Cognitive behavioral therapies for informal caregivers of patients with cancer and cancer survivors: a systematic review and meta-analysis

Mia S. O’Toole¹ *, Robert Zachariae², Megan E. Renna³, Douglas S. Mennin⁴ and Allison Applebaum⁵

¹Unit for Psychooncology and Health Psychology, Institute of Psychology and Behavioral Science, School of Business and Social Sciences, Aarhus University, Aarhus, Denmark
²Unit for Psychooncology and Health Psychology, Aarhus University and Aarhus University Hospital, Aarhus, Denmark
³Department of Psychology, CUNY Hunter College and the Graduate Center, New York, USA
⁴Department of Psychology, CUNY Hunter College and the Graduate Center, New York, USA
⁵Department of Psychiatry and Behavioral Sciences, Memorial Sloan Kettering Cancer Center, New York, USA

*Correspondence to:
Unit for Psychooncology and Health Psychology, Institute of Psychology and Behavioral Science, School of Business and Social Sciences, Bartholins Allé 9, 8000 Aarhus C, Denmark
E-mail: mio@psy.au.dk

Abstract

Objective: Informal caregivers (ICs) of patients with cancer and cancer survivors report a number of psychological and physical complaints because of the burden associated with providing care. Given the documented effect of Cognitive Behavioral Therapy (CBT) on ICs’ common psychological complaints, such as anxiety and depression, the objective was to conduct a meta-analysis on the effect of CBTs for adult ICs.

Methods: A literature search was conducted in order to identify all intervention studies on adult ICs that employed at least one therapeutic component defined as a CBT component.

Results: Literature searches revealed 36 unique records with sufficient data. These studies were subjected to meta-analyses using random effects models. A small, statistically significant effect of CBTs (Hedge’s $g = 0.08$, $p = 0.014$) was revealed, which disappeared when randomized controlled trials were evaluated alone ($g = 0.04$, $p = 0.200$). A number of variables were explored as moderators. Only the percentage of female participants was positively associated with the effect size.

Conclusions: Based on the negligible effect of CBTs across outcomes, future studies should consider moving beyond traditional CBT methods as these do not appear efficacious. It is suggested that future interventions orient towards advances in the basic affective sciences and derived therapies in order to better understand and treat the emotional struggles experienced by ICs.

© 2016 The Authors. Psycho-Oncology published by John Wiley & Sons Ltd.

Background

There is growing recognition that informal caregivers (ICs) of chronically ill patients are themselves in need of care. Historically, research on caregiver burden has focused on ICs of patients with a variety of dementia, such as Alzheimer’s and Parkinson’s disease. More recently, the burden experienced by ICs of patients with cancer is receiving increased attention, which may in part be because of the rising incidence of cancer globally [1]. Such caregivers face the concurrent stress of significant role transitions and the responsibilities of managing patient needs, in addition to existing responsibilities, which commonly results in caregiver burden. Given et al. [2] describe caregiver burden as a ‘multidimensional biopsychosocial reaction resulting from an imbalance of care demands relative to caregivers’ personal time, social roles, physical and emotional states, financial resources, and formal care resources given the other multiple roles they fulfill’. Carrying this burden often comes with psychological and physical complaints [3,4]. ICs have been found to have high levels of psychological distress, and longitudinal studies have shown that caregiver burden is significantly associated with anxiety and depression over time [e.g. 5,6]. Examples of specific complaints by ICs include feeling overwhelmed by taking on the patient’s responsibilities, fear of losing the patient, and uncertainty about the future [7]. Such complaints are likely to persist into survivorship, as 30–40% of caregivers continue to experience clinical levels of anxiety and depression if their loved ones survives cancer [8,9]. Caregiver burden is also associated with a range of physical health complications, including sleep difficulties and fatigue [10,11], cardiovascular disease [12,13], poor immune functioning [14,15], and increased mortality [16,17]. Together, the psychological and physical symptoms associated with providing care to a patient with cancer place caregivers at particular risk for experiencing negative outcomes and hence are in urgent need of effective interventions. Despite this fact, the
state of science of intervention development for ICs of patients with cancer is in its infancy [18,19].

