
Goodwin 2001 <sup>18</sup>	235 (158/77)	50.2	Breast (Stage IV)	Meaning-enhancing	GS, EE, CT	Group	78	Mixed	Active	7	1.060; 95% CI [0.780, 1.450]
Fawzy 1993 <sup>181</sup> / 2003 <sup>82</sup>	68 (34/34)	42.5	Melanoma (Stage I)	CBT-based	ME, SM, RT, CT, PRS, PS, GS	Group	9	Medical	Passive	10	0.348; 95% CI [0.122, 0.992]
Kissane 2004 <sup>80</sup>	303 (154/149)	46.3	Breast (Stage I-II)	CBT-based	EE, CR, CT, PRS, GOS, GS	Group	30	Mixed	Active	5	1.350; 95% CI [0.760, 2.390]
Boessen 2007 <sup>83</sup>	258 (128/130)	45.9	Melanoma (Stage I)	CBT-based	ME, SM, RT, PRS, CT, GS, PS	Group	12	Mixed	Passive	5	0.990; 95% CI [0.400, 2.600]
Kissane 2007 <sup>19</sup>	227 (147/80)	51.7	Breast (Stage IV)	Meaning-enhancing	ESS, CT, EE, GS	Group	78	Mixed	Active	6	0.920; 95% CI [0.690, 1.240]
Kuchler 1999 <sup>84</sup> / 2007 <sup>85</sup>	271 (136/135)	56.8	Mixed (Stage I-IV)	Other counseling	ME, CS, PS, ESS, EE, GOS, RT, PT	Individual	4	Psychological	Passive	10	0.653; 95% CI [0.495, 0.861]

**Table 4 (continued) | Important study characteristics of the 32 included RCTs that investigated psychosocial interventions and survival in cancer patients**

Name and year of publication	Sample size (int./con.)	Mean age (years)	Cancer type (and stage)	Intervention type	Intervention components	Intervention format	Intervention dose (hours)	Intervention provider	Nature of control group	Follow-up (years)	HR and 95% CI
Spiegel 2007 <sup>30</sup>	125 (64/61)	53.2	Breast (Stage IV)	Meaning-enhancing	EE, IC, ESS, HY, GS	Group	78	Mixed	Active	14	0.930; 95% CI [0.620, 1.400]
Andersen 2008 <sup>186</sup>	227 (114/113)	50.8	Breast (Stage II-III)	CBT-based	RT, PRS, ESS, IC, CT, GS	Group	39	Psychological	Passive	11	0.510; 95% CI [0.280, 0.930]
Bakitas 2009 <sup>87</sup>	322 (161/161)	65.1	Mixed (Stage III-IV)	Other counseling	ME, PRS, ASM, GS	Individual	43	Medical	Passive	0.9	0.889; 95% CI [0.691, 1.144]
Ross 2009 <sup>91</sup>	249 (125/124)	68.5	Colorectal (Dukes Stage A-D)	Other counseling	PS, ME	Individual	15	Medical	Passive	8	1.070; 95% CI [0.770, 1.510]
Temel 2010 <sup>82</sup>	151 (77/74)	64.9	Lung (Stage IV)	Other counseling	GOS, ADM, CR, CS, ME	Individual	4	Medical	Passive	3.1	0.588; 95% CI [0.394, 0.877]
Choi 2012 <sup>127</sup>	237 (118/119)	59.6	Mixed (Stage I-IV)	CBT-based	ME, PRS, IC, CO, ASM, EE	Individual	10	Medical	Passive	7.5	1.070; 95% CI [0.750, 1.530]
Z. Guo 2013 <sup>37a</sup>	178 (89/89)	47.0	Mixed (Stage I-III)	CBT-based	ME, PE, CR, EE, ESS, MBC	Group	10	Medical	Passive	2.1	1.044; 95% CI [0.504, 2.162]
X.D. Zhang 2013 <sup>183a</sup>	55 (27/28)	59.0	Esophageal (Stage I-III)	CBT-based	EE, ME, PS, CT, SM	Individual	10	Medical	Passive	4	1.369; 95% CI [0.592, 3.164]
Julião 2015 <sup>183a</sup>	80 (39/41)	66.1	Mixed (Stage IV)	Narrative-driven	MBC, TS, RL	Individual	3	Medical	Passive	0.1	0.350; 95% CI [0.130, 0.920]
Stegl 2015 <sup>189</sup>	240 (120/120)	50.3	Breast (Stage 0-III)	CBT-based	CR, CT, ESS, RT, GS	Group	15	Psychological	Active	11	0.210; 95% CI [0.050, 0.930]
Ye 2017 <sup>165a</sup>	212 (108/104)	50.2	Breast (Stage IV)	Meaning-enhancing	PE, ME, RT, EE, GS	Group	96	Mixed	Active	5	0.736; 95% CI [0.525, 1.133]
Bao 2019 <sup>193a</sup>	220 (110/110)	43.9	Leukemia (Stage M1-6)	Other counseling	ME, PE, PRS, CS, CO	Individual	27	Medical	Passive	1.7	0.653; 95% CI [0.427, 0.998]

**Table 4 (continued) | Important study characteristics of the 32 included RCTs that investigated psychosocial interventions and survival in cancer patients**

Name and year of publication	Sample size (int./con.)	Mean age (years)	Cancer type (and stage)	Intervention type	Intervention components	Intervention format	Intervention dose (hours)	Intervention provider	Nature of control group	Follow-up (years)	HR and 95% CI
J. Wang 2019 <sup>191a</sup>	136 (68/68)	59.4	Liver (Stage A-B)	Other counseling	ME, PS, EE, CT	Individual	46	Medical	Active	2.9	0.646; 95% CI [0.441, 0.947]
L. Zhang 2021 <sup>192a</sup>	160 (80/80)	59.8	Gastric (Stage I-IV)	Narrative-driven	RM, GS	Group	24	Medical	Active	3.6	0.512; 95% CI [0.172, 1.522]
X. Zhao 2021 <sup>193a</sup>	150 (75/75)	49.8	Brain (Grade I-IV)	Narrative-driven	RM, GS	Group	24	Medical	Active	3	0.664; 95% CI [0.381, 1.156]
Zhou & Sun 2021 <sup>194a</sup>	210 (105/105)	64.3	Colorectal (Stage I-III)	Narrative-driven	RM, GS	Group	24	Medical	Active	3	0.640; 95% CI [0.382, 1.070]
L. Chen 2022 <sup>195a</sup>	86 (44/42)	64.9	Thyroid (Stage I-IV)	Narrative-driven	ME, RM, GS	Group	24	Medical	Active	2.9	0.318; 95% CI [0.045, 2.259]
Q. Guo 2022 <sup>196a</sup>	138 (69/69)	68.3	Lung (Stage I-III)	Narrative-driven	ME, RM, GS	Group	48	Medical	Active	3	0.668; 95% CI [0.348, 1.285]
Huang 2022 <sup>197a</sup>	108 (55/53)	62.7	Prostate (Stage II-IV)	Narrative-driven	ME, RM, GS	Group	40	Medical	Active	3	0.697; 95% CI [0.244, 1.990]
Kirkegaard 2023 <sup>89a</sup>	198 (96/102)	51.4	Breast (Stage I-III)	CBT-based	EE, CR, CT, PPS, GOS, GS	Group	32	Mixed	Passive	18	0.680; 95% CI [0.410, 1.140]

HR < 1 indicates a survival benefit of the intervention; HR > 1 indicates a survival disadvantage; HR = 1 indicates no effect. The 95% CIs listed here were taken from the articles but may differ in the decimal places from those that appear in the forest plot (see Fig. 4), as they had to be transformed to SEs for the meta-analysis and were then back-transformed by Metafor.

Int. intervention group, Con control group, 95% CI 95% confidence interval, CBT cognitive behavior therapy,

I. Psychological Core Components (Mechanisms): Cognitive-behavioral/self-regulation: PE psychoeducation, GOS goal setting, PPS problem-solving, CR cognitive reappraisal/restructuring, RT relaxation techniques, CT coping skills/techniques, SM stress management,

HY hypnosis, Meaning- & emotion-focused: MBC meaning-based component, EE expression of emotions, Narrative/identity-focused: TS tell their story, RL reflect on their life, RM reminiscence on past life events.

II. Non-specific or unspecific psychological interventions (Formats): PS psychological support, CS coping support, CO counseling, PT psychotherapy, GS group support.

III. Non-psychological or supportive components (contextual): medical/informational: ME medical education; ASIM aid with symptom management. Communication/decision: IC improvement of Communication; ADM aid with decision making. Social support: ESS enhancement of social support, OSS other social support (e.g., web-based).

<sup>a</sup>The HR of these RCTs were not included in any prior meta-analysis.

two parameters)<sup>100</sup>. For the purpose of describing the likely range in which effect sizes of future studies could fall<sup>95</sup>, we present a 95% prediction interval (PI) estimated according to ref. 101.

For the presentation of the meta-analytic results, three forest plots with studies ordered according to effect size (for an overview), precision (to search for small-study effects), and publication year (to visualize time trends) were produced. The robustness of the overall effect was evaluated with a sensitivity analysis (leave-one-out method). The development of the overall effect over time was assessed using a cumulative meta-analysis<sup>102</sup>, starting with the effect of the first published study and adding the other effects one at a time in their order of publication.

In order to provide a more easily grasped effect size for clinicians and patients, a weighted overall median survival (time) difference between intervention and control groups (in months) was calculated, again using an RE model and REML estimator, with the R package *Metamedian* 1.0.0<sup>78,79</sup>. This method estimates study weights with the help of the first as well as third quartiles (if available) and/or the minimum and maximum values of survival time, which may provide good approximations, but not necessarily perfectly accurate weights. Therefore, the robustness of this estimate was also analyzed with sensitivity analysis (leave-one-out method). We emphasize that the weighted median survival difference only provides a rough overall statistic, as it can only be based on trials that actually provide median survival values (those with survival rates <50% in the intervention and control groups) and, hence, this metric was not used for any hypothesis testing.

### Risk of bias within studies

**Overall risk of bias.** The risk-of-bias tool 2 (RoB 2), developed by ref. 103, systematically assesses RCTs (primary studies) for biases in five domains: (1) the randomization process, (2) deviations from intended interventions, (3) missing outcome data, (4) outcome measurement, and (5) selection of reported results. Risk of bias was assessed independently by two authors (KDAH, KS), and differences in opinion were settled by discussion and consensus. Interrater reliability was calculated using Cohen's Kappa ( $\kappa$ ), based on the initial independent ratings prior to consensus, and using interpretation guidelines by ref. 104.

**Study quality.** Study quality was assessed with reference to the outcome of interest (survival) by a combination of checklists (see Supplementary Table A4 for a list of study quality items) consisting of the Quality Assessment of Controlled Intervention Studies (Items 1–14<sup>105</sup>), the DELPHI list (Item 15<sup>106</sup>), and the CONSORT checklist (Items 16–18<sup>107</sup>). A score of +1 was given to each item if the respective criterion was met, -1 if it was not met, and 0 if unknown (range: +18 to -18 points). Literature that was explicitly referred to in the articles of primary studies to contain further information was also checked for assessing the study quality of each trial (see Supplementary Table A5 for a list of additional literature used). Study quality was evaluated independently by two authors (KDAH and AG), and any differences were resolved by discussion and consensus.

**Post hoc study power.** To date, there is no approved method to estimate study-specific post hoc power, as every study may represent a slightly different population. Each population effect is required to be known to calculate study-specific post hoc power, but *specific* population effects cannot be reliably estimated from single studies<sup>108,109</sup>. However, it can be argued that meta-analytic post hoc power follows a more plausible rationale, as its assessment is not based on study-specific effects, but the assumed *common* population effect estimated by the meta-analysis.

Hence, we calculated the meta-analytic post hoc power of each RCT according to an (adjusted) R code from ref. 110 by taking each sample size ( $n$ ), each proportion of patients in the intervention group ( $n_{int}/n$ ), the estimated overall effect HR of our meta-analysis, and the probability of events (deaths/ $n$ ) within each trial into account.

**Sponsorship bias.** Studies were grouped according to similar categories proposed by ref. 111, although in a more refined manner: no funding involved (coded 0), government funding only (coded 1), (also) public charity funding involved (coded 2), (also) private foundation funding involved (coded 3), (also) for-profit organization funding involved (coded 4), and for-profit organization funding only (coded 5). The analysis of funding sources compared the subgroup summary effects of all groups with each other (codes 0–5), whereas the analysis of risk of sponsorship bias compared the subgroup summary effects of groups with either low (coded 0–3), moderate (coded 4), or high risk (coded 5) of sponsorship bias.

**Moderator analysis.** Potential causes of heterogeneity were investigated with a systematic approach to effect-moderator analysis that involved: (1) all moderator variables explored by the ten prior meta-analyses; (2) factors that (significantly) affected survival in the included RCTs; (3) factors identified in additional literature (e.g., on important prognostic factors); and (4) the risk of bias factors study quality, post hoc power, and sponsorship bias (see Supplementary Table A6 for a moderator disclosure table).

A mixed-effects (ME) meta-analytic model was used in the subgroup analysis to identify moderators that may generalize to other populations. It is a standard approach in subgroup analysis to model the variation within subgroups as a random effect and the variation between subgroups as a fixed effect<sup>38</sup>.

For categorical data, we formed subgroups containing at least two studies ( $k \geq 2$ ), calculated summary effects for each subgroup, and subsequently tested whether any of these differed significantly from zero (test for subgroup effects; TfSE). Next, we tested whether subgroup summary effects differed significantly from each other (test for subgroup differences; TfSD). A categorical moderator was identified only when both tests delivered significant results.

For continuous data, we tested each predictor (moderator) and the outcome (survival) for a significant correlation (test of moderator; ToM) in individual meta-regression analyses if at least three studies were available ( $k \geq 3$ ). Note that results of subgroup analyses (due to lack of statistical power) and meta-regressions (due to risk of overfitting) that include less than ten studies ( $k < 10$ ) should be interpreted with caution<sup>93</sup>. Additionally, studies should be evenly distributed among subgroups in subgroup analysis.

Residual heterogeneity was assessed with the  $Q$  test, and the relative amount of variance is reported by  $I^2$ . A Benjamini–Hochberg procedure<sup>112</sup>, with a false discovery rate (FDR) of 10%, was applied to all subgroups ( $p$  values of TfSD) and meta-regression analyses ( $p$  values of ToM) in order to control for the accumulation of type-1 errors from multiple testing.

### Exploratory machine-learning-based search for moderators

An exploratory search for potential moderators was performed with a recently developed machine-learning and weighted bootstrap sampling approach called *MetaForest*<sup>61,113</sup>, which derives from an adaptation of the random forest algorithm<sup>114,115</sup>. The method is robust against overfitting, captures linear and non-linear relationships, and can also be used to detect higher-order interactions. Importantly, this procedure allows reliable identification of relevant moderators from a larger set of potential moderators in meta-analyses, even when the number of moderators is larger than the actual number of included studies, due to higher power than conventional moderator analyses<sup>61,113</sup>. We followed the approach for small-sample meta-analyses as outlined by ref. 113.

First, a data file was created in R containing only the moderators of interest with  $\leq 3$  missing values ( $\leq 10\%$ ) replaced by their mode (for categorical variables) or median values (for continuous variables), applying simple imputation, in order for *MetaForest* to work properly. Moderators with  $> 3$  missing values ( $> 10\%$ ) were excluded to avoid bias. We refrained from using multivariate imputation techniques due to the low inter-correlations between many of the moderators, which would have

compromised the quality of the imputed values. Second, the optimal number of trees was identified by running a test model with a high tree number (20,000), which allowed us to determine the number of trees at which the model's out-of-bag (OOB) cumulative mean squared (prediction) error ( $MSE_{\text{Oob}}$ , a measure of predictive accuracy<sup>61</sup>) converged. Third, a final model with the optimal number of trees (the value at which the test model converged) was run, in order to avoid unnecessary model complexity.

Fourth, in order to eliminate irrelevant moderators, a 100-fold recursive preselection algorithm was applied, which drops the moderator with the most negative variable importance from the model before rerunning the analysis until only moderators with positive variable importance (in >10% of replications) remain. Fifth, only the moderators for which the 50th percentile interval (range between the 25th and 75th percentiles) from the variable importance metrics did not include zero were kept in the model. Sixth, optimal tuning parameters (type of weights, number of candidate variables per split, and minimum terminal node size) were determined via machine-learning techniques implemented in the R package *caret*<sup>16</sup>, which uses tenfold clustered cross-validation to find the optimal model with the smallest (cross-validation) root mean squared error ( $R^2_{\text{cv}}$ ).

Seventh, for the final model with optimal tuning parameters, a convergence plot (which visualizes at what number of trees the cumulative  $MSE_{\text{Oob}}$  stabilized<sup>13</sup>), a variable importance plot (which shows the relative importance of each relevant moderator in predicting the effect size<sup>13</sup>), and a partial dependence plot (which displays the relationship between a relevant moderator and the effect size, whilst averaging over all other moderators<sup>61</sup>) were created for the presentation of results. Additionally, first-order interactions between identified moderators were explored by generating partial dependence plots of important bivariate interactions.

### Reporting bias across studies

Reporting bias across trials was investigated with several contemporary methods, as each of these addressed different features of the data and their combination was thus considered more informative than each method alone.

First, a contour-enhanced funnel plot<sup>117,118</sup> was visually inspected to search for hints of publication bias (i.e., if studies with high standard errors were selectively missing at one side of the funnel plot).

Second, we tested for the presence of publication bias with *p*-uniform<sup>119</sup>, which examines the distribution of significant *p* values and FAT-PET-PEESE (exploratory), which regresses observed effect sizes either on their standard errors (PET) weighted by the least squares method or on their sampling variances (PEESE) and tests for a significant slope (FAT; test for publication bias).

Third, *p*-uniform\* (an advanced version of *p*-uniform<sup>120</sup>), which estimates the true overall effect size adjusted for publication bias (if present) by modeling the distribution of both significant and not significant *p* values and PET-PEESE, which uses regression modeling between effect size and precision (standard errors) by either fitting a linear regression (PET; used when the effect is estimated to be zero) or a curvilinear regression (PEESE; used when the overall effect is estimated to be non-zero<sup>65,121</sup>) were planned to adjust the overall effect estimate when two conditions were met: (1) Publication bias could be clearly identified; and (2) the Meta Showdown Explorer<sup>66</sup> attested a higher likelihood of accurately and precisely estimating the adjusted overall effect for one of the two methods compared to the simple RE model, based on pre-computed simulation data under conditions similar to the given dataset.