Cognitive behavioral therapies (CBTs) have been found to be effective in treating individuals presenting with symptoms of anxiety and depression [20,21] – common complaints of ICs – and hence it is likely an appropriate first choice treatment for ICs. A number of systematic reviews, including one meta-analysis, of psychological interventions for ICs have been conducted. Most reviews have evaluated intervention feasibility and quality of design, and only a few have attempted to evaluate intervention efficacy and effectiveness [22–25]. Two reviews have specifically evaluated the effect of CBTs on a number of different outcomes in ICs such as quality of life and burden [19,26]. In one meta-analysis [27], 29 RCTs for ICs were evaluated. Although studies were categorized according to intervention framework, that is, their self-stated main treatment orientation, the specific effect of the 7 studies categorized as employing CBTs was not evaluated. In a later narrative, systematic review of 49 existing psychological intervention studies for ICs of patients with cancer, the three interventions categorized as CBTs all had a positive effect [19]. Across these reviews, the manner in which interventions were categorized should be highlighted. Northouse et al. [27] categorized interventions according to the ‘primary framework’ as stated by the authors. However, a ‘stress and coping’ framework may not look much different than a ‘cognitive-behavioral’ framework in terms of the therapeutic methods used. Likewise, categorizing interventions according to their ‘primary focus’ [19] may instill arbitrary differences between interventions that use similar techniques. In contrast, a narrative systematic review by O’Toole et al. [26] employed a definition of CBT according to the intervention strategies actually employed (i.e. cognitive restructuring, imaginal or in vivo exposure, coping skills training, problem-solving, behavior activation, behavioral experiments, structured homework, acceptance-based strategies, stress and anxiety management through relaxation, or mindfulness [cf. 28]). Thirty-nine studies belonging to this umbrella of CBTs were evaluated, and results showed that about half of the studies produced at least one positive outcome, whereas 33% did not detect any effect of the intervention, and 15% did not report any inferential statistics because of a small sample size, or did not report the relevant statistical analyses. The review concluded that meta-analytic efforts would be an important next step in evaluating the effect of CBTs for ICs.

The primary aim of the present study was therefore to evaluate the overall effect of interventions using CBT components, which has not previously been meta-analytically evaluated. The goal was to evaluate the effect of CBTs, tested in both randomized and open designs, on a number of outcomes, including mastery and well-being (i.e. psychological, physical and social well-being) in ICs of patients with cancer. We hypothesized that CBTs would be effective across outcomes. A secondary aim was to explore possible moderators of this effect, including trial design, outcomes evaluated, demographic variables, intervention duration and modality of delivery, illness-related variables among patients, and study quality.

Methods

The study was conducted in accordance with the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) recommendations [29].

Selection criteria

Included studies were peer-reviewed and (a) investigated the effect of a CBT for ICs of patients with cancer or cancer survivors, (b) employed at least one quantitative measure of psychological, physical, or interpersonal functioning/well-being of the IC both pre- and post-intervention, (c) enrolled adult samples (age ≥18 years), (d) reported results that could be converted into an effect size, and (e) were written in English. An intervention was considered a CBT if it included at least one of the following components: cognitive restructuring, imaginal or in vivo exposure, coping skills training, problem-solving, behavior activation, behavioral experiments, structured homework, acceptance-based strategies, stress and anxiety management through relaxation, or mindfulness [cf. 28]. All papers were evaluated independently by authors MSO and MR, and disagreement on the inclusion/exclusion of a study was resolved by consensus.

Search strategy

A keyword-based search in the electronic databases of PsychINFO, Cochrane, CINAHL, and Embase was conducted. Keywords related to oncology (cancer OR neoplasm OR oncology OR palliative care OR palliative medicine OR malignancy) were combined with keywords related to the population (caregiver* OR carer* OR caregiver OR spouse OR relative OR partner OR family) and the intervention (intervention OR coping skills OR psychosocial OR problem-solving OR iCBT OR CBT OR cognitive therapy OR behavioral intervention OR cognitive intervention OR home practice OR e-Health OR cognitive restructuring OR exposure OR mindfulness, OR meditation OR relaxation training OR cognitive behavior therapy OR cognitive behavioral therapy OR online therapy OR online treatment OR internet treatment, internet-based therapy OR psychotherapy). Filters limiting the search to peer-reviewed studies on the adult, human population written in English were employed.

Two independent searches were conducted by MSO and MR for the period from the earliest time available
CJTs for cancer caregivers

through January 2014. In addition, a backward search
(snowballing) was conducted of reference lists of identi-
ified articles and earlier systematic reviews together with
a forward search (citation tracking) until no additional
relevant articles were found.

Data extraction

Studies were coded and rated for type of outcome (mas-
tery, psychological well-being, interpersonal well-being,
physical well-being, and generic quality of life), caregiver
characteristics (mean age, percent women), intervention
characteristics (explicit CBT framework [yes, no], treat-
ment recipient [IC only or couple/dyad], treatment format
[individual or group], treatment modality [face-to-face,
web/phone-based, combination], treatment duration
[weeks from pre to post therapy], number of treatment ses-
sions, number of cognitive-behavioral treatment compo-
nents), patient characteristics (disease stage [early (i.e. I
or II), late (i.e. III or IV), survivors], time since diagnosis),
and study quality characteristics (trial type [RCT, open
trial; OT]), control type [active control group vs. non-
active], and quality (Jadad score; [30]).