Fourth, *p*-curve<sup>122,123</sup>, which, like *p*-uniform, explores the distribution of significant *p* values, was used to examine the data for evidence of *p*-hacking (testing for left-skewness of the *p*-curve).

Fifth, it was tested whether the studies contained (any) evidential value (testing for right-skewness of the *p*-curve), for which either the test for half the curve (*p* values of primary studies <0.025) has to be significant ( $p < 0.05$ ) or both, the test for half the curve (*p* values of primary studies <0.025) and the full curve (*p* values of primary studies <0.05), have to yield  $p < 0.10$ .

Sixth, it was tested whether the studies' evidential value was inadequately low (testing for flatness of the *p*-curve; power <33%) for which either the test for the full curve (*p* values of primary studies <0.05) has to be significant ( $p < 0.05$ ) or both, the test for half the curve (*p* values of primary studies <0.025) and the binomial test (share of results <0.025), have to yield  $p < 0.10$ <sup>122,123</sup>.

### Specification curve and multiverse meta-analysis

The ambiguity in the results of prior meta-analyses was investigated via specification curve and multiverse meta-analysis<sup>76</sup>.

First, the WHICH (data universe) and HOW factors (model universe) of the prior meta-analyses were identified (see Supplementary Table A7 for a WHICH and HOW factor disclosure table). Some prior meta-analyses focused on specific (sub-)populations, such as (non-)metastatic female breast cancer patients, whereas others had no such restrictions. Some analyses based the decision of the utilized meta-analytic model (FE or RE model) on the presence of statistical heterogeneity (the result of the *Q* test), whereas others decided to use the RE model on a priori theoretical grounds; the latter is nowadays considered appropriate<sup>38</sup>. Accordingly, the factors sex of patients (female only vs. all patients), cancer type (breast cancer only vs. all cancer types) and presence of metastasis (metastatic vs. non-metastatic vs. all stages) were defined as WHICH factors, whereas the meta-analytic model (FE vs. RE) was defined as a HOW factor. Note that other theoretically possible categories (e.g., male only or other cancer types) were not included, as no prior meta-analyses had examined them, and available RCTs were also limited in this regard.

Additionally, prior meta-analyses using ORs or RRs calculated multiple overall effects at various follow-up endpoints (at 1, 2, 4, 5, 6, or 10 years), unlike meta-analyses using HRs, which always made use of the full study periods (see Table 1). However, different follow-up endpoints were not used as a WHICH factor, because (1) with six different follow-up endpoints and  $k = 32$  trials each level would have included only a handful of studies (at best) and (2) this approach would not have adequately reflected what prior meta-analyses (using ORs or RRs) had actually done (i.e., coding multiple effect sizes at various follow-up endpoints for each and every trial; where data were available).

Hence, the follow-up endpoint was coded as a HOW factor for ORs and RRs by calculating multiple overall effects at 1, 2, 4, 5, 6, and 10 years for each study, where data were available, but not for HRs (as the latter always makes use of the full study period available). This entailed that the number of HOW factors differed between ORs, RRs, and HRs, which is why separate descriptive specification curves had to be plotted for ORs, RRs, and HRs, respectively.

Although the full dataset underlying the multiverse meta-analysis included multiple effect sizes from some studies across different follow-up intervals, these were never analyzed together within the same model. Thus, while some dependencies might have existed in the broader context, they did not violate the assumption of independence within any single meta-analysis reported in the multiverse. Each analysis was conducted using standard random-effects models, which are appropriate under these conditions.

The three WHICH factors sex of patients (2 levels), cancer type (2 levels), and presence of metastasis (3 levels) as well as the HOW factor meta-analytic model (2 levels) and the HOW factor follow-up endpoint (6 levels; for OR and RR) resulted in 24 ( $2 \times 2 \times 3 \times 2$ ) possible specifications for the HR and 144 ( $2 \times 2 \times 3 \times 2 \times 6$ ) possible specifications for the OR and RR. Note that the algorithm of the R code only kept unique specifications with  $k \geq 2$  studies; this means that the actual number of specifications shown in the resulting plots might be lower than the maximum number possible.

Second, the original code for the specification curve and multiverse meta-analysis<sup>76</sup> was adjusted to fit the conditions of the current dataset (see Supplemental File S3 for the data sheets and the R scripts of the multiverse meta-analysis at <https://osf.io/wnxdb/files>).

Third, descriptive meta-analytic specification curve plots for ORs, RRs, and HRs were created that displayed the observed overall effect sizes, their

95% CIs, and the included number of studies for each specification in ascending order.

Fourth, separate  $p$  value histograms were generated for ORs, RRs, and HRs that allowed us to examine whether or not the overall distribution of  $p$  values diverged from what would be expected from true zero overall effects (i.e., if  $p$  values were evenly distributed or if there was an accumulation of significant  $p$  values >5%).

Fifth, inferential meta-analytic specification curve plots for ORs, RRs, and HRs with overall effects and their 95% CI bands from 1000 simulated specifications with true overall effects of zero were constructed (via bootstrapping) and contrasted against the *actually* observed overall effects from the descriptive specification curve plots; all based on the corpus of available studies. Regions where the curve of observed overall effects lay outside the 95% CI band of simulated true zero overall effects were interpreted as evidence for the existence of a true (nonzero) overall effect for the respective specifications.

Note that separate data files and separate inferential specification curve plots for follow-up intervals at 1, 2, 4, 5, 6, and 10 years (regarding ORs and RRs) were created, as otherwise the data contained missing values for certain trials and specific follow-up intervals (e.g., not every trial had a 10-year follow-up), which could not be handled otherwise by the multiverse meta-analysis R code.

### Certainty in the evidence (GRADE rating)

The certainty of evidence was evaluated using the GRADE framework, which assesses the confidence in the estimated overall effect across five key domains: risk of bias, inconsistency, indirectness, imprecision, and publication bias. Given that the present analysis was based exclusively on RCTs, the starting level of certainty was high, as recommended by the GRADE Working Group<sup>24</sup>. We then considered each potential reason for downgrading:

I. Risk of bias was evaluated using the RoB 2 tool<sup>103</sup> across all five domains: (1) bias due to the process of randomization, (2) deviation from the intended intervention, (3) missingness of the outcome data, (4) measurement of the outcomes, (5) selection of the reported result, and an “overall risk of bias” judgment. Downgrading was considered if a substantial proportion of studies had either “some concerns” or “high risk” ratings.

II. Inconsistency was evaluated based on statistical heterogeneity indicators, particularly  $I^2$ ,  $\tau^2$ ,  $\tau$ , and the 95% PI. Inconsistency was judged serious if variation in effect estimates could not be fully explained by moderators.

III. Indirectness was considered if the included studies deviated from the population, intervention, comparator, or outcome of interest. No downgrading was planned unless systematic differences were evident.

IV. Imprecision was assessed primarily based on whether the 95% CI of the overall effect HR included the null. A secondary measure, such as the overall median survival (time) difference, was not part of hypothesis testing and was therefore not considered.

V. Reporting bias (such as publication bias and  $p$ -hacking) was evaluated using multiple approaches (contour-enhanced funnel plot,  $p$ -uniform, FAT-PET-PEESE, and  $p$ -curve) and downgraded only in case of consistent evidence for a reporting bias.

Additionally, upgrading factors such as (1) a large magnitude of effect, (2) a dose-response relationship, or (3) plausible confounding reducing the effect size were considered.

### Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

## Results

### Study selection

Literature search for RCTs in electronic databases yielded a total of 6848 publications (see Fig. 3 for a PRISMA flow chart of the search for trials). Following the removal of 1427 duplicate entries, a total of 5421 articles were

retained for further analysis. Subsequently, titles and abstracts were screened, and 5312 articles were removed because they did not fit the inclusion criteria (three of which were retracted papers). A total of 109 articles were retrieved for full inspection, of which 77 articles were omitted for meeting exclusion criteria (see Supplementary Table A8 and Note 1 for details on excluded articles). Finally, 32 RCTs remained and were included for analysis. Literature search for meta-analyses identified ten prior meta-analyses (see Supplementary Fig. A3 for a PRISMA flow chart of the search for meta-analyses), and search for gray literature resulted in 0 articles (see Supplementary Fig. A4 for a PRISMA flow chart of the search for gray literature).

### Study characteristics

In total, 32 RCTs involving 5704 patients were included in the meta-analysis, and the HRs of 14 of these trials had not been included in any of the prior meta-analyses reviewed (see Table 4 for a list of study characteristics for all RCTs). Sample sizes varied from 55 to 375 patients ( $M = 178$ ). Mean ages ranged from 43 to 68 years ( $M = 56$ ). Cancer types were brain ( $k = 1$ ), breast ( $k = 11$ ), colorectal ( $k = 2$ ), esophageal ( $k = 1$ ), gastric ( $k = 1$ ), leukemia ( $k = 1$ ), liver ( $k = 1$ ), lung ( $k = 2$ ), melanoma ( $k = 2$ ), prostate ( $k = 1$ ), thyroid ( $k = 1$ ), and mixed cancer type samples ( $k = 8$ ). Intervention types were CBT-based ( $k = 11$ ), meaning-enhancing ( $k = 5$ ), narrative-driven ( $k = 8$ ), and other counseling ( $k = 8$ ). Interventions were delivered either individually ( $k = 11$ ) or in groups ( $k = 21$ ), and dosages ranged from 3 to 96 h ( $M = 33$ ). Control groups were either active ( $k = 14$ ) or passive ( $k = 18$ ). Follow-ups ranged from 0.1 to 18 years ( $M = 6$ ).

### Individual studies

HRs of the intervention group (see Supplementary Table A9 for an effect size and  $p$  value disclosure table) were available for 16 RCTs, calculated from the inverse HRs of the control group for two RCTs, taken from prior meta-analyses for four RCTs, and calculated from Kaplan–Meier survival curves for ten RCTs. The observed effect sizes ranged from  $\log HR = -1.56$  to 0.35 ( $HR = 0.21$  to 1.42) with 72% of them below zero (HRs below 1; see Fig. 4 for a forest plot sorted according to effect size).

### Meta-analytic synthesis

There was a small overall effect favoring longer survival in the intervention group, with  $\log HR = -0.23$ , 95% CI  $[-0.35, -0.11]$ ,  $SE = 0.06$ ,  $z = -3.68$ ,  $p < 0.001$ , which translates into  $HR = 0.80$ , 95% CI  $[0.71, 0.90]$ . The  $Q$  test for the presence of true effect size heterogeneity was significant with  $Q(31) = 59.27$ ,  $p = 0.002$ . The proportion of true variance to total variance was  $I^2 = 48\%$ , 95% CI  $[19, 77]$ ; the variance of true effects (log scale) was  $\tau^2 = 0.05$ , 95% CI  $[0.01, 0.18]$ ; and the standard deviation (log scale) was  $\tau = 0.23$ , 95% CI  $[0.11, 0.43]$ . The 95% PI ranged from HR 0.49 to 1.29. Consequently, although there was evidence of a small overall effect across all studies, effects in future RCTs could (potentially) range from substantially positive to slightly unfavorable; however, none of the included trials reported a significant negative effect on survival. Sensitivity analysis showed that the estimated overall effect was robust, with effect sizes ranging between  $HR = 0.78$ , 95% CI  $[0.69, 0.88]$  to  $HR = 0.81$ , 95% CI  $[0.72, 0.92]$  when refs. 125,126 were removed, respectively.

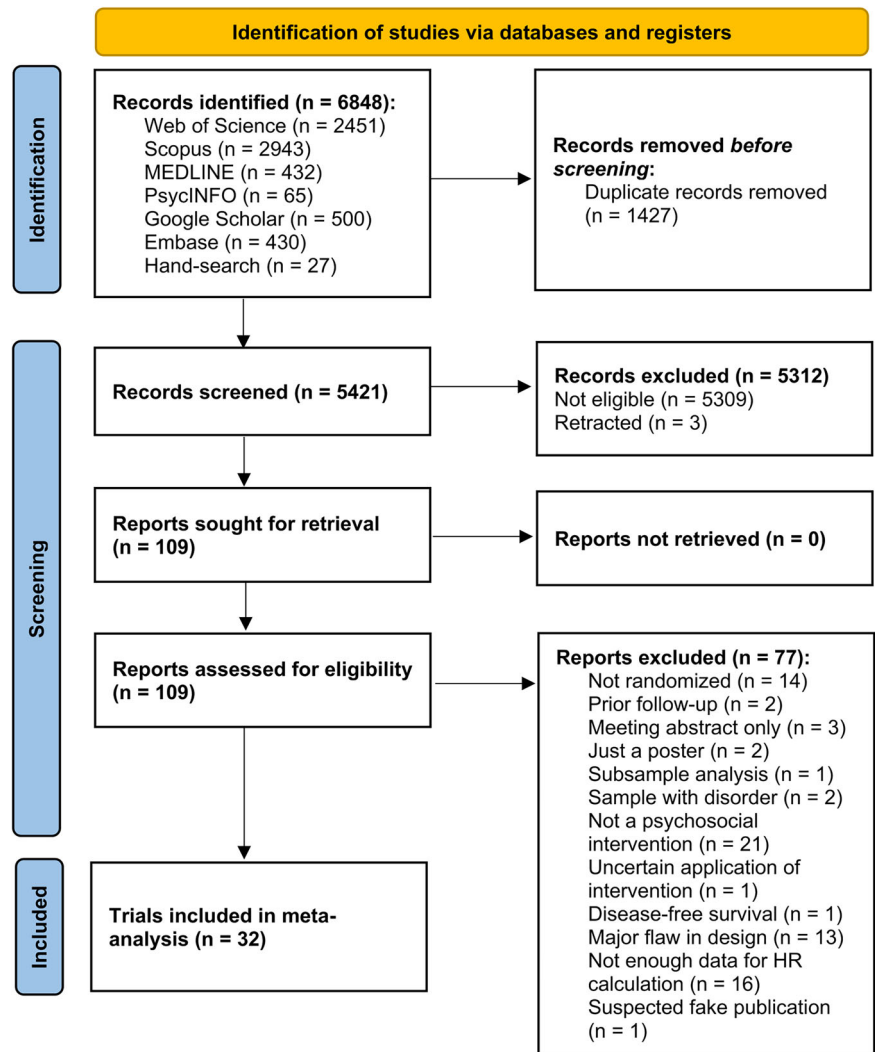
### Small-study effects

The pattern in the forest plot sorted by precision did not fully align with a classic small-study effect or publication bias. Although the smallest, least precise studies tended to report the largest effects, a stable trend appeared in the middle section, where studies with good to moderate precision also consistently showed positive effects (see Supplementary Fig. B1 for a forest plot sorted according to precision).

### Time trend

A small time trend appeared to be present in the forest plot sorted by publication year. In earlier studies (before 2015), the HRs were relatively mixed, with values both smaller and greater than 1 (see Supplementary Fig.

**Fig. 3** | PRISMA flow chart 2020 displaying the selection process of the systematic literature search for RCTs.



B2 for a forest plot sorted according to publication year). However, from 2015 onward, the HRs consistently fell below 1, suggesting a potential systematic shift in effect sizes over time.

Similarly, cumulative meta-analysis indicated that only from the year 2015 onwards were enough trials available to provide sufficient precision and statistical power for the small overall effect HR to become significant (see Supplementary Fig. B3 for a forest plot of the cumulative meta-analysis).

### Overall median survival difference

The overall median survival difference for cancer patients treated with psychosocial interventions showed a trend with a survival benefit of 3.9 months, 95% CI [-0.7, 8.5], SE = 2.34,  $z = 1.66$ ,  $p = 0.10$ ; based on articles that provided sufficient data ( $k = 16$ ). Sensitivity analysis showed that the overall median survival difference ranged between 2.1 months, 95% CI [-0.6, 4.8] to 4.9 months, 95% CI [0.7, 9.2] when refs. 91,127 were removed, respectively.

### Active psychological components

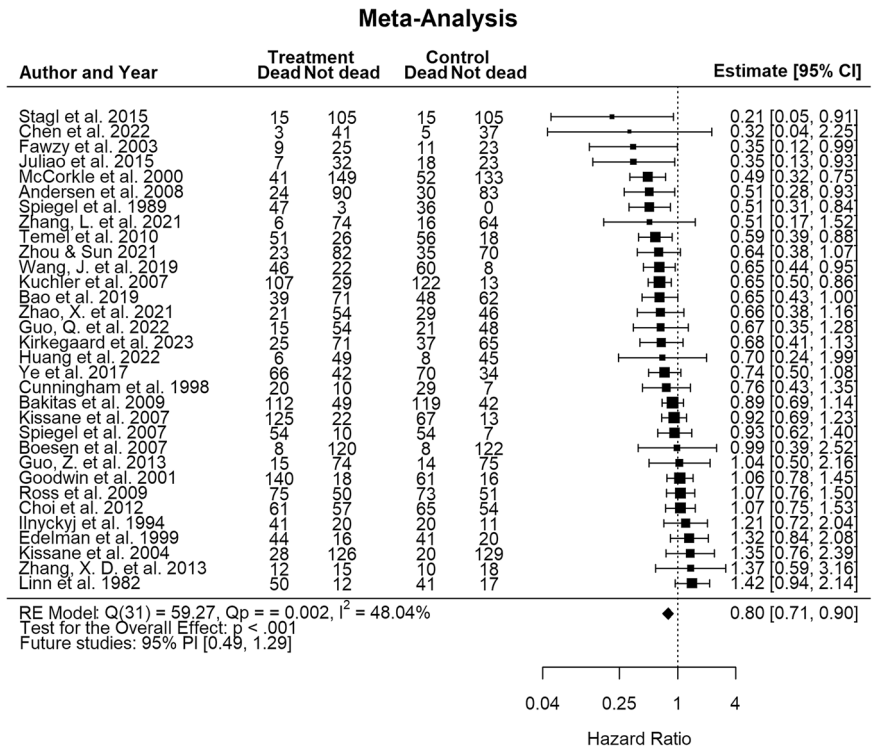
Psychosocial interventions in this meta-analysis were consistently multi-component and delivered either in a group-based format ( $k = 21$ ; four trials with significant/17 with non-significant effects), which typically entailed elements of group-based social support, or in an individual-based format ( $k = 11$ ; 6/5), delivered on a one-to-one basis. With regard to other active psychological components, the most frequently included elements were medical education ( $k = 16$ ; 6/10), coping skills/techniques ( $k = 13$ ; 5/8), and

expression of emotions ( $k = 12$ ; 3/9). Components of moderate frequency comprised relaxation techniques ( $k = 8$ ; 4/4), problem-solving ( $k = 8$ ; 3/5), (non-specific) psychological support ( $k = 7$ ; 4/3), enhancement of social support ( $k = 6$ ; 3/3), cognitive reappraisal or restructuring ( $k = 6$ ; 2/4), goal setting ( $k = 6$ ; 2/4), reminiscence of past life events ( $k = 6$ ; 0/6), and improvement of communication ( $k = 5$ ; 2/3). Components of lower frequency were mindfulness-based components ( $k = 4$ ; 2/2), psychoeducation ( $k = 4$ ; 1/3), coping support ( $k = 3$ ; 3/0), counseling ( $k = 3$ ; 2/1), and stress management ( $k = 3$ ; 1/2). Less frequently reported elements are listed elsewhere (see Table 5 for a full list of psychological components).

### Risk of bias within studies

The overall risk of bias ranged from some concerns ( $k = 30$ ) to high risk ( $k = 2$ ), primarily due to insufficient information on allocation concealment and the absence of a preregistration protocol (see Supplementary Fig. B4 for a Risk-of-Bias Plot). Interrater reliability between the two raters for the initial risk of bias assessment was high, with Cohen's  $\kappa = 0.94$ , indicating almost perfect agreement<sup>104</sup>. For included RCTs, average total score of study quality was moderate with  $M = 7$  points (range 1 to 11; see Supplementary Tables B1–B3 for a list with study quality items), whereas average meta-analytic post hoc power was very low with  $M = 17\%$  (range 6 to 40%; see Supplementary Table B4 for a meta-analytic post hoc power calculation sheet). The corpus of primary studies contained five unfunded and 21 funded trials, while no information was given in six trials.

**Fig. 4 | Forest plot of the random-effects meta-analysis showing survival data of all 32 included RCTs (ordered according to effect size).** HR <1 indicates a survival benefit for the intervention group; HR >1 indicates a survival benefit for the control group; HR = 1 indicates no effect. The 95% CIs in the forest plot may differ in the decimal places from those stated in the articles (see Table 4), as they had to be transformed to SEs for the meta-analysis and were then back-transformed for the forest plot by Metafor. RE model random-effects model,  $Q = Q$  test for heterogeneity,  $Q_p = Q$  test  $p$  value,  $I^2 =$  Proportion of true variance to total variance, 95% PI = 95% prediction interval,  $k = 32$  studies.



Risk of sponsorship bias was low for 21 studies and moderate for five, while no information was present for the same six trials (see Supplementary Table B5 for a funding source and risk of sponsorship bias rating sheet).

**Moderator analysis**

Analysis of 25 moderators (one already excluded beforehand due to no data) and four risk of bias factors (see Supplementary Table C1 for an overview) was conducted (see Supplementary Fig. C1–C29 for specific forest and bubble plots). The moderator *analysis type* revealed that studies using ITT analysis reported a significant subgroup effect (HR = 0.74, 95% CI [0.65, 0.85]), whereas studies that did not apply ITT analysis did not report a significant effect (HR = 0.96, 95% CI [0.78, 1.18]; see Supplementary Fig. C12 for a forest plot of the moderator analysis type). However, the  $p$  value for this moderator (TISD,  $p = 0.038$ ) exceeded the Benjamini–Hochberg critical threshold of  $p = 0.003$  for this rank, and thus did not remain significant after correction for multiple comparisons.

Therefore, this finding should be interpreted with caution. Another moderator, *surgery prior to study entry*, indicated a trend (ToM  $p = 0.063$ ): studies with a higher percentage of patients who had undergone surgery before study entry showed greater effects (see Supplementary Fig. C22 for a bubble plot of the moderator surgery prior to study entry).

**Identifying relevant moderators via machine learning**

The machine-learning-based moderator search using MetaForest included 15 moderators and two risk of bias factors (each with  $\leq 3$  missing values), whereas 11 moderators and two other risk of bias factors (with  $> 3$  missing values) were excluded (see Supplementary Table C2A, B for a list of included and excluded moderators). The final model was run with 5000 regression trees, showed a cumulative  $MSE_{\text{ob}} \sim 0.22$ , and converged at  $\sim 2000$  trees (see Supplementary Fig. D1 for a convergence plot of the final model). The optimal tuning parameters were weights = uniform, candidate variables per split = 2, and minimum terminal node size = 3.

The model yielded a cross-validation prediction error of  $MSE_{\text{ob}} = 0.19$ , indicating moderate error between the model’s predictions and actual values. It had a positive but small predictive value in unseen data,

with  $R^2_{\text{ob}} = 0.12$ , suggesting that only 12% of the variance in new data were explained—indicating weak predictive performance, which was expected due to the high number of moderators and the limited number of studies.

In contrast, it achieved moderate predictive performance using cross-validation with  $R^2_{\text{cv}} = 0.28$ , suggesting that the model explained 28% of the variance within the resampling procedure. This suggests the model performed better in internal validation than on genuinely new data. The gap between  $R^2_{\text{ob}}$  and  $R^2_{\text{cv}}$  indicates mild overfitting to the training data.

Identified moderators sorted according to variable importance (VI) were *analysis type* (VI =  $\sim 0.038$ ), *tumor phase* (VI =  $\sim 0.029$ ), and *study quality* (VI =  $\sim 0.025$ ; see Supplementary Fig. D2 for a variable importance plot). Note that clear cutoff values for MetaForest analysis do not exist, but a general rule of thumb in random forest analyses considers VI values of 0.01–0.05 as low, 0.05–0.1 as moderate, and  $> 0.1$  as high<sup>128</sup>.

Partial dependence plots, displaying the relationships of each identified moderator with the effect size while averaging over the other moderators, were generated in combination with the raw data and the 95% percentile intervals of the predictions of individual trees from the model (see Supplementary Fig. D3 for partial dependence plots of the identified moderators). The plots for analysis type and study quality indicated slight tendencies that studies applying an intention-to-treat (ITT) analysis or achieving a higher quality score reported greater life-prolonging effects. Partial dependence plots displaying bivariate relationships between the identified moderators did not indicate any interactions (see Supplementary Fig. D4 for partial dependence plots of bivariate interactions).

**Reporting bias across studies**

**Publication bias.** There was some visual indication of publication bias in the contour-enhanced funnel plot (see Fig. 5 for a contour-enhanced funnel plot), as trials with larger standard errors (usually those with small sample sizes) seemed to be missing in the lower right area of the plot (i.e., effect sizes in the direction opposite to the favored outcome). However, the test for publication bias of  $p$ -uniform (based on  $k = 10$  significant effect sizes) was not significant  $p = 0.266$ . Similarly, the funnel plot asymmetry test of FAT-PET-PEESE (exploratory; based on all  $k = 32$  effect sizes) did not confirm publication bias but only showed a trend ( $p = 0.091$ ).

**Table 5 | Psychological components of the 32 included RCTs that investigated psychosocial interventions and survival in cancer patients ordered according to absolute and relative frequency (number of significant vs. non-significant studies)**

Psychological components	Absolute frequency	Relative frequency
Group support	21	4/17
Medical education	16	6/10
Coping skills/techniques	13	5/8
Expression of emotions	12	3/9
Relaxation techniques	8	4/4
Problem-solving	8	3/5
(Undefined) psychological support	7	4/3
Enhancement of social support	6	3/3
Cognitive reappraisal/restructuring	6	2/4
Goal setting	6	2/4
Reminiscence of past life events	6	0/6
Improvement of communication	5	2/3
Mindfulness-based component	4	2/2
Psychoeducation	4	1/3
Coping support	3	3/0
Counseling	3	2/1
Stress management	3	1/2
Aid with decision making	2	1/1
Hypnosis	2	1/1
Tell their story	2	1/1
Aid with symptom management	2	0/2
Reflect on their life	2	1/1
Psychotherapy	1	1/0
Other social support	1	0/1

**Correcting for publication bias.** The Meta Showdown Explorer was tuned to the conditions of the current dataset: severity of publication bias = none (as indicated by  $p$ -uniform and FAT-PET-PEESE; the option small was not available, only none, medium, and high), heterogeneity  $\tau = 0.2$ , number of included studies  $k = 30$ , true effect size  $SMD = 0.2$  (a small effect), questionable research practices = none (as indicated by  $p$ -curve).

The outcome suggested that the (naïve) RE meta-analysis would outperform  $p$ -uniform and FAT-PET-PEESE in estimating the overall effect with a lower false-positive rate (type-1 error probability) and a higher true-positive rate (indicative of statistical power) than both methods (see Supplementary Fig. E1, E2 for a Meta Showdown Explorer Plot of the false- and true-positive rate).

Regarding the precision in estimating a true zero effect ( $\delta = 0$ ) and a small true (nonzero) effect ( $\delta = 0.2$ ), the naïve (RE) meta-analysis would again outperform  $p$ -uniform and be just as precise as FAT-PET-PEESE (see Supplementary Figs. E3, E4 for a Meta Showdown Explorer Plot of correctly estimating a zero and small overall effect).

For the sake of completeness, the results of FAT-PET-PEESE are reported. PET estimated an overall effect of 1.02, 95% CI [0.78, 1.35],  $p = 0.871$ , while PEESE estimated an overall effect of 0.89, 95% CI [0.77, 1.02],  $p = 0.096$ , corrected for an undetectable (and potentially distorted) reporting bias (see Supplementary Fig. E5 for a FAT-PET-PEESE funnel plot).

**$p$ -Hacking and tests of evidentiality.** The  $p$ -curve analysis (based on  $k = 10$  significant  $p$  values) showed no visual indication of  $p$ -hacking (see Fig. 6 for the  $p$ -curve of all 10 significant  $p$  values). Two separate tests were conducted as part of the  $p$ -curve analysis.

First, the test for evidential value, which examines whether the  $p$ -curve is significantly right-skewed (indicating the presence of true effects rather than selective reporting or randomness), was not significant for the full curve ( $p$  values  $< 0.05$ ;  $p = 0.458$ ), but significant for the half curve ( $p$  values  $< 0.025$ ;  $p = 0.029$ ). According to ref. 123, evidential value is supported if either the full or the half-curve test is significant, indicating that the primary studies showed evidential value.

Second, the test for inadequate evidential value due to low statistical power (i.e.,  $< 33\%$  power) was significant for the full curve ( $p = 0.043$ ), but not for the half-curve ( $p = 0.720$ ) and the binomial test (proportion of  $p$  values  $< 0.025$ ;  $p = 0.935$ ). Following ref. 123, evidence is considered inadequately powered if either the full-curve test is  $p < 0.05$ , or if both the half-curve and binomial tests are  $p < 0.10$ , suggesting that the primary studies may be underpowered.

Taken together, the studies analyzed thus appeared to provide some evidential value, but may suffer from low statistical power, meaning that exact replications could have a low probability of detecting similar effects. Importantly, the number of significant  $p$  values included in this analysis was at the minimum recommended for  $p$ -curve ( $k = 10$ ), and the method is known to sometimes overestimate evidential value under such conditions<sup>129</sup>, limiting the robustness and generalizability of the inferences. Consequently, these results should be interpreted with caution and considered preliminary rather than definitive.

Similarly, the power of the significant  $p$  values from the individual HRs to detect an overall effect, corrected for selective reporting (for possible publication bias and  $p$ -hacking; ref. 130), was very low at 5%, 90% CI = [5, 31%]. Again, because this estimate is based on the minimum recommended number of significant  $p$  values ( $k = 10$ ), it may not precisely reflect the true evidential power. Therefore,  $p$ -curve might provide a distorted power estimate<sup>66</sup>.

### Specification curve and multiverse meta-analysis

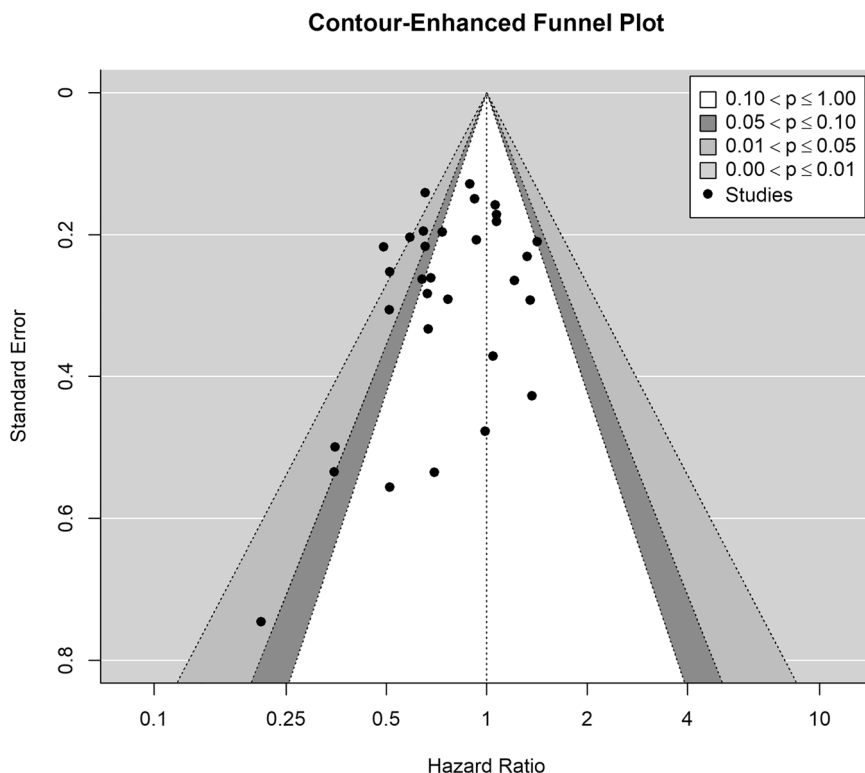
Descriptive specification curve plots (see Figs. 7–9 for descriptive specification curve plots separated into HRs, ORs, and RRs) showed that those meta-analytic specifications restricting the corpus of primary studies ( $k = 32$ ) to a subset of trials (for logHRs:  $k < 12$ ; for logORs and logRRs:  $k < 18$ )—for example, setting a focus on metastatic or non-metastatic female breast cancer patients (for all effect sizes) or on follow-up intervals longer than two years (only for logORs and logRRs)—frequently resulted in non-significant overall effects (with the 10-year follow-up being the exception).

Moreover, the amount of significant  $p$  values from the observed overall effects were  $\sim 42\%$  (logHRs),  $\sim 13\%$  (logORs), and  $\sim 22\%$  (logRRs; see Fig. 10 for histograms of  $p$  value distributions), which is much higher than what would be expected if a general zero overall effect across all specifications was true (in that case the expected amount of significant  $p$  values would be  $\sim 5\%$  for each effect size). This finding suggests the existence of a general true (nonzero) overall effect of psychosocial interventions enhancing the survival time of cancer patients across the specifications examined in prior meta-analyses and the present review; based on the currently available corpus of primary studies.

Notably, significant  $p$  values for logHRs were much more likely than for logORs and logRRs, due to the fact that effect sizes of the corpus of primary studies were pooled over all follow-ups for the calculation of logHRs (thus incorporating a relatively high amount of trials for each specification leading to a higher precision and, therefore, a higher likelihood of overall effects to turn out significant), in contrast to pooling effect sizes from subsets of the corpus of primary studies for the calculation of 1-, 2-, 4-, 5-, 6-, or 10-year follow-up intervals for logORs and logRRs (thus incorporating a gradually smaller number of trials with longer follow-up intervals); note that almost all studies had at least a 1-year follow-up, but only half of all studies had a 5-year follow-up period.

Remarkably, the inferential (bootstrap) specification curve plot for logHRs (see Fig. 11 for an inferential specification curve plot of logHRs) confirmed that almost all specifications of observed overall effects deviated clearly (and significantly) from the 95% CI band of 1000 simulated overall

**Fig. 5 | Contour-enhanced (RE) funnel plot with summary effects from all 32 included studies.** HR <1 indicates a survival benefit for the intervention group; HR >1 indicates a survival benefit for the control group; HR = 1 indicates no effect. The reference line was drawn at HR = 1 (no effect) with significance contours at  $p < 0.10, 0.05, 0.01$ ,  $k = 32$  studies.



true zero effects, which can be viewed as evidence for the existence of a *general* true (nonzero) overall effect across specifications for logHRs.

Likewise, the inferential (bootstrap) specification curve plots for logORs and logRRs, split into 1-, 2-, 4-, 5-, 6-, and 10-year follow-ups (see Fig. 12 and 13 for inferential specification curve plots of logORs and logRRs), showed that the observed overall effects of several specifications of the 1-, 2-, and 10-year follow-up interval deviated slightly (yet significantly) from the 95% CI band of true zero effects, which can be seen as evidence for the existence of a *general* true (nonzero) overall effect across specifications for logORs and logRRs as well (at least for the mentioned follow-up intervals).

**Certainty in the evidence (GRADE rating)**

Applying the GRADE framework, the overall certainty of evidence was judged to be moderate for the conclusion that psychosocial interventions prolong survival in cancer patients (see Supplementary Figs. E6 and E7 for a Summary of findings table and a GRADE evidence profile).

I. Risk of bias was deemed not serious, as 30 of the 32 included RCTs were rated as having some concerns, primarily due to unclear allocation concealment or missing preregistration. These issues reflected that the corpus of primary studies encompassed trials published over four decades, during which reporting standards evolved: allocation concealment reporting and preregistration became standard only after the mid-2000s and the mid-2010s, respectively. Only two studies were rated as high risk due to excessive participant exclusion. Given the nature and distribution of these issues, no downgrading was applied.

II. Inconsistency was rated as serious. Although the direction of the effect consistently favored the intervention group, moderate to substantial heterogeneity ( $I^2 = 48%$ ) and a wide 95% PI (HR 0.49 to 1.29) suggested clinically relevant variation in treatment effects. Accordingly, one downgrade was applied.

III. Indirectness was judged not serious, as all included studies addressed the relevant population, intervention, and outcome.

IV. Imprecision was considered not serious. The overall effect estimate (HR = 0.80, 95% CI [0.71, 0.90]) was precise and did not cross the threshold of no effect.

V. Reporting bias was judged not serious, as multiple robust detection methods failed to find evidence of publication bias or selective reporting.

In conclusion, the GRADE approach identified some inconsistency, leading to a moderate certainty rating regarding the beneficial effect of psychosocial interventions on survival in cancer patients.