All outcomes were categorized according to type as fol-
loows: Mastery refers to appraisal efforts, self-efficacy,
coping skills, knowledge about cancer, and ability to per-
form caregiver related tasks of assisting the patient. Psy-
chological well-being refers to mood, distress (e.g.
anxiety and depression), and overall mental quality of
life. Interpersonal well-being concerns social support,
quality of communication with family and cancer patient,
intimacy, sexual satisfaction, and overall quality of rela-
tionship. Physical well-being refers to the presence of
physical symptoms, exercise habits, physical aspects of
sexual performance, and overall physical quality of life.
Finally, generic quality of life concerns global measures
of quality of life that could not be categorized as either
psychological, interpersonal, or physical well-being. Re-
garding control type, a control condition was considered
active if participants received psychoeducation but not
one of the methods described as defining CBT, or if
they received other planned, non-specific psychosocial
support.

To determine the quality score, the original 11 Jadad
criteria were used [30]. Five criteria relevant for the type
of studies reviewed in the present paper were added: (a)
Was an active control condition included (other than
waitlist)? (b) Was there a clear description of the control
(comparison) group(s)? (c) Was there a clear description
of therapist/interventionist background and level of com-
petency? (d) Were the statistical methods clearly de-
scribed? (e) Are study reports free of suggestion of
selective outcome reporting [cf. 31])? These five criteria
address methodological clarity regarding the intervention,
the degree of the findings’ specificity to CBT, and
potential biases in reporting. Together, the quality ratings
yielded a total modified Jadad score ranging between 0
and 16. Quality scores were not used as weights when
calculating effect sizes, as this is not recommended [32].

Fifteen studies did not report an effect size or means
and standard deviations. Therefore, the authors of those
studies were contacted with a request for the relevant in-
formation or data. Twelve authors responded out of which
five were able to provide data. For studies where it was
possible, an effect size was computed based on statistics
other than a mean and standard deviation, for instance a
\( r \) and \( p \)-value. It was not possible to calculate an effect size
for four studies, in which cases it was set to 0.

All codings and ratings were provided by the first (MO)
and third (MR) author. Disagreements were discussed and
solved by consensus. Literature search and data extraction
protocols are available upon request.

Analytic overview

Meta-analyses were performed to determine both the
pooled effects size for the effect of the CBTs on the com-
bined and individual outcomes based on random-effects
models. Effects were averaged within and across out-
comes so that any given study in any given analysis was
only represented once in order to satisfy the assumption
of independence between observations [33]. Given the
large number of participants, resulting in small effect sizes
being significant, results were mainly interpreted with
regard to the produced effect size.

A number of moderation analyses were conducted with
meta-regression analyses, based on random-effects
models and estimated with the Maximum Likelihood
method. All moderators were analyzed both individually
(unadjusted models) and together in models combining
moderators concerning caregiver, intervention, or patient
characteristics (adjusted models). Regarding study charac-
teristics, these moderators were not evaluated in a com-
bined, adjusted model. This would not be meaningful
because of overlap between variables, where most OTs
had no control group, and because design and control
characteristics were included in the modified Jadad-score.

Effect sizes were expressed as Hedge’s \( g \) instead of
Cohen’s \( d \), given the former (and not the latter) adjusts
for a potential bias to overestimate the effect size in small
samples [34], and a \( p \)-value < 0.05 was considered signif-
ificant. Positive effect sizes indicate an effect of CBTs in
the expected direction. Each effect size was weighted by
its precision (inverse variance). Attrition was large in
several of the studies, and when available, the \( N \) used in
the calculation was the \( N \) in the final analysis for each
outcome.

Heterogeneity was explored using \( Q \) and \( I^2 \) statistics.
\( Q \)-tests concern the probability that results reflect systematic
between-study differences. Because of the generally low
statistical power of heterogeneity tests, a \( p \)-value \( \leq 0.10 \) was used to determine significant heterogeneity [35]. The \( F^2 \) statistic is an estimate of the degree of observed heterogeneity unexplained by sampling error and is unaffected by the number of studies. \( F^2 \) values of 0%, 25%, 50%, and 75% are considered negligible, low, moderate, and high, respectively [36].

Positive and negative findings may not be equally likely to get published, thereby introducing risk of publication bias. The distribution of effect sizes was visually inspected by means of funnel plots [37], and tested with Egger’s test [38]. When a possible publication bias was indicated, an adjusted effect size was estimated using Duval and Tweedie’s [39] trim-and-fill method, which imputes missing results and recalculates the effect size. The fail-safe number refers to the number of unidentified or unpublished studies with null findings that will reduce the pooled result to statistical non-significance [40]. If the fail-safe number exceeded \( 5K + 10 \), with \( K \) being the number of studies included in the meta-analysis, the results were considered sufficiently robust in the face of possible publication bias [41].