**Discussion**

**Overall effect and heterogeneity**

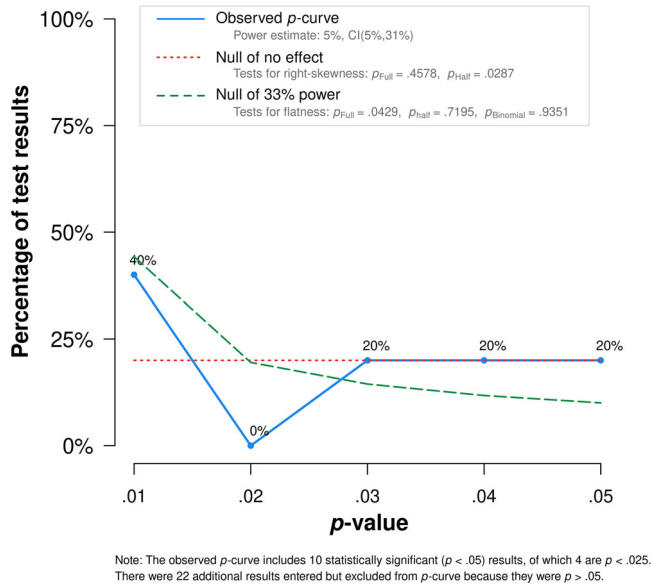
We found evidence for a small but robust overall effect HR 0.80, 95% CI [0.71, 0.90], suggesting that psychosocial interventions prolong the survival of cancer patients. Moderate to substantial heterogeneity was present across the distribution of true effects, resulting in a wide 95% PI of HR 0.49 to 1.29. This indicates that future primary studies could (potentially) observe effects ranging from substantial benefit to slightly unfavorable in certain populations or under particular conditions. Importantly, none of the included RCTs reported a significant negative impact on survival time.

The observed heterogeneity may, in part, be attributable to the clinical and methodological diversity of the included trials. Given the limited availability of RCTs in this research area, it was necessary to incorporate a broad range of psychosocial interventions, cancer types, and tumor stages. While this broad inclusion strategy may enhance generalizability, it inevitably introduces additional heterogeneity.

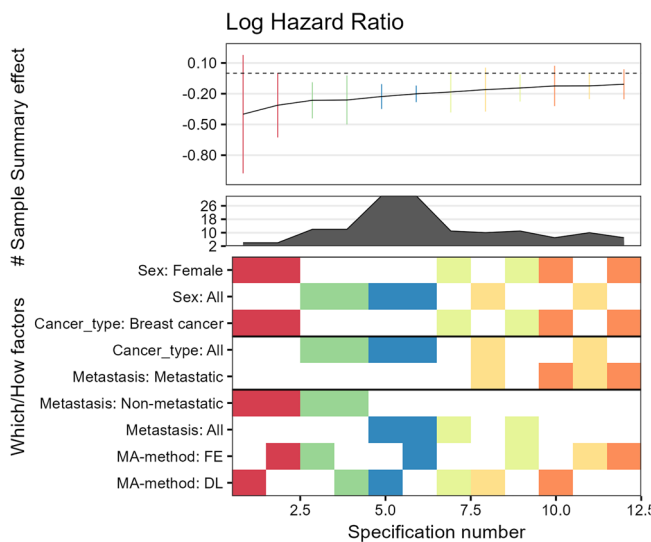
In addition, many of the trials were underpowered to detect effects of the magnitude of the overall effect observed in our current analysis, which may have contributed considerably to greater variability and imprecision. Consequently, the certainty of the evidence was rated as moderate, reflecting the current body of research while adhering to scientific rigor.

**An easy-to-grasp effect size for survival**

The gain in survival time was also estimated by an easy-to-grasp effect size to enhance interpretability for clinicians and patients: The overall median survival (time) benefit for treated patients compared to controls showed a trend of 3.9 months, 95% CI [-0.7, 8.5], although this metric was not used



**Fig. 6 | *p*-Curve displaying the distribution of all ten significant *p* values from HRs.** Right-skewness (indicating evidential value of the corpus of primary studies) was present, but the observed *p*-curve (blue line) proceeded close to the reference line of low power of 33% (green dashed line). Left-skewness (indicating *p*-hacking) was not evident,  $k = 10$  significant *p* values.



**Fig. 7 | Descriptive meta-analytic specification curve plot for logHRs across all meta-analytic specifications from prior meta-analyses using the currently available corpus of primary studies.** LogHR < 0 indicates a survival benefit of the intervention group; logHR > 0 indicates a survival benefit of the control group; logHR = 0 indicates no effect. Interpret the curve vertically. The top panel shows each specification’s overall effect (horizontal line) and 95% CI (vertical line) relative to the dashed no-effect line (logHR = 0). The middle panel shows the number of trials per specification. The bottom panel depicts the combination of “How” and “Which” factors for each specification. Colors indicate sample size, from larger (violet, blue, green) to smaller (yellow, orange, red). FE fixed effects, DL DerSimonian and Laird method,  $k = 32$  studies.

for hypothesis testing, as it could only be based on a subgroup of trials ( $k = 16$ ) that provided the necessary data. However, for cancer patients and their families, an estimated survival gain of four months—particularly if accompanied by preserved or improved quality of life—may represent an invaluable asset.

**Clinical considerations of psychosocial interventions for cancer patients**

**Comparable survival benefits to medical cancer treatments.** Contextualizing our findings within the broader oncology literature, the observed overall effect HR of the present meta-analysis is, indeed, clinically meaningful. Interestingly, it closely mirrors the effect of being married versus unmarried, HR of 0.80, 95% CI [0.79–0.81], in a large-scale 5-year epidemiological study of 734,889 cancer patients with ten major cancer types<sup>131</sup>.

Moreover, our observed overall effect in prolonging survival appears comparable to those reported for established medical cancer treatments. Recent meta-analyses on chemotherapy for pancreatic cancer and metastatic breast cancer reported small overall effect sizes with an HR of 0.71, 95% CI [0.62–0.82]<sup>132</sup> and an HR of 0.87, 95% CI [0.78–0.97]<sup>133</sup>, respectively. A meta-analysis on radiotherapy in breast cancer did not find a statistically significant overall effect with an RR of 0.93, 95% CI [0.79, 1.09]<sup>134</sup>, whereas a meta-analysis of radiotherapy in gallbladder cancer showed a small overall effect with an OR of 0.63, 95% CI [0.50–0.81]<sup>135</sup>. A meta-analysis on hormone therapy in breast cancer showed a small overall effect with an HR of 0.79, 95% CI [0.69, 0.90]<sup>136</sup>.

Note that the categorization of small, moderate, and large effect sizes differs between HRs and ORs (see Supplementary Table C4); to our knowledge, no universally accepted cutoffs exist for RRs.

These comparisons demonstrate that the survival benefit associated with psychosocial interventions for cancer patients, when provided on top of standard cancer treatments, lies within the same order of magnitude as that of commonly accepted chemo-, radio-, and hormone therapy, yet these interventions come without the adverse side effects typically associated with medical cancer treatments. Accordingly, psychosocial cancer care can no longer be considered optional or purely supportive from a scientific standpoint.

**Comparable certainty in the evidence to medical cancer treatments.**

Although meta-analyses of medical cancer treatments often draw from a larger pool of studies, which might allow for more focused and homogeneous estimates, a closer look reveals a different picture. For instance, the 95% PIs of the mentioned meta-analyses on chemotherapy (95% PI [HR 0.58 to 0.87]<sup>132</sup>; 95% PI [HR 0.66 to 1.15]<sup>133</sup>), radiotherapy (95% PI [RR 0.72 to 1.21]<sup>134</sup>; 95% PI [OR 0.26 to 1.54]<sup>135</sup>), and hormone therapy (95% PI [HR 0.49 to 1.28]<sup>136</sup>) frequently crossed the line of no effect, and in some cases could even indicate unfavorable survival outcomes in future studies.

It is worth noting that we often had to calculate the 95% PIs ourselves for the meta-analyses on chemo-, radio-, and hormone therapy, as these were frequently not reported directly (see Supplementary Table C3 for data on how the 95% PIs were calculated). This omission may partly account for the apparent lack of awareness that the level of certainty in the scientific evidence regarding medical cancer treatments is actually lower than one might expect.

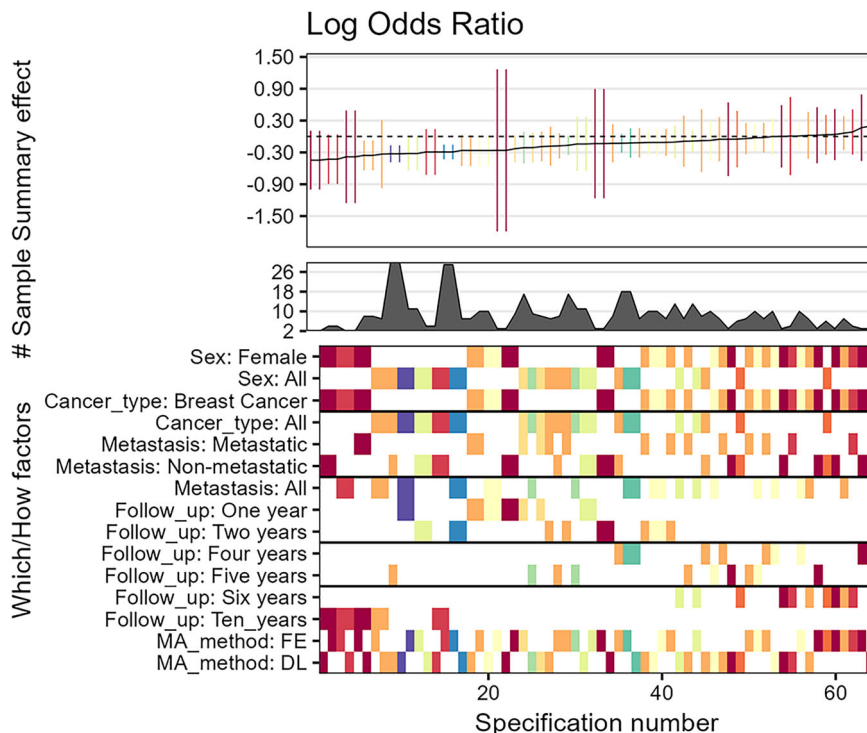
Nonetheless, these kinds of treatments continue to be administered as standard care for cancer patients despite known adverse side effects such as anemia, alopecia, anxiety, cognitive dysfunction, fatigue, insomnia, immunosuppression, nausea and vomiting, low mood, mucositis, pain, and weight loss<sup>137–139</sup>. Whereas psychosocial interventions appear to achieve comparable survival benefits with similar certainty in the evidence but without these harms, they are still frequently relegated to a merely supportive role—highlighting a critical evidence practice gap in oncology.

**Additional benefits of psychosocial treatments for cancer patients.**

Furthermore, psychosocial interventions are unique in offering additional benefits beyond survival. Several meta-analyses have shown that these kinds of interventions simultaneously improve well-being<sup>140</sup>, quality of life<sup>141</sup>, anxiety<sup>142</sup>, depression<sup>143</sup>, stress<sup>144</sup>, fatigue<sup>145</sup>, and pain<sup>146</sup> in cancer patients. This multifaceted benefit profile clearly distinguishes

**Fig. 8 | Descriptive meta-analytic specification curve plot for logORs across all meta-analytic specifications from prior meta-analyses using the currently available corpus of primary studies.**

LogOR <0 indicates a survival benefit of the intervention group; logOR >0 indicates a survival benefit of the control group; logOR = 0 indicates no effect. Interpret the curve vertically. The top panel shows each specification's overall effect (horizontal line) and 95% CI (vertical line) relative to the dashed no-effect line (logOR = 0). The middle panel shows the number of trials per specification. The bottom panel depicts the combination of "How" and "Which" factors for each specification. Colors indicate sample size, from larger (violet, blue, green) to smaller (yellow, orange, red). The "How" factor Follow-up was added here, as logORs in the prior analyses were calculated after 1-, 2-, 4-, 5-, 6-, and 10-year intervals for every trial that provided data (in contrast to logHRs). FE fixed effects, DL DerSimonian and Laird method,  $k = 32$  studies.



psychosocial care from standard medical therapies, which primarily target tumor progression and survival, often at the expense of well-being and quality of life.

On top of that, psychosocial interventions may act as catalysts for broader biopsychosocial changes in cancer patients. Although they do not directly target malignant cells, they can influence psychological (e.g., thoughts, emotions), behavioral (e.g., social interactions, health behaviors, treatment adherence), and physiological processes, including immune, endocrine, and stress responses<sup>5-7,147-149</sup>—all of which could contribute to improved survival outcomes.

Therefore, the modest overall HR observed in this meta-analysis may underestimate the broader causal pathways through which psychosocial interventions exert their effects across psychological, behavioral, and biological domains.

**Recognizing the role of psycho-oncology in clinical decision-making.** It is notable that in oncology, medical treatments with modest or even borderline effect sizes are routinely embraced—despite high financial cost, considerable toxicity, and high psychological as well as physiological burden. In contrast, psychosocial interventions often remain sidelined, even though they are low-cost, low-risk, and scalable, while simultaneously improving psychological well-being, enhancing immune and endocrine function, and promoting healthier behaviors.

This discrepancy reveals a persistent bias: the prevailing biomedical model continues to undervalue interventions that do not fit neatly into pharmacological or surgical paradigms, despite their capacity to influence survival through interconnected biopsychosocial pathways.

Our findings challenge this framing, as the identified survival benefits of psychosocial interventions for cancer patients are not only statistically and clinically meaningful—they appear to be comparable in both magnitude and certainty of evidence to many widely accepted medical cancer therapies, such as chemo-, radio-, and hormone therapy, while also carrying minimal risk and lower costs.

Thus, even from a purely economic standpoint, psychosocial interventions deserve greater recognition and broader implementation. For example, median costs for health insurance of standard chemotherapy among breast cancer patients in the United States (between 2008 and 2012)

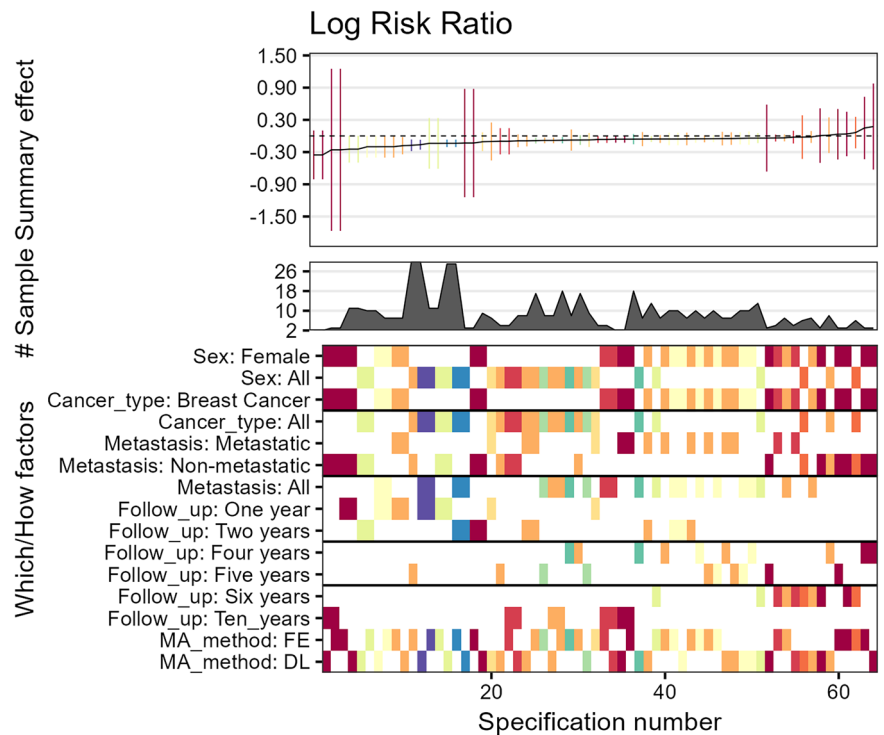
varied substantially by treatment type, ranging from \$62,302 (CMF regimen) to \$104,197 (TAC regimen) per patient and patient out-of-pocket expenses from \$2104 to \$3383, respectively—which represents a considerable financial burden incurred on the health care system and patients<sup>150</sup>. Median costs for one course of radiotherapy among breast cancer patients in the United States (between 2000 and 2009) were \$8100<sup>151</sup>.

In contrast, average costs for standard psychosocial interventions during a comparable time frame were substantially lower: such as six weekly 120-min sessions of mindfulness-based stress reduction in the United States (2015) cost \$666 per patient or \$111 per patient and session<sup>152,153</sup>, six weekly 90-min sessions of cognitive-behavioral group therapy in the Netherlands (2008) cost \$406 (€311.85) per patient or \$45 (34.65€) per patient and session<sup>153,154</sup>, and around 25 weekly 90-min sessions of supportive-expressive group therapy in Canada (2006) cost \$1394 per patient or \$55.46 per patient and session<sup>153,155</sup>.

In sum, these findings highlight the urgent need for guideline committees and oncology policymakers to formally establish psychosocial (psychological) care alongside medical therapy as a standard component of cancer treatment, fostering a more integrative, patient-centered approach that enhances survival, psychological well-being, and quality of life, while remaining cost-effective.

**Psychological components in psychosocial interventions for cancer patients.** The analysis of psychological components included in the psychosocial interventions should be interpreted exploratorily and descriptively, as all interventions comprised multiple components. Consequently, no causal inferences regarding the effectiveness of individual components can be drawn. Rather, the observed frequency patterns suggest that effective psychosocial interventions commonly integrate educational components (such as medical education and psychoeducation), cognitive-behavioral strategies (e.g., coping skills/techniques, relaxation techniques, problem-solving, cognitive reappraisal/restructuring, goal setting), and emotion-focused elements (e.g., expression of emotions, meaning-based components), alongside other supportive elements (e.g., group support, other social support, aid with decision making, aid with symptom management). The observed patterns show some heterogeneity in component reporting, suggesting that

**Fig. 9 | Descriptive meta-analytic specification curve plot for logRRs across all meta-analytic specifications from prior meta-analyses using the currently available corpus of primary studies.** LogRR <0 indicates a survival benefit of the intervention group; logRR >0 indicates a survival benefit of the control group; logRR = 0 indicates no effect. Interpret the curve vertically. The top panel shows each specification's overall effect (horizontal line) and 95% CI (vertical line) relative to the dashed no-effect line (logRR = 0). The middle panel shows the number of trials per specification. The bottom panel depicts the combination of "How" and "Which" factors for each specification. Colors indicate sample size, from larger (violet, blue, green) to smaller (yellow, orange, red). The "How" factor Follow-up was added here, as logRRs in the prior analyses were calculated after 1-, 2-, 4-, 5-, 6-, and 10-year intervals for every trial that provided data (in contrast to logHRs). FE fixed effects, DL DerSimonian and Laird method,  $k = 32$  studies.



component overlap and sometimes unclear definitions may limit comparability across trials.

From a clinical and intervention-design perspective, the findings support the use of multicomponent approaches and may inform the development of future psychosocial interventions for cancer patients, while emphasizing the need for formal moderator or dismantling studies. Overall, our findings highlight the importance of theory-driven intervention design and standardized reporting of psychological components.

**Methodological considerations of the meta-analysis and multi-verse meta-analysis**

**Small-study effects.** The pattern observed in the forest plot sorted by study precision did not fully correspond to what would be expected from typical small-study effects or publication bias. While the least precise studies tended to report the largest effects (HRs <1), a consistent trend of HRs below 1 was also evident among studies with moderate to good precision. This suggests that the overall survival benefit was not solely driven by small-study effects or selective publication and supports the robustness of psychosocial interventions, which contradicts claims that observed benefits are merely artefacts of study design or selective reporting.

Instead, this finding likely reflects a genuine effect, although the variation among the most precise studies indicated that additional factors—such as patient characteristics, intervention protocols, or differences in study design—may have contributed to the observed heterogeneity.

**Time trend.** A temporal trend appeared to emerge when considering the progression of studies over time. In studies published before 2015, HRs were mixed, with values both above and below 1, indicating inconsistent evidence for the existence of an overall survival benefit. However, beginning with 2015, studies consistently reported HRs <1, suggesting more uniform evidence for an overall survival benefit.

Supporting this interpretation, the cumulative meta-analysis demonstrated that only after 2015 did the growing number of available trials provide enough statistical power and precision for the overall effect to reach statistical significance. Consequently, meta-analyses performed prior to this

date could not reliably detect the (small) overall effect estimated by our current analysis, even if they had used the HR as the effect size, which is already a more precise effect measure than the OR and RR.

This shift likely reflects more than just increased sample sizes; it may also represent a systematic change in effect sizes over time, driven by evolving patient populations, improved study designs, and more rigorous reporting and publication standards. Together, these factors strengthen the clearer signal-to-noise ratio, while changes in medical treatment contexts may further moderate psychosocial intervention effects.

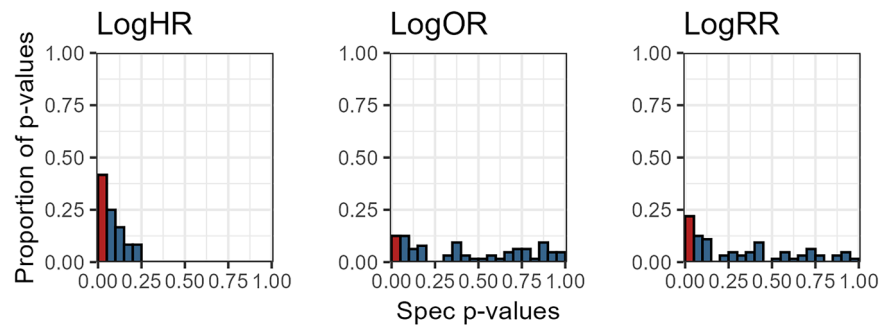
**Risk of bias factors.** Overall risk of bias of studies ranged from moderate ( $k = 30$ ) to high ( $k = 2$ ), mostly due to a lack of information on allocation concealment or absence of a preregistration protocol, which reflects the fact that trials were published over the span of four decades (from 1982 to 2023) and specific reporting guidelines were only established from the mid-2000s or mid-2010s on, respectively.

Average study quality was moderate ( $M = 7$ ; range  $-18$  to  $+18$  points) and risk of sponsorship bias was low for most studies, as no for-profit organization funding was involved, whereas average meta-analytic post hoc power of included RCTs, using the overall effect HR of our meta-analysis as the population effect, was very low at 17%. In general, a (post hoc) power of at least 80% would be considered adequate, otherwise not significant (but also significant) results cannot be trusted with high confidence, due to high statistical error probabilities<sup>156,157</sup>. This implies that most trials could only reliably detect considerably larger effects than the small overall effect found across the corpus of primary studies. This may not only explain the inconsistency of the results of individual trials but (in part) also of the prior meta-analyses reviewed.

Consequently, we propose that future trials on cancer survival should be designed with sufficient power to identify even small but still clinically relevant effects, or definitively disprove them. Although, it should be noted that none of the risk of bias factors seemed to significantly moderate the effects of included trials.

**Moderators.** The investigation of the 25 substantive and methodological moderators identified one significant moderator, namely analysis type. Studies that used ITT analysis showed a significant overall effect (HR =

**Fig. 10 | Histograms of  $p$  value distributions for observed overall effects (logHRs, logORs, and logRRs) across all meta-analytic specifications from prior meta-analyses.** The proportion of significant  $p$  values ( $p < 0.05$ ) is indicated by the red columns, while the proportions of not significant  $p$  values ( $p > 0.05$ ) are marked with the blue columns in each histogram,  $k = 32$  studies.



0.74, 95% CI [0.65, 0.85]) compared to studies that did not use ITT analysis (HR = 0.96, 95% CI [0.78, 1.18]). The criteria for ITT analysis were very strict in our present meta-analysis and did not allow trials to exclude any participants after randomization—even if they were later found to meet predefined exclusion criteria, which would typically qualify as a modified ITT analysis. In addition, this finding did not withstand the Benjamini–Hochberg procedure and should therefore be interpreted with appropriate caution.

Surprisingly, none of the moderators identified by prior meta-analyses (see Table 1) such as age and marital status<sup>40</sup>, intervention type<sup>68</sup>, or intervention format<sup>158,159</sup>—could be replicated by our analysis. Moreover, several moderators reported by ref. 68, for which the investigated subgroups actually had overlapping 95% CIs, such as cancer type, presence of metastasis, intervention provider, and length of follow-up—could also not be confirmed.

This may relate to the number of available studies being relatively limited in the light of the clinical and methodological heterogeneity present in our analysis or to the possibility that some moderators of prior meta-analyses reviewed may have been identified using fixed-effect models. However, this assumption could not be verified, as information on the applied statistical model was most often lacking. If fixed-effect models had been used, the generalizability of such findings should be viewed with caution anyway<sup>38</sup>.

**Exploratory search for moderators.** Exploratory moderator analysis with the machine learning and bootstrap-based approach MetaForest identified three moderators, namely analysis type, tumor phase, and study quality, with consistently positive but low variable importance ( $0.01 < VI < 0.05$ ) and showed a weak overall predictive performance in unseen data ( $R^2_{\text{ob}} = 0.12$ ) as well as a slightly higher predictive performance using cross-validation ( $R^2_{\text{cv}} = 0.28$ ).

This suggests that these variables might be promising candidates for future hypothesis-driven research. However, the discrepancy between out-of-bag performance (12%) and cross-validation (28%) indicated some degree of overfitting, which emphasizes the need for replication in future independent datasets. As such, the current findings provide preliminary evidence for potential moderators, but should not be overinterpreted as conclusive.

**Reporting bias across studies.** Visual inspection of the contour-enhanced funnel plot revealed some indication of small-study effects (or publication bias), as studies with larger standard errors appeared to be missing in the lower right area of the plot, where studies with effect sizes in the direction opposite to the favored outcome would typically be located.

The  $p$ -uniform test, conducted on significant  $p$  values ( $k = 10$ ), was not statistically significant ( $p = 0.266$ ), indicating no clear evidence of publication bias. However, given the limited number of significant  $p$  values, the test may have lacked power to detect publication bias reliably, and this outcome should be interpreted with caution. Likewise, the funnel asymmetry test (FAT), conducted as part of the PET-PEESE procedure using all 32 effect

sizes, did not yield statistically significant evidence for publication bias either ( $p = .091$ ), but only suggested a trend toward asymmetry.

Hence, given the moderate to substantial heterogeneity ( $I^2 = 48\%$ ;  $\tau^2 = 0.05$  [log scale];  $\tau = 0.23$  [log scale]) and relatively small sample sizes in some studies, it is possible that true heterogeneity rather than publication bias explains the funnel plot asymmetry<sup>160</sup>. In light of this, the current results do not provide robust evidence of publication bias, but they do not entirely rule it out either. Visual indication of publication bias in the funnel plot, a non-significant  $p$ -uniform result with limited power, and a non-significant trend in the FAT all in the presence of heterogeneity collectively call for a cautious interpretation of the present findings. In addition, there was no indication of  $p$ -hacking (no left-skewness) observable in the  $p$ -curve.

However, the statistical power of the  $p$ -curve may have been low too, as only the minimum recommended number of significant  $p$  values ( $k = 10$ ) was available for  $p$ -curve analysis<sup>129</sup>, limiting the conclusiveness of these results.

#### Causes for the heterogeneity in the results of prior meta-analyses.

Descriptive specification curve plots indicated that the contradictory results of prior meta-analyses reviewed by our analysis (four reporting a survival benefit, whereas six reporting no benefit) were a consequence of the application of different effect size measures with varying degrees of precision (HR vs. OR vs. RR) as well as the use (or nonuse) of various selection criteria for the inclusion of studies (e.g., a focus on subpopulations such as [non-]metastatic female breast cancer patients or on different follow-up endpoints such as 1, 2, 4, 5, 6, or 10 years), which often resulted in a small number of trials to be analyzed within a given specification and thereby limiting statistical power—especially in the presence of (already) small overall effects. Consequently, meta-analytic specifications, which included a higher amount of trials (those with less or without any restrictions), frequently resulted in significant overall effects.

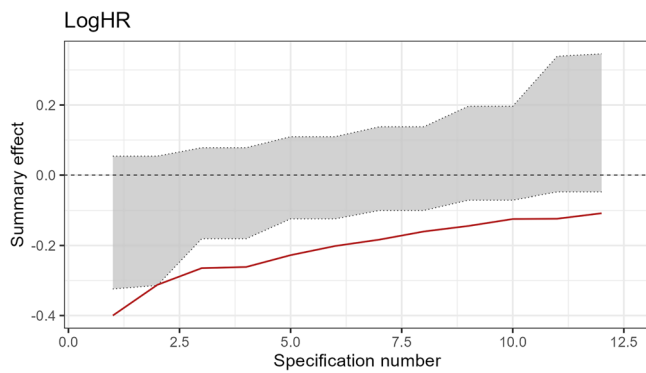
Strikingly, the amount of specifications with significant overall effects (logHRs ~42%, logORs ~13%, and logRRs ~22%) was much higher than what would be expected if a *general* zero overall effect was true (in that case it would be ~5% for each effect size), which emphasized the existence of a *general* true (nonzero) overall effect across reasonable meta-analytic specifications.

On top of that, inferential specification curve plots (bootstrap-based approach) confirmed that most overall effects of the logHR as well as some of the overall effects of the 1-, 2-, and 10-year follow-ups of the logOR and logRR deviated significantly from 1000 simulated true zero overall effects and their 95% CI bands.

Taken together, these results provide strong evidence that the observed overall effects are unlikely to be statistical artefacts and instead reflect a consistent pattern across a broad range of reasonable meta-analytic choices.

#### Limitations

**Overall effect.** Meta-analysis is a powerful tool to address certain research questions, but it can be misleading in the face of clinical and



**Fig. 11 | Inferential specification curve plot of observed overall effects (logHRs) across reasonable meta-analytic specifications from prior meta-analyses (red line) against the 95% CI band of 1000 simulated true-zero effects (gray area) using the currently available corpus of primary studies.** LogHR <0 indicates a survival benefit of the intervention group; logHR >0 indicates a survival benefit of the control group. The black dotted line indicates a zero effect (logHR = 0). The red line shows the actually observed overall effects across specifications of prior meta-analyses (sorted by size) plotted against the 95% CI band of 1000 simulated overall effects for which a zero effect was true (gray area). When the red line falls outside the gray area, it suggests evidence for a robust and generalizable true overall effect across reasonable meta-analytic specifications,  $k = 32$  studies.

statistical heterogeneity<sup>55,56</sup>. Both forms of heterogeneity were present in the current and all prior meta-analyses reviewed (see Table 1), which could mean that the overall effect might not be overly meaningful<sup>161</sup>.

Investigating samples from one and the same population (e.g., same cancer type and stage) receiving the exact same intervention in a single meta-analysis, while conducting separate analyses for the other populations, could be considered an ideal meta-analytic goal. It could provide the most reliable conclusions of what works for whom.

However, this ideal goal was not feasible with respect to the research question, as currently there were not enough primary studies available to follow such a strict approach, which may partly explain the moderate to substantial heterogeneity and wide 95% PI evident in our meta-analysis. Instead, our approach encompassed different kinds of psychosocial interventions, cancer types, and tumor stages while applying rigorous methods of modern evidence synthesis to address possible sources of bias directly and thereby limiting researchers' degrees of freedom.

**Data collection process.** The analysis had the following methodological limitations. First, the search was restricted to studies published in English, which may have introduced language bias. Second, although we contacted study authors to obtain missing or unclear data, some did not respond, potentially limiting the completeness of our data extraction. Finally, despite efforts to include gray literature, there was still an indication of potential publication bias among the included studies.

**Calculation of effect size.** The proportional hazards (PH) assumption, that the HR between two groups remains constant over time, is vital for the Cox regression model. Both primary studies and previously published meta-analyses have adopted this assumption.

Due to the lack of individual patient data, we have assessed the PH assumption by visually inspecting Kaplan–Meier survival curves, a standard method when individual data are unavailable. Established guidelines suggest that if these curves do not cross and maintain a (somewhat) constant distance over time, the PH assumption is likely valid.

In our analysis, no major violations have been observed but we acknowledge the limitations of the visual inspection approach, which, although widely used in the absence of individual data, may not be as robust as formal statistical tests.

**Moderators.** In general, it is recommended that subgroup analyses should only be conducted if there are at least ten studies included in the meta-analysis to provide sufficient statistical power and yield meaningful results<sup>93</sup>. However, a minimum number of studies required for each subgroup is not specified, and caution is advised when studies are unevenly distributed among subgroups.

Additionally, it is recommended that meta-regressions include at least ten studies to minimize the risk of overfitting and ensure that the results can be interpreted meaningfully. Therefore, results of meta-regression analyses involving fewer than ten studies should also be interpreted with caution.

Consequently, the statistical power to reliably detect certain effect moderators in subgroup or meta-regression analysis might have been low, and any null findings should thus be considered as preliminary and interpreted accordingly (“Absence of evidence is not evidence of absence.”<sup>162</sup>).

Regarding our exploratory moderator analysis with MetaForest, we used simple imputation (median or mode) for variables with up to 10% of missing data. While this approach does not account for the uncertainty inherent in imputed values and may have attenuated true associations between moderators and effect sizes, multivariate imputation was not a viable option in the present context. Most moderators were only weakly intercorrelated, and MetaForest cannot deal with multiple imputed datasets. However, to reduce bias, all moderators with more than 10% of missing values were excluded from the analysis.

**Reporting bias across studies.** Several limitations apply to the methods used. First, the contour-enhanced funnel plot suggested potential for a minor small-study effect (or publication bias), though its visual nature limits its reliability.

Second, the non-significant  $p$ -uniform test, which included only  $k = 10$  significant  $p$  values, had low statistical power, reducing its ability to detect publication bias.

Third, the funnel asymmetry test, while indicating a trend for asymmetry, may have been affected by moderate to substantial heterogeneity ( $I^2 = 48\%$ ;  $\tau^2 = 0.05$  [log],  $\tau = 0.23$  [log]) and small sample sizes rather than an actual publication bias.

Lastly, the  $p$ -curve analysis was constrained by the minimum number of significant  $p$  values available ( $k = 10$ ), which is the lower bound for meaningful analysis. Additionally,  $p$ -curve can systematically overestimate evidential value, especially in the presence of heterogeneous effect sizes or low-powered studies<sup>129</sup>. Under such conditions,  $p$ -curve is known to sometimes overestimate evidential value and has low statistical power, limiting the robustness and generalizability of the inferences.

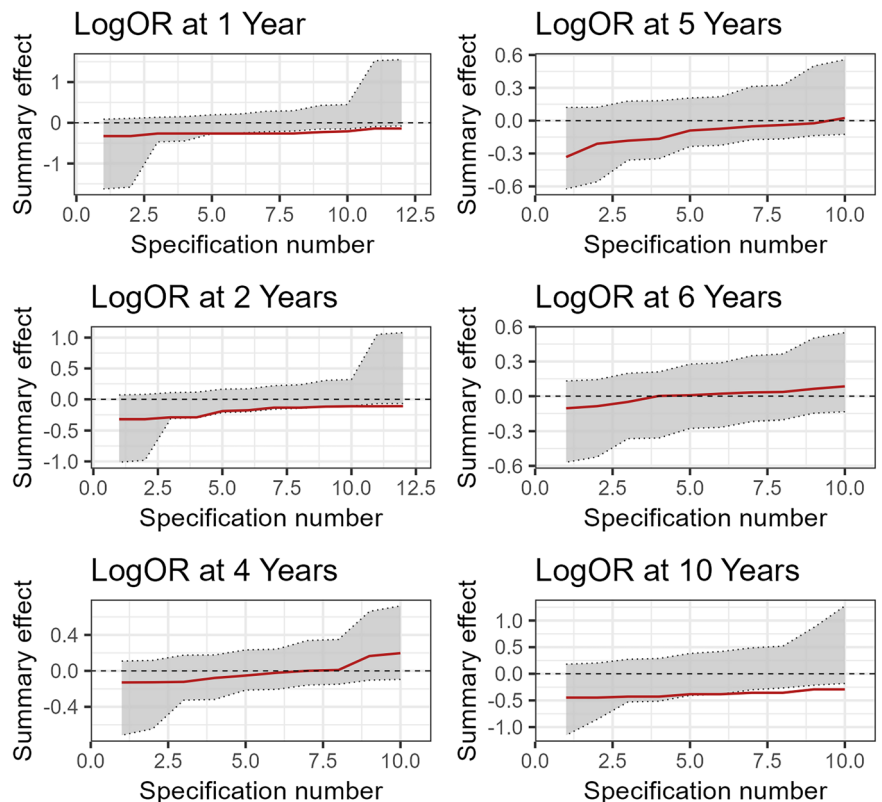
Consequently, the absence of detected  $p$ -hacking should be interpreted cautiously, and these results should be considered preliminary and indicative rather than definitive.

**Lack of control for treatment confounding.** Regarding the history of research in the field of psycho-oncology, one must also consider that patient populations might have changed over the years. Nowadays, cancer patients may be much more willing to ask for psychological support themselves, compared to four decades ago, which might have led control patients in more recent studies to attend support outside the study protocol more frequently than in older studies, which may mitigate the effects to some degree<sup>20</sup>.

Most included studies (27 out of 32) did not provide information on the use of psychological or psychotherapeutic services during the post-intervention follow-up period. As a result, it remains unclear to what extent participants in either group may have received additional psychosocial care outside the study protocols. Only five studies<sup>18,20,163–165</sup> monitored such follow-up care and checked for group differences, but reporting was limited in the other studies and detailed data were mostly unavailable.

This lack of control may have introduced treatment confounding over time and could have attenuated the observed effects, particularly if participants in the control group sought out supportive care after the study as well. In this sense, the identified overall effect of our analysis may only reflect

**Fig. 12 | Inferential specification curve plot of observed overall effects (logORs) across reasonable meta-analytic specifications from prior meta-analyses (red line) against the 95% CI band of 1000 simulated true-zero effects (gray area) using the currently available corpus of primary studies (1-, 2-, 4-, 5-, 6-, and 10-year follow-ups). LogOR <0 indicates a survival benefit of the intervention group; logOR >0 indicates a survival benefit of the control group. The black dotted line indicates a zero effect (logOR = 0). The red line shows the actually observed overall effects across specifications of prior meta-analyses (sorted by size) plotted against the 95% CI band of 1000 simulated overall effects for which a zero effect was true (gray area). When the red line falls outside the gray area, it suggests evidence for a robust and generalizable true overall effect across reasonable meta-analytic specifications,  $k = 32$  studies (for every panel).**



a conservative estimate. Likewise, medical treatment of cancer patients has advanced in the past decades (e.g., the development of hormone therapy), which may benefit certain populations of cancer patients (those with ER/PR-positive tumors) but not others (those with ER/PR-negative tumors). In this context, it has been argued that the effects of hormone therapy might supersede the effects of psychosocial interventions in the former group of patients in recent studies<sup>20</sup>.

**Ethical constraints.** The entire field of psycho-oncologic research is challenged by the ethical and practical difficulties of designing RCTs that can definitively answer whether psychosocial interventions prolong survival while being ethical and upholding scientific rigor. Strict designs would require long-term follow-ups, strict control over co-interventions, and comparisons to minimal or no-treatment groups—conditions that are ethically problematic when psychological support is known to have substantial benefits for emotional well-being.

Withholding such support for the sake of scientific rigor may be considered unethical, particularly in vulnerable patient populations. As a result, many existing RCTs rely on active control groups, which reduces the risk of ethical violations but may attenuate observable survival effects. This trade-off between methodological rigor and ethical responsibility represents an inherent limitation in interventional research and complicates causal interpretations.

We have addressed this issue to the extent possible by including the type of control group as a moderator in our analysis; however, fully disentangling its impact remains difficult given the necessary ethical constraints.

### Suggestions

**Suggestions for future primary research.** The problem of underpowered studies is well known. About 70% of all meta-analyses might be solely based on studies too underpowered to detect a small effect<sup>49</sup>. Reasons for this practice often lie in difficulties recruiting participants because of rare diseases or limited amounts of time and resources. However, from a statistical point of view, underpowered studies are

simply ill-suited to clarify pending research questions<sup>166</sup> and future RCTs should thus be planned in such a manner that even small but clinically meaningful effects can be detected with sufficient power (i.e.,  $\geq 80\%$ ).

Suggestions on how to deal with this power problem include the following:

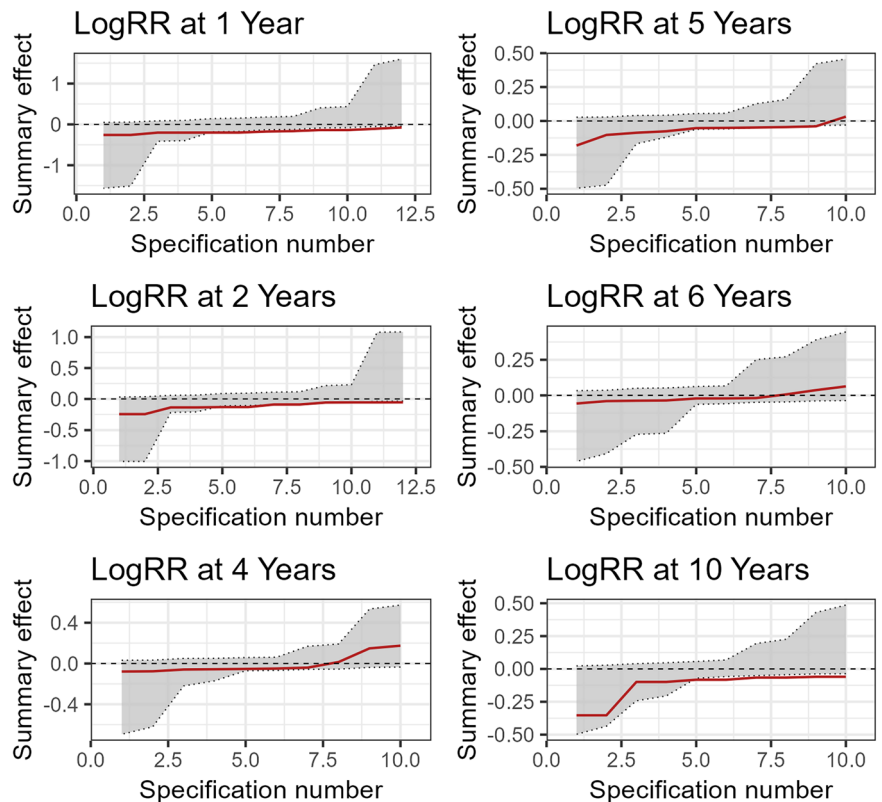
(1) Increasing collaborative efforts, as this would enhance statistical power (by increasing sample size) as well as external validity of findings (by including diverse populations<sup>167</sup>).

(2) RCTs that solely investigate the psychological effects of psychosocial interventions in cancer patients should also collect data on survival (and tumor progression) on a regular basis. This could even be done retrospectively for trials that have already been conducted with comparably little effort and would yield invaluable information.

(3) From a statistical perspective, investigating more aggressive types of cancer, which may generate more events over the same period and thereby increase statistical power—although it still remains an open question whether psychosocial interventions prolong the survival of patients with more aggressive cancers as well. Besides that, primary studies should also conduct more replications and should focus on the most common cancer types (e.g., breast, lung, colorectal, prostate, and stomach<sup>1</sup>) and certain stages (e.g., stage I and IV), which would enable secondary research to draw more definitive conclusions about what works for whom.

Moreover, only a handful of trials provided data on some important baseline characteristics investigated by our analysis (e.g., estrogen and progesterone receptor status, lymph node status, or prior hormone therapy) and (almost) none of the trials delivered data on patient characteristics after the start of the experiments (e.g., subsequent surgery, chemo-, radio-, or hormone therapy); although there is evidence from the original landmark study of Spiegel et al. in 1989<sup>17</sup> that medical therapy given to cancer patients after randomization did not differ and therefore could not explain the benefits of survival for the intervention group in this trial<sup>168</sup>. However, for a clearer picture regarding moderating factors, it would be necessary for trials to collect these very important data until the end of the study.

**Fig. 13 | Inferential specification curve plot of observed overall effects (logRRs) across reasonable meta-analytic specifications from prior meta-analyses (red line) against the 95% CI band of 1000 simulated true-zero effects (gray area) using the currently available corpus of primary studies (1-, 2-, 4-, 5-, 6-, and 10-year follow-ups). LogRR <0 indicates a survival benefit of the intervention group; logRR >0 indicates a survival benefit of the control group. The black dotted line indicates a zero effect (logRR = 0). The red line shows the actually observed overall effects across specifications of prior meta-analyses (sorted by size) plotted against the 95% CI band of 1000 simulated overall effects for which a zero effect was true (gray area). When the red line falls outside the gray area, it suggests evidence for a robust and generalizable true overall effect across reasonable meta-analytic specifications,  $k = 32$  studies (for every panel).**



Furthermore, future RCTs should routinely assess and report the use of additional psychosocial care outside the study intervention during follow-up. Only a handful of trials in the present meta-analysis provided information on whether participants sought or received other psychological support after the intervention phase.

Without such data, it remains difficult to draw firm conclusions about the long-term effects of the studied interventions. Regular monitoring of external psychosocial care—through self-report, medical records, or standardized check-ins—would help clarify whether observed effects are attributable to the intervention itself or potentially confounded by post-treatment support.

**Suggestions for future meta-analytic research.** The growing number of often low-quality systematic reviews and meta-analyses in recent years represents a threat to the actual validity of scientific research and is driven by various contributors with different motives<sup>42</sup>. It has been argued that the majority of meta-analyses are either flawed, misleading, redundant, not useful, or remain unpublished<sup>42</sup>.

Therefore, comprehensive approaches may offer a useful step toward a more valid overall picture, especially in research fields where the existing evidence is fragmented or inconclusive. Novel approaches, such as specification curve and multiverse meta-analysis, may help researchers to uncover the causes for controversial results of prior meta-analytic research<sup>77</sup>. In essence, we would like to encourage fellow meta-analytic researchers to adopt contemporary methods of evidence synthesis in their future work.

**Suggestions for policymakers, clinicians, and patients.** Our findings suggest that psychosocial (psychological) interventions should be made accessible to cancer patients whenever this seems feasible and encourage policymakers to take appropriate action, as psychosocial interventions appear to offer survival benefits on par with, and on top of, current medical cancer treatments; have no known adverse side effects; show no evidence of

harm; provide substantial psychological benefits—including improved well-being and quality of life as well as reductions in anxiety, depression, stress, fatigue, and pain; are of relatively low cost compared to conventional treatments, and are supported by converging evidence from epidemiological, experimental, and meta-analytic research. We therefore also suggest that patients actively engage in psychosocial, psychological, or psychotherapeutic interventions, as their mental, emotional, and physical health—and that of their families—may be affected considerably.

Psychosocial interventions analyzed in our meta-analysis were typically multicomponent and commonly included educational components (e.g., medical education and psychoeducation), cognitive-behavioral strategies (e.g., coping skills, relaxation techniques, cognitive restructuring), emotion-focused approaches (e.g., expression of emotions, meaning-based components), and social support (e.g., group support). These overarchingly categorized components may serve as practical reference points for the design and structuring of multicomponent psychosocial programs for cancer patients, without implying any differential effectiveness of individual elements. The observed inclusion patterns highlight frequently applied elements that could guide hypothesis-driven interventional designs in future trials. Overall, these insights emphasize how the field can benefit from standardized reporting and systematic evaluation of component combinations, providing a framework to inform future research and intervention design.

#### **A proposal for a future research topic**

The psychosocial interventions examined in our present meta-analysis primarily aimed to support cancer patients during a time of acute psychological distress following diagnosis and treatment. However, current psychobiological research suggests that the onset and progression of cancer may be influenced by complex interactions between chronic stress, hormonal dysregulation, and immune functioning<sup>6,12,14,148,169–173</sup>.

Future research may consider investigating psychological and psychotherapeutic interventions that explicitly target severe emotional conflicts

and unresolved (lifetime) trauma not directly related to the cancer diagnosis itself. Such interventions could focus on the long-term psychobiological impact of early-life adversity and chronic stress, aiming to modulate endocrine, immune, and inflammatory pathways implicated in cancer development and progression. Assessing the effects of these interventions on biomarkers and, ultimately, on recovery rates and survival, could yield valuable insights for psycho-oncology, as well as for the broader field of chronic disease prevention and treatment.

Moreover, it may be valuable to examine whether specific types of unresolved psycho-emotional conflicts are differentially associated with the development of distinct cancer types, potentially moderated by individual differences in coping capacity, resilience, stress reactivity, or responsiveness to trauma-focused therapy. Such moderating factors may have obscured relevant associations in previous (epidemiologic) research, where these psychological dimensions were often not accounted for. Clarifying these pathways could deepen our understanding of cancer etiology and lead to more personalized, psychologically informed intervention strategies. It may also help identify patient subgroups that are particularly likely to benefit from trauma-focused or emotion-centered psychotherapeutic approaches.

## Conclusion

The overarching outcome of our analysis suggests that the ongoing debate on whether psychosocial interventions are able to prolong the survival of cancer patients may have been limited by a serious statistical power problem due to: (1) true effect variation in the population; (2) underpowered primary studies (RCTs) with sample sizes too small and/or probability of events (death rates in trials) too low to reliably detect the (probably often small) effects observed within these studies; (3) the use of different effect size measures with varying degrees of precision (ORs vs. RRs vs. HRs), both in primary studies and meta-analyses; (4) diverse inclusion criteria for trials in prior meta-analyses reviewed resulting in varying numbers of included RCTs and, therefore, differing levels of statistical power; and (5) the existence of a rather small meta-analytic overall effect, which requires a sufficient number of trials for its reliable detection.

Altogether, our findings have shed light on the causes for heterogeneity in the results of primary studies and prior meta-analyses and provided robust evidence for the existence of a *generalizable* true overall survival-enhancing effect according to both (1) our contemporary meta-analysis and (2) our multiverse meta-analysis encompassing reasonable meta-analytic specifications and corresponding overall effects; based on the currently available corpus of primary studies. Taken together, the evidence supports a paradigm shift: psychosocial interventions should no longer be viewed as optional support but as a fundamental, survival-relevant component of comprehensive cancer care.

In order to gain a clearer understanding of which types of psychosocial interventions are most effective for whom, further high-quality studies, followed by meta-analytic syntheses, are necessary.

## Data availability

All numerical data underlying plots and analyses, as well as other study materials, are openly available as Excel and Word files at the Open Science Framework: <https://osf.io/wnxdb/files>.

## Code availability

All custom R code used for analyses in this study is openly available at the Open Science Framework: <https://osf.io/wnxdb/files>.

Received: 5 May 2024; Accepted: 28 January 2026;

Published online: 11 February 2026

## References

- Sung, H. et al. Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J. Clin.* **71**, 209–249 (2021).
- Fox, B. H. The role of psychological factors in cancer incidence and prognosis. *Oncology* **9**, 245–253 (1995).
- Garszen, B. Psychological factors and cancer development: evidence after 30 years of research. *Clin. Psychol. Rev.* **24**, 315–338 (2004).
- Giese-Davis, J. et al. Decrease in depression symptoms is associated with longer survival in patients with metastatic breast cancer: a secondary analysis. *J. Clin. Oncol.* **29**, 413–420 (2011).
- Eckerling, A., Ricon-Becker, I., Sorski, L., Sandbank, E. & Ben-Eliyahu, S. Stress and cancer: mechanisms, significance and future directions. *Nat. Rev. Cancer* **21**, 767–785 (2021).
- Roche, K., Cooper, D., Armstrong, T. S. & King, A. S. The link between psychological distress and survival in solid tumor patients: a systematic review. *Cancer Med.* **12**, 3343–3364 (2023).
- Wang, Y. H. et al. Depression and anxiety in relation to cancer incidence and mortality: a systematic review and meta-analysis of cohort studies. *Mol. Psychiatry* **25**, 1487–1499 (2020).
- Penedo, F. J., Benedict, C. & McGregor, B. Cancer: psychosocial treatment. In *Encyclopedia of Behavioral Medicine* (eds Gellman, M. D. & Turner, J. R.) 319–325 (Springer, 2013).
- Spiegel, D. Minding the body: psychotherapy and cancer survival. *Br. J. Health Psychol.* **19**, 465–485 (2013).
- Benedict, C., Walsh, E. A. & Penedo, F. J. Psychosocial interventions in cancer. In *Psychological Aspects of Cancer: A Guide to Emotional and Psychological Consequences of Cancer, their Causes, and their Management* (eds Carr, B. I. & Steel, J.) 159–169 (Springer, 2022).
- Mészáros Crow, E. et al. Psychosocial interventions reduce cortisol in breast cancer patients: Systematic review and meta-analysis. *Front. Psychol.* **14**, 1–15 (2023).
- Chida, Y., Hamer, M., Wardle, J. & Steptoe, A. Do stress-related psychosocial factors contribute to cancer incidence and survival? *Nat. Rev. Clin. Oncol.* **5**, 466–475 (2008).
- Moreno-Smith, M., Lutgendorf, S. K. & Sood, A. K. Impact of stress on cancer metastasis. *Future Oncol.* **6**, 1863–1881 (2010).
- Sephton, S. E. Diurnal cortisol rhythm as a predictor of breast cancer survival. *J. Natl. Cancer Inst.* **92**, 994–1000 (2000).
- Sephton, S. E. et al. Diurnal cortisol rhythm as a predictor of lung cancer survival. *Brain Behav. Immun.* **30**, 163–170 (2013).
- Nowels, M. A. et al. Palliative care interventions effects on psychological distress: a systematic review & meta-analysis. *J. Pain. Symptom Manag.* **65**, 691–713 (2023).
- Spiegel, D., Kraemer, H., Bloom, J. & Gotthel, E. Effect of psychosocial treatment on survival of patients with metastatic breast cancer. *Lancet* **334**, 888–891 (1989).
- Goodwin, P. J. et al. The effect of group psychosocial support on survival in metastatic breast cancer. *N. Engl. J. Med.* **345**, 1719–1726 (2001).
- Kissane, D. W. et al. Supportive-expressive group therapy for women with metastatic breast cancer: survival and psychosocial outcome from a randomized controlled trial. *Psychooncology* **16**, 277–286 (2007).
- Spiegel, D. et al. Effects of supportive-expressive group therapy on survival of patients with metastatic breast cancer: a randomized prospective trial. *Cancer* **110**, 1130–1138 (2007).
- Cunningham, A. J. & Edmonds, C. V. Possible effects of psychological therapy on survival duration in cancer patients. *J. Clin. Oncol.* **23**, 5263–5263 (2005).
- Kraemer, H. C., Kuchler, T. & Spiegel, D. Use and misuse of the consolidated standards of reporting trials (CONSORT) guidelines to assess research findings: comment on Coyne, Stefanek, & Palmer (2007). *Psychol. Bull.* **135**, 173–178 (2009).
- Spiegel, D. Can psychotherapy prolong cancer survival? *Psychosomatics* **31**, 361–366 (1990).
- Spiegel, D. Mind matters in cancer survival. *JAMA* **305**, 502 (2011).

25. Spiegel, D. & Giese-Davis, J. Examining the evidence that psychotherapy improves the survival of cancer patients. *Biol. Psychiatry* **56**, 62–64 (2004).
26. Coyne, J. C., Stefanek, M. & Palmer, S. C. Psychotherapy and survival in cancer: the conflict between hope and evidence. *Psychol. Bull.* **133**, 367–394 (2007).
27. Coyne, J. C., Thombs, B. D., Stefanek, M. & Palmer, S. C. Time to let go of the illusion that psychotherapy extends the survival of cancer patients: reply to Kraemer, Kuchler, and Spiegel (2009). *Psychol. Bull.* **135**, 179–182 (2009).
28. Fox, B. H. Some problems and some solutions in research on psychotherapeutic intervention in cancer. *Support. Care Cancer* **3**, 257–263 (1995).
29. Fox, B. H. A hypothesis about Spiegel et al.'s 1989 paper on psychosocial intervention and breast cancer survival. *Psychooncol.* **7**, 361–370 (1998).
30. Edelman, S., Craig, A. & Kidman, A. D. Can psychotherapy increase the survival time of cancer patients? *J. Psychosom. Res.* **49**, 149–156 (2000).
31. Goodwin, P. J. Psychosocial support for women with advanced breast cancer. *Breast Cancer Res. Treat.* **81**, 103–110 (2003).
32. Kissane, D. Beyond the psychotherapy and survival debate: The challenge of social disparity, depression and treatment adherence in psychosocial cancer care. *Psychooncology* **18**, 1–5 (2009).
33. Palmer, S. C. & Coyne, J. C. Examining the evidence that psychotherapy improves the survival of cancer patients. *Biol. Psychiatry* **56**, 61–62 (2004).
34. Palmer, S. C., Stefanek, M. E., Thombs, B. D. & Coyne, J. C. Psychologic intervention and survival: wishing does not make it so—Letter. *Clin. Cancer Res.* **16**, 5364–5365 (2010).
35. Stefanek, M. E. Psychotherapy and cancer survival: a cautionary note. *Psychosomatics* **32**, 237–238 (1991).
36. Stefanek, M. E., Palmer, S. C., Thombs, B. D. & Coyne, J. C. Finding what is not there. *Cancer* **115**, 5612–5616 (2009).
37. Rosner, A. L. Evidence-based medicine: revisiting the pyramid of priorities. *J. Bodyw. Mov. Ther.* **16**, 42–49 (2012).
38. Borenstein, M., Hedges, L. V., Higgins, J. P. T. & Rothstein, H. R. *Introduction to Meta-Analysis* (John Wiley & Sons, 2021).
39. Bothwell, L. E., Greene, J. A., Podolsky, S. H. & Jones, D. S. Assessing the gold standard—lessons from the history of RCTs. *N. Engl. J. Med.* **374**, 2175–2181 (2016).
40. Mirosevic, S. et al. Not just another meta-analysis”: sources of heterogeneity in psychosocial treatment effect on cancer survival. *Cancer Med.* **8**, 363–373 (2019).
41. Simonsohn, U., Simmons, J. & Nelson, L. D. Above averaging in literature reviews. *Nat. Rev. Psychol.* **1**, 551–552 (2022).
42. Ioannidis, J. P. The mass production of redundant, misleading, and conflicted systematic reviews and meta-analyses. *Milbank Q.* **94**, 485–514 (2016).
43. Chow, E., Tsao, M. N. & Harth, T. Does psychosocial intervention improve survival in cancer? A meta-analysis. *Palliat. Med.* **18**, 25–31 (2004).
44. Spruance, S. L., Reid, J. E., Grace, M. & Samore, M. Hazard ratio in clinical trials. *Antimicrob. Agents Chemother.* **48**, 2787–2792 (2004).
45. Banks, G. C., Rogelberg, S. G., Woznyj, H. M., Landis, R. S. & Rupp, D. E. Evidence on questionable research practices: the good, the bad, and the ugly. *J. Bus. Psychol.* **31**, 323–338 (2016).
46. Ioannidis, J. P. Why most published research findings are false. *PLoS Med.* **2**, e124 (2005).
47. Conn, V. S. & Rantz, M. J. Research methods: managing primary study quality in meta-analyses. *Res. Nurs. Health* **26**, 322–333 (2003).
48. Onwuegbuzie, A. J. & Leech, N. L. Post hoc power: a concept whose time has come. *Underst. Stat.* **3**, 201–230 (2004).
49. Turner, R. M., Bird, S. M. & Higgins, J. P. T. The impact of study size on meta-analyses: examination of underpowered studies in Cochrane reviews. *PLoS ONE* **8**, e59202 (2013).
50. Bekelman, J. E., Li, Y. & Gross, C. P. Scope and impact of financial conflicts of interest in biomedical research: a systematic review. *JAMA* **289**, 454–465 (2003).
51. Lundh, A., Lexchin, J., Mintzes, B., Schroll, J. B. & Bero, L. Industry sponsorship and research outcome: systematic review with meta-analysis. *Intensive Care Med.* **44**, 1603–1612 (2018).
52. Ioannidis, J. P. Meta-research: the art of getting it wrong. *Res. Synth. Methods* **1**, 169–184 (2010).
53. Ioannidis, J. P. Meta-research: why research on research matters. *PLoS Biol.* **16**, e2005468 (2018).
54. Higgins, J. P. T. et al. The Cochrane Collaboration's tool for assessing risk of bias in randomised trials. *BMJ* **343**, d5928 (2011).
55. Int'Hout, J., Ioannidis, J. P., Rovers, M. M. & Goeman, J. J. Plea for routinely presenting prediction intervals in meta-analysis. *BMJ Open* **6**, e010247 (2016).
56. Kizilirmak, P., Özdemir, O. & Öngen, Z. The most critical question when reading a meta-analysis report: Is it comparing apples with apples or apples with oranges? *Anatol. J. Cardiol.* **15**, 701–708 (2015).
57. Gospodarowicz, M. & O'Sullivan, B. Prognostic factors in cancer. *Semin. Surg. Oncol.* **21**, 13–18 (2003).
58. Zimmermann, T., Heinrichs, N. & Baucom, D. Does one size fit all? Moderators in psychosocial interventions for breast cancer patients: a meta-analysis. *Ann. Behav. Med.* **34**, 225–239 (2007).
59. Teo, I., Krishnan, A. & Lee, G. L. Psychosocial interventions for advanced cancer patients: a systematic review. *Psychooncology* **28**, 1394–1407 (2019).
60. Lipsey, M. W. Those confounded moderators in meta-analysis: good, bad, and ugly. *Ann. Am. Acad. Pol. Soc. Sci.* **587**, 69–81 (2003).
61. van Lissa, C. J. MetaForest. exploring heterogeneity in meta-analysis using random forest. Preprint at PsyArxiv <https://doi.org/10.31234/osf.io/myg6s> (2017).
62. Richardson, M., Garner, P. & Donegan, S. Interpretation of subgroup analyses in systematic reviews: a tutorial. *Clin. Epidemiol. Glob. Health* **7**, 192–198 (2019).
63. McClelland, G. H., Lynch Jr, J. G., Irwin, J. R., Spiller, S. A. & Fitzsimons, G. J. Median splits, type II errors, and false-positive consumer psychology: don't fight the power. *J. Consum. Psychol.* **25**, 679–689 (2015).
64. Marks-Anglin, A. & Chen, Y. A historical review of publication bias. *Res. Synth. Methods* **11**, 725–742 (2020).
65. van Aert, R. C., Wicherts, J. M. & van Assen, M. A. L. M. Publication bias examined in meta-analyses from psychology and medicine: a meta-meta-analysis. *PLoS ONE* **14**, e0215052 (2019).
66. Carter, E. C., Schönbrodt, F. D., Gervais, W. M. & Hilgard, J. Correcting for bias in psychology: a comparison of meta-analytic methods. *Adv. Methods Pract. Psychol. Sci.* **2**, 115–144 (2019).
67. Wicherts, J. M. et al. Degrees of freedom in planning, running, analyzing, and reporting psychological studies: a checklist to avoid p-hacking. *Front. Psychol.* **7**, 1832 (2016).
68. Oh, P. J., Shin, S. R., Ahn, H. S. & Kim, H. J. Meta-analysis of psychosocial interventions on survival time in patients with cancer. *Psychol. Health* **31**, 396–419 (2016).
69. Egger, M., Smith, G. D., Schneider, M. & Minder, C. Bias in meta-analysis detected by a simple, graphical test. *BMJ* **315**, 629–634 (1997).
70. Stanley, T. D. & Doucouliagos, H. Meta-regression approximations to reduce publication selection bias. *Res. Synth. Methods* **5**, 60–78 (2014).
71. Rosenthal, R. The file drawer problem and tolerance for null results. *Psychol. Bull.* **86**, 638–641 (1979).

72. Duval, S. & Tweedie, R. Trim and fill: a simple funnel-plot-based method of testing and adjusting for publication bias in meta-analysis. *Biometrics* **56**, 455–463 (2000).
73. Pérez-Bracchiglione, J. et al. Graphical representation of overlap for OVEviews: GROOVE tool. *Res. Synth. Methods* **13**, 381–388 (2022).
74. Simonsohn, U., Simmons, J. P. & Nelson, L. D. Specification curve: descriptive and inferential statistics on all reasonable specifications. SSRN <https://doi.org/10.2139/ssrn.2694998> (2015).
75. Steegen, S., Tuerlinckx, F., Gelman, A. & Vanpaemel, W. Increasing transparency through a multiverse analysis. *Perspect. Psychol. Sci.* **11**, 702–712 (2016).
76. Voracek, M., Kossmeier, M. & Tran, U. S. Which data to meta-analyze, and how? A specification-curve and multiverse-analysis approach to meta-analysis. *Z. f.ür. Psychologie* **227**, 64–82 (2019).
77. Pietschnig, J., Gerdemann, D., Zeiler, M. & Voracek, M. Of differing methods, disputed estimates and discordant interpretations: the meta-analytical multiverse of brain volume and IQ associations. *R. Soc. Open Sci.* **9**, 211621 (2022).
78. McGrath, S., Zhao, X., Qin, Z. Z., Steele, R. & Benedetti, A. One-sample aggregate data meta-analysis of medians. *Stat. Med.* **38**, 969–984 (2018).
79. McGrath, S., Sohn, H., Steele, R. & Benedetti, A. Meta-analysis of the difference of medians. *Biom. J.* **62**, 69–98 (2019).
80. Hartmann, K. E., Matchar, D. B. & Chang, S. M. In *Methods Guide for Medical Test Reviews* (eds Chang, S. M., Matchar, D. B. & Smetana, G. W.) (National Library of Science, 2012).
81. Page, M. J. et al. Updating guidance for reporting systematic reviews: development of the PRISMA 2020 statement. *J. Clin. Epidemiol.* **134**, 103–112 (2021).
82. LaMorte, W. W. Estimating the survival function. [https://sphweb.bumc.bu.edu/otlt/mph-modules/bs/bs704\\_survival/bs704\\_survival4.html](https://sphweb.bumc.bu.edu/otlt/mph-modules/bs/bs704_survival/bs704_survival4.html) (2016).
83. Tierney, J. F., Stewart, L. A., Ghersi, D., Burdett, S. & Sydes, M. R. Practical methods for incorporating summary time-to-event data into meta-analysis. *Trials* **8**, 16 (2007).
84. Kossmeier, M., Tran, U. S. & Voracek, M. Visualizing meta-analytic data with R package metaviz. <https://cran.uni-muenster.de/web/packages/metaviz/vignettes/metaviz.html> (2020).
85. Dusseldorp, E., van Genugten, L., van Buuren, S., Verheijden, M. W. & van Empelen, P. Combinations of techniques that effectively change health behavior: evidence from meta-CART analysis. *Health Psychol.* **33**, 1530–1540 (2014).
86. Bramer, W. M., Rethlefsen, M. L., Kleijnen, J. & Franco, O. H. Optimal database combinations for literature searches in systematic reviews: a prospective exploratory study. *Syst. Rev.* **6**, 245 (2017).
87. Guo, Z. et al. The benefits of psychosocial interventions for cancer patients undergoing radiotherapy. *Health Qual. Life Outcomes* **11**, 121 (2013).
88. Julião, M., Nunes, B. & Barbosa, A. Dignity therapy and its effect on the survival of terminally ill Portuguese patients. *Psychother. Psychosom.* **84**, 57–58 (2015).
89. Kirkegaard, A. M. et al. Effects on long-term survival of psychosocial group intervention in early-stage breast cancer: Follow-up of a randomized controlled trial. *Acta Oncol.* **62**, 422–428 (2023).
90. Kissane, D. W. et al. Effect of cognitive-existential group therapy on survival in early-stage breast cancer. *J. Clin. Oncol.* **22**, 4255–4260 (2004).
91. Ross, L. et al. No effect on survival of home psychosocial intervention in a randomized study of Danish colorectal cancer patients. *Psychooncology* **18**, 875–885 (2009).
92. Temel, J. S. et al. Early palliative care for patients with metastatic non-small-cell lung cancer. *N. Engl. J. Med.* **363**, 733–742 (2010).
93. Higgins, J. P. T. et al. *Cochrane Handbook for Systematic Reviews of Interventions (Version 6.5)* (Cochrane, 2024).
94. GetData graph digitizer v. 2.26 <https://getdata-graph-digitizer.software.informer.com/> (2013).
95. Georgiev, G. Z. Hazard ratio calculator. <https://www.gigacalculator.com/calculators/hazard-ratio-calculator.php> (2026).
96. Azuero, A. A note on the magnitude of hazard ratios. *Cancer* **122**, 1298–1299 (2016).
97. Viechtbauer, W. Conducting meta-analyses in R with the metafor package. *J. Stat. Softw.* **36**, 1–48 (2010).
98. Langan, D. et al. A comparison of heterogeneity variance estimators in simulated random-effects meta-analyses. *Res. Synth. Methods* **10**, 83–98 (2018).
99. Viechtbauer, W. Bias and efficiency of meta-analytic variance estimators in the random-effects model. *J. Educ. Behav. Stat.* **30**, 261–293 (2005).
100. Deeks, J. J., Higgins, J. P. T., Altman, D. G. & Cochrane Statistical Methods, G. in *Cochrane Handbook for Systematic Reviews of Interventions* (eds Higgins, J. P. T. et al.) Ch. 10 (Wiley-Blackwell, 2019).
101. Spinelli, L. M. & Pandis, N. Prediction interval in random-effects meta-analysis. *Am. J. Orthod. Dentofacial Orthop.* **157**, 586–588 (2020).
102. Chalmers, T. C. & Lau, J. Meta-analytic stimulus for changes in clinical trials. *Stat. Methods Med. Res.* **2**, 161–172 (1993).
103. Sterne, J. A. C. et al. RoB 2: a revised tool for assessing risk of bias in randomised trials. *BMJ* **366**, 14898 (2019).
104. Landis, J. R. & Koch, G. G. The measurement of observer agreement for categorical data. *Biometrics* **33**, 159–174 (1977).
105. National Heart, L. & Blood Institute. Study quality assessment tools: quality assessment of controlled intervention studies. <https://www.nhlbi.nih.gov/health-topics/study-quality-assessment-tools> (2020).
106. Verhagen, A. P. et al. The Delphi list: a criteria list for quality assessment of randomized clinical trials for conducting systematic reviews developed by Delphi consensus. *J. Clin. Epidemiol.* **51**, 1235–1241 (1998).
107. Schulz, K. F., Altman, D. G., Moher, D. & Consort, G. CONSORT 2010 statement: updated guidelines for reporting parallel group randomised trials. *Trials* **11**, 32 (2010).
108. Hoenig, J. M. & Heisey, D. M. The abuse of power: the pervasive fallacy of power calculations for data analysis. *Am. Stat.* **55**, 19–24 (2001).
109. Zhang, Y. et al. Post hoc power analysis: Is it an informative and meaningful analysis? *Gen. Psychiatry* **32**, e100069 (2019).
110. Consulting, H. Test time-to-event data Cox PH, 2-sided equality: power and sample size calculators. <http://powerandsamplesize.com/Calculators/Test-Time-To-Event-Data/Cox-PH-2-Sided-Equality> (2013).
111. Lexchin, J. Sponsorship bias in clinical research. *Int. J. Risk Saf. Med.* **24**, 233–242 (2012).
112. Thissen, D., Steinberg, L. & Kuang, D. Quick and easy implementation of the Benjamini–Hochberg procedure for controlling the false positive rate in multiple comparisons. *J. Educ. Behav. Stat.* **27**, 77–83 (2002).
113. van De Schoot, R. & Miočević, M. (eds.) *Small Sample Size Solutions: A Guide for Applied Researchers and Practitioners* (CRC Press, 2020).
114. Breiman, L. Random forests. *Mach. Learn.* **45**, 5–32 (2001).
115. Strobl, C., Malley, J. & Tutz, G. An introduction to recursive partitioning: rationale, application, and characteristics of classification and regression trees, bagging, and random forests. *Psychol. Methods* **14**, 323–348 (2009).
116. Kuhn, M. Building predictive models in R using the caret package. *J. Stat. Softw.* **28**, 1–26 (2008).
117. Palmer, T. M., Sutton, A. J., Peters, J. L. & Moreno, S. G. Contour-enhanced funnel plots for meta-analysis. *Stata J. Promoting Commun. Stat. Stata* **8**, 242–254 (2008).

118. Peters, J. L., Sutton, A. J., Jones, D. R., Abrams, K. R. & Rushton, L. Contour-enhanced meta-analysis funnel plots help distinguish publication bias from other causes of asymmetry. *J. Clin. Epidemiol.* **61**, 991–996 (2008).
119. van Assen, M. A. L. M., van Aert, R. & Wicherts, J. M. Meta-analysis using effect size distributions of only statistically significant studies. *Psychol. Methods* **20**, 293–309 (2015).
120. van Aert, R. C. & van Assen, M. A. L. M. Correcting for publication bias in a meta-analysis with the P-uniform\* method. Preprint at *MetaArxiv* <https://osf.io/ebq6m/> (2018).
121. Alinaghi, N. & Reed, W. R. Meta-analysis and publication bias: How well does the FAT-PET-PEESE procedure work? *Res. Synth. Methods* **9**, 285–311 (2018).
122. Simonsohn, U., Nelson, L. D. & Simmons, J. P. P-curve: a key to the file-drawer. *J. Exp. Psychol.* **143**, 534–547 (2014).
123. Simonsohn, U., Simmons, J. P. & Nelson, L. D. Better p-curves: making p-curve analysis more robust to errors, fraud, and ambitious p-hacking, a reply to Ulrich and Miller (2015). *J. Exp. Psychol.* **144**, 1146–1152 (2015).
124. Schünemann, H. J. et al. The development methods of official GRADE articles and requirements for claiming the use of GRADE—a statement by the GRADE guidance group. *J. Clin. Epidemiol.* **159**, 79–84 (2023).
125. Linn, M. W., Linn, B. S. & Harris, R. Effects of counseling for late stage cancer patients. *Cancer* **49**, 1048–1055 (1982).
126. McCorkle, R. et al. Advancing geriatric nursing practice: a specialized home care intervention improves survival among older post-surgical cancer patients. *J. Am. Geriatr. Soc.* **48**, 1707–1713 (2000).
127. Choi, J. et al. Cognitive behavioral symptom management intervention in patients with cancer: survival analysis. *Support. Care Cancer* **20**, 1243–1250 (2012).
128. Gregorutti, B., Michel, B. & Saint-Pierre, P. Correlation and variable importance in random forests. *Stat. Comput.* **27**, 659–678 (2017).
129. Morey, R. D. & Davis-Stober, C. P. On the poor statistical properties of the P-curve meta-analytic procedure. *J. Am. Stat. Assoc.* **00**, 1–13 (2025).
130. Simonsohn, U., Nelson, L. D. & Simmons, J. P. P. -curve and effect size: correcting for publication bias using only significant results. *Perspect. Psychol. Sci.* **9**, 666–681 (2014).
131. Aizer, A. A. et al. Marital status and survival in patients with cancer. *J. Clin. Oncol.* **31**, 3869–3876 (2013).
132. Evans, D. et al. Meta-analysis of adjuvant chemotherapy versus no adjuvant chemotherapy for resected stage I pancreatic cancer. *Surgery* **175**, 1470–1479 (2024).
133. Ren, W. et al. Clinical evidence of chemotherapy or endocrine therapy maintenance in patients with metastatic breast cancer: meta-analysis of randomized clinical trials and propensity score matching of multicenter cohort study. *Cancer Res. Treat.* **54**, 1038–1052 (2022).
134. Garg, P. K., Jakhetiya, A., Pandey, R., Chishi, N. & Pandey, D. Adjuvant radiotherapy versus observation following lumpectomy in ductal carcinoma in situ: a meta-analysis of randomized controlled trials. *Breast J.* **24**, 233–239 (2017).
135. Ren, B. et al. A meta-analysis of the efficacy of postoperative adjuvant radiotherapy versus no radiotherapy for extrahepatic cholangiocarcinoma and gallbladder carcinoma. *Radiat. Oncol.* **15**, 15 (2020).
136. Yu, X. et al. Hormone replacement therapy and breast cancer survival: a systematic review and meta-analysis of observational studies. *Breast Cancer* **24**, 643–657 (2017).
137. De Ruysscher, D. et al. Radiotherapy toxicity. *Nat. Rev. Dis. Primers* **5**, 13 (2019).
138. Fleming, L. et al. The impact of medication side effects on adherence and persistence to hormone therapy in breast cancer survivors: a quantitative systematic review. *Breast* **64**, 63–84 (2022).
139. Nurgali, K., Jagoe, R. T. & Abalo, R. Editorial: Adverse effects of cancer chemotherapy: anything new to improve tolerance and reduce sequelae? *Front. Pharmacol.* **9**, 245 (2018).
140. McLouth, L. E. et al. A systematic review and meta-analysis of effects of psychosocial interventions on spiritual well-being in adults with cancer. *Psychooncology* **30**, 147–158 (2020).
141. de la Torre-Luque, A., Gambará, H., López, E. & Cruzado, J. A. Psychological treatments to improve quality of life in cancer contexts: a meta-analysis. *Int. J. Clin. Health Psychol.* **16**, 211–219 (2016).
142. Sanjida, S. et al. Are psychological interventions effective on anxiety in cancer patients? A systematic review and meta-analyses. *Psychooncology* **27**, 2063–2076 (2018).
143. Hart, S. L. et al. Meta-analysis of efficacy of interventions for elevated depressive symptoms in adults diagnosed with cancer. *J. Natl. Cancer Inst.* **104**, 990–1004 (2012).
144. Zhao, C. et al. The effects of acceptance and commitment therapy on the psychological and physical outcomes among cancer patients: a meta-analysis with trial sequential analysis. *J. Psychosom. Res.* **140**, 110304 (2021).
145. Xunlin, N., Lau, Y. & Klainin-Yobas, P. The effectiveness of mindfulness-based interventions among cancer patients and survivors: a systematic review and meta-analysis. *Support. Care Cancer* **28**, 1563–1578 (2019).
146. Warth, M. et al. Psychosocial interventions for pain management in advanced cancer patients: a systematic review and meta-analysis. *Curr. Oncol. Rep.* **22**, 3 (2020).
147. Andersen, B. L., Shelby, R. A. & Golden-Kreutz, D. M. RCT of a psychological intervention for patients with cancer: I. Mechanisms of change. *J. Consult. Clin. Psychol.* **75**, 927 (2007).
148. Antoni, M. H. et al. The influence of bio-behavioural factors on tumour biology: pathways and mechanisms. *Nat. Rev. Cancer* **6**, 240–248 (2006).
149. Antoni, M. H. et al. Reduction of cancer-specific thought intrusions and anxiety symptoms with a stress management intervention among women undergoing treatment for breast cancer. *Am. J. Psychiatry* **163**, 1791–1797 (2006).
150. Giordano, S. H. et al. Estimating regimen-specific costs of chemotherapy for breast cancer: observational cohort study. *Cancer* **122**, 3447–3455 (2016).
151. Boero, I. J. et al. The impact of radiotherapy costs on clinical outcomes in breast cancer. *Radiother. Oncol.* **117**, 393–399 (2015).
152. Lengacher, C. A. et al. A cost-effective mindfulness stress reduction program: a randomized control trial for breast cancer survivors. *Nurs. Econ.* **33**, 210–232 (2015).
153. Jansen, F., Zwieter, V. V., Coupé, V. M., Leemans, C. R. & Verdonck-de Leeuw, I. M. A review on cost-effectiveness and cost-utility of psychosocial care in cancer patients. *Asia Pac. J. Oncol. Nurs.* **3**, 125–136 (2016).
154. Mewes, J. C. et al. Cost-effectiveness of cognitive behavioral therapy and physical exercise for alleviating treatment-induced menopausal symptoms in breast cancer patients. *J. Cancer Surviv.* **9**, 126–135 (2014).
155. Lemieux, J., Topp, A., Chappell, H., Ennis, M. & Goodwin, P. J. Economic analysis of psychosocial group therapy in women with metastatic breast cancer. *Breast Cancer Res. Treat.* **100**, 183–190 (2006).
156. Christley, R. M. Power and error: increased risk of false positive results in underpowered studies. *Open Epidemiol. J.* **3**, 16–19 (2010).

157. Maxwell, S. E. The persistence of underpowered studies in psychological research: causes, consequences, and remedies. *Psychol. Methods* **9**, 147–163 (2004).
158. Fu, W. W. et al. The impact of psychosocial intervention on survival in cancer: a meta-analysis. *Ann. Palliat. Med.* **5**, 93–106 (2016).
159. Smedslund, G. & Ringdal, G. I. Meta-analysis of the effects of psychosocial interventions on survival time in cancer patients. *J. Psychosom. Res.* **57**, 123–131 (2004).
160. Ioannidis, J. P. A. & Trikalinos, T. A. The appropriateness of asymmetry tests for publication bias in meta-analyses: a large survey. *Can. Med. Assoc. J.* **176**, 1091–1096 (2007).
161. Esterhuizen, T. M. & Thabane, L. Con: meta-analysis: some key limitations and potential solutions. *Nephrol. Dial. Transplant.* **31**, 882–885 (2016).
162. Altman, D. G. & Bland, J. M. Statistics notes: absence of evidence is not evidence of absence. *BMJ* **311**, 485 (1995).
163. Cunningham, A. et al. A randomized controlled trial of the effects of group psychological therapy on survival in women with metastatic breast cancer. *Psychooncology* **7**, 508–517 (1998).
164. Edelman, S., Bell, D. R. & Kidman, A. D. A group cognitive behaviour therapy programme with metastatic breast cancer patients. *Psychooncology* **8**, 295–305 (1999).
165. Ye, Z. J. et al. Effect of a mentor-based, supportive-expressive program, be resilient to breast cancer, on survival in metastatic breast cancer: a randomised, controlled intervention trial. *Br. J. Cancer* **117**, 1486–1494 (2017).
166. Crutzen, R. & Peters, G. J. Y. Targeting next generations to change the common practice of underpowered research. *Front. Psychol.* **8**, 1184 (2017).
167. Munafò, M. R. et al. A manifesto for reproducible science. *Nat. Hum. Behav.* **1**, 1–9 (2017).
168. Kogon, M. M., Biswas, A., Pearl, D., Carlson, R. W. & Spiegel, D. Effects of medical and psychotherapeutic treatment on the survival of women with metastatic breast carcinoma. *Cancer* **80**, 225–230 (1997).
169. Antoni, M. H. & Dhabhar, F. S. The impact of psychosocial stress and stress management on immune responses in patients with cancer. *Cancer* **125**, 1417–1431 (2019).
170. Hassan, S. et al. Behavioral stress accelerates prostate cancer development in mice. *J. Clin. Investig.* **123**, 874–886 (2013).
171. Obradović, M. M. S. et al. Glucocorticoids promote breast cancer metastasis. *Nature* **567**, 540–544 (2019).
172. Reiche, E. M. V., Nunes, S. O. V. & Morimoto, H. K. Stress, depression, the immune system, and cancer. *Lancet Oncol.* **5**, 617–625 (2004).
173. Shields, G. S., Spahr, C. M. & Slavich, G. M. Psychosocial interventions and immune system function: a systematic review and meta-analysis of randomized clinical trials. *JAMA Psychiatry* **77**, 1031–1043 (2020).
174. Zabalegui, A., Sanchez, S., Sanchez, P. D. & Juando, C. Nursing and cancer support groups. *J. Adv. Nurs.* **51**, 369–381 (2005).
175. Edwards, A. G., Hulbert-Williams, N. & Neal, R. Psychological interventions for women with metastatic breast cancer. *Cochrane Database Syst. Rev.* <https://doi.org/10.1002/14651858.cd004253.pub3> (2008).
176. Mustafa, M., Carson-Stevens, A., Gillespie, D. & Edwards, A. G. Psychological interventions for women with metastatic breast cancer. *Cochrane Database Syst. Rev.* <https://doi.org/10.1002/14651858.CD004253.pub4> (2013).
177. Xia, Y. et al. Psychosocial and behavioral interventions and cancer patient survival again: hints of an adjusted meta-analysis. *Integr. Cancer Ther.* **13**, 301–309 (2014).
178. Jassim, G. A., Whitford, D. L., Hickey, A. & Carter, B. Psychological interventions for women with non-metastatic breast cancer. *Cochrane Database Syst. Rev.* <https://doi.org/10.1002/14651858.cd008729.pub2> (2015).
179. Ilnyckyj, A., Farber, J., Cheang, M. & Weinerman, B. H. A randomized controlled trial of psychotherapeutic intervention in cancer patients. *Ann. R. Coll. Phys. Surg. Can.* **27**, 93–96 (1994).
180. Edelman, S., Lemon, J., Bell, D. R. & Kidman, A. D. Effects of group CBT on the survival time of patients with metastatic breast cancer. *Psychooncology* **8**, 474–481 (1999).
181. Fawzy, F. I. et al. Malignant melanoma: effects of an early structured psychiatric intervention, coping, and affective state on recurrence and survival 6 years later. *Arch. Gen. Psychiatry* **50**, 681–689 (1993).
182. Fawzy, F. I., Canada, A. I. & Fawzy, N. W. Malignant melanoma: effects of a brief, structured psychiatric intervention on survival and recurrence at 10-year follow-up. *Arch. Gen. Psychiatry* **60**, 100–103 (2003).
183. Boesen, E. H. et al. Survival after a psychoeducational intervention for patients with cutaneous malignant melanoma: a replication study. *J. Clin. Oncol.* **25**, 5698–5703 (2007).
184. Kuchler, T. H. et al. Impact of psychotherapeutic support on gastrointestinal cancer patients undergoing surgery: survival results of a trial. *Hepatogastroenterology* **46**, 322–335 (1999).
185. Kuchler, T., Bestmann, B., Rappat, S., Henne-Bruns, D. & Wood-Dauphinee, S. Impact of psychotherapeutic support for patients with gastrointestinal cancer undergoing surgery: 10-year survival results of a randomized trial. *J. Clin. Oncol.* **25**, 2702–2708 (2007).
186. Andersen, B. L. et al. Psychologic intervention improves survival for breast cancer patients: a randomized clinical trial. *Cancer* **113**, 3450–3458 (2008).
187. Bakitas, M. et al. Effects of a palliative care intervention on clinical outcomes in patients with advanced cancer: the project ENABLE II randomized controlled trial. *JAMA* **302**, 741–749 (2009).
188. Zhang, X. D. et al. Perioperative comprehensive supportive care interventions for Chinese patients with esophageal carcinoma: a prospective study. *Asian Pac. J. Cancer Prev.* **14**, 7359–7366 (2013).
189. Stagl, J. M. et al. A randomized controlled trial of cognitive-behavioral stress management in breast cancer: survival and recurrence at 11-year follow-up. *Breast Cancer Res. Treat.* **154**, 319–328 (2015).
190. Bao, H., Chen, Y., Li, M., Pan, L. & Zheng, X. Intensive patient’s care program reduces anxiety and depression as well as improves overall survival in de novo acute myelocytic leukemia patients who underwent chemotherapy: a randomized, controlled study. *Transl. Cancer Res.* **8**, 212–227 (2019).
191. Wang, J., Yan, C. & Fu, A. A randomized clinical trial of comprehensive education and care program compared to basic care for reducing anxiety and depression and improving quality of life and survival in patients with hepatocellular carcinoma who underwent surgery. *Medicine* **98**, e17552 (2019).
192. Zhang, L. L. et al. Reminiscence therapy exhibits alleviation of anxiety and improvement of life quality in postoperative gastric cancer patients: a randomized, controlled study. *Medicine* **100**, e26821 (2021).
193. Zhao, X. Reminiscence therapy-based care program for reducing anxiety and depression in glioma survivors. *Medicine* **100**, e23056 (2021).
194. Zhou, L. & Sun, H. The effect of reminiscence therapy-involved care on anxiety, depression, life quality and survival in colorectal cancer patients. *Clin. Res. Hepatol. Gastroenterol.* **45**, 101546 (2021).
195. Chen, L., Yang, X., Ren, X. & Lin, Y. Reminiscence therapy care program as a potential nursing intervention to relieve anxiety, depression, and quality of life in older papillary thyroid carcinoma patients: a randomized, controlled study. *Front. Psychol.* **13**, 1–10 (2022).

196. Guo, Q., Li, T., Cao, T. & Ma, C. Reminiscence therapy involved care programs as an option to improve psychological disorders and patient satisfaction in elderly lung cancer patients: a randomized, controlled study. *J. Cancer Res. Ther.* **18**, 1937–1944 (2022).
197. Huang, T., Su, H., Zhang, S. & Huang, Y. Reminiscence therapy-based care program serves as an optional nursing modality in alleviating anxiety and depression, improving quality of life in surgical prostate cancer patients. *Int. Urol. Nephrol.* **54**, 2467–2476 (2022).

## Acknowledgements

This research project received no external funding. We acknowledge institutional access to academic databases and computational resources that supported this work. Open-access publication fees were covered by the University of Vienna. The University had no role in study design, data collection, statistical analysis, decision to publish, or preparation of the manuscript. We thank Johannes K. Vilsmeier for reviewing and independently verifying the R code of the meta-analysis and multiverse meta-analysis. Artificial intelligence (ChatGPT and OpenAI) was used solely for language refinement (writing style, grammar, spelling, and clarity) and did not contribute to the development of scientific content, results, analyses, or interpretations. This article is dedicated to the memory of my mother, Hitomi Asakawa, once a world-famous principal dancer (prima ballerina) in Maurice Béjart's Ballet of the 20th Century, whose remarkable life and journey with cancer sparked the idea for this 7-year research project. Her grace, strength, and resilience continue to guide and inspire me to this day. I also wish to thank my father, Joachim Haas, whose unwavering belief in me and steadfast support have helped me navigate the ongoing challenges on the path of scientific research.

## Author contributions

K.D.A.-H.: Conceptualization, methodology, software, validation, formal analysis, investigation, data curation, writing—original draft, writing—review and editing, visualization, and project administration; D.S.: Conceptualization, validation, and writing—review and editing; L.B., A.G., and K.S.: Validation, investigation, and writing—review and editing; M.V.: Methodology and writing—review and editing; U.S.T.: Conceptualization, methodology, validation, writing—original draft, writing—review and editing, and supervision.

## Competing interests

The authors declare no competing interests.

## Additional information

**Supplementary information** The online version contains supplementary material available at <https://doi.org/10.1038/s44271-026-00414-x>.

**Correspondence** and requests for materials should be addressed to Kenji D. Asakawa-Haas or Ulrich S. Tran.

**Peer review information** *Communications Psychology* thanks Wim van Den Noortgate, Geir Smedslund, Wolfgang Linden and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. Primary Handling Editor: Jennifer Bellingtier. A peer review file is available.


**Reprints and permissions information** is available at <http://www.nature.com/reprints>

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

**Open Access** This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

© The Author(s) 2026

**Kenji D. Asakawa-Haas**  , **David Spiegel** <sup>2</sup>, **Lukas Bossert** <sup>1</sup>, **Aleksandra Garic** <sup>1</sup>, **Katrin Schwartz** <sup>1</sup>, **Martin Voracek** <sup>1</sup> & **Ulrich S. Tran** <sup>1</sup> 

<sup>1</sup>Department of Cognition, Emotion, and Methods in Psychology, Faculty of Psychology, University of Vienna, Vienna, Austria. <sup>2</sup>Department of Psychiatry & Behavioral Sciences, Stanford University School of Medicine, Stanford, California, CA, USA.  e-mail: [kenji.asakawa.haas@gmail.com](mailto:kenji.asakawa.haas@gmail.com); [ulrich.tran@univie.ac.at](mailto:ulrich.tran@univie.ac.at)