All analyses were conducted using the Comprehensive Meta-Analysis program, version 3.3.070 [42].

**Results**

**Search results and study characteristics**

Information flow of study selection with reasons for exclusion is presented in Figure 1. The literature searches yielded 1131 unique records of which 36 independent studies were included and subjected to meta-analytic evaluation.

**Study characteristics**

The studies reviewed included a total of 4746 ICs with a mean sample size of 131. Final data were analyzed for 3820 ICs with a mean sample size of 106. For study characteristics see Supplemental Appendix 1. Most studies were RCTs \( (K=27) \) comparing CBTs with a non-active control condition \( (K=21) \). The most common type of treatment was individual (as opposed to group) \( (K=28) \) therapy for couples/dyads \( (K=28) \), and delivered face-to-face \( (K=22) \). Most ICs were providing care to patients with mixed stages of cancer \( (K=14) \). Twelve studies explicitly stated that the intervention adhered to a cognitive behavioral framework. The most commonly employed treatment components (see definition above) were coping skills training \( (K=24) \), problem-solving \( (K=15) \), cognitive restructuring \( (K=14) \), structured homework \( (K=11) \), and relaxation \( (K=10) \). Regarding quality ratings, the two raters showed good inter-rater agreements, agreeing between 80 and 100% on the individual quality criteria. Each disagreement was solved by consensus. The final mean quality rating was 10.1 \( (SD=2.3; \text{range: 5–14 on the 0–16 scale}) \). The lowest scores were found for the criteria of masking (i.e. masking the condition to the participants; \( K=0 \)), blinding (i.e. concealing group allocation to the researchers; \( K=6 \)), and \textit{a priori} power calculations \( (K=8) \).

*Figure 1. Flow chart of study selection*
Pooled effect sizes and between-study differences

The overall combined effect across studies was negligible (Hedge’s $g=0.08$, 95% CI [0.02–0.14]), also when adjusted for publication bias (0.01). See results below and in Table 1. The largest statistically significant effect was found for psychological well-being ($g=0.16$; $K=31$, $p<0.001$). The effect sizes for interpersonal well-being ($g=0.13$; $K=16$, $p=0.006$) and physical well-being ($g=0.13$; $K=18$, $p=0.012$) also reached statistical

Table 1. Pooled effect sizes across outcomes and levels of moderator variables

<table>
<thead>
<tr>
<th>Outcome</th>
<th>Sample size</th>
<th>Heterogeneity</th>
<th>Global effect sizes</th>
<th>Failsafe N$^d$</th>
<th>Criterion</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$K$</td>
<td>$N$</td>
<td>$Q^b$</td>
<td>df</td>
<td>$p$</td>
</tr>
<tr>
<td>Overall combined effect</td>
<td>36</td>
<td>3820</td>
<td>36.1</td>
<td>35</td>
<td>0.016</td>
</tr>
<tr>
<td>Adjusted for publication bias</td>
<td>(49)</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Psychological well-being</td>
<td>31</td>
<td>3044</td>
<td>43.7</td>
<td>30</td>
<td>0.050</td>
</tr>
<tr>
<td>Adjusted for publication bias</td>
<td>(34)</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Interpersonal well-being</td>
<td>16</td>
<td>1664</td>
<td>18.4</td>
<td>15</td>
<td>0.907</td>
</tr>
<tr>
<td>Adjusted for publication bias</td>
<td>(21)</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Physical well-being</td>
<td>18</td>
<td>1812</td>
<td>23.3</td>
<td>17</td>
<td>0.139</td>
</tr>
<tr>
<td>Adjusted for publication bias</td>
<td>(24)</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Generic QoL</td>
<td>10</td>
<td>1292</td>
<td>30.4</td>
<td>9</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Mastery</td>
<td>20</td>
<td>2616</td>
<td>28.5</td>
<td>19</td>
<td>0.075</td>
</tr>
</tbody>
</table>

Proposed moderators

Study design

Open trials | 9 | 334 | 45.8 | 8 | 0.013 | 0.0 | 0.21 | 0.07 – 0.34 | 0.002 |
RCTs | 27 | 3486 | 26.8 | 26 | 0.048 | 3.1 | 0.04 | 0.02 – 0.11 | 0.200 |

Control condition

Active control | 9 | 1213 | 14.3 | 8 | 0.075 | 43.9 | 0.03 | — — | 0.725 |
Non-active control | 21 | 2508 | 6.7 | 20 | 0.098 | 0.0 | 0.07 | — — | 0.068 |

Therapeutic framework

CBT | 12 | 1100 | 8.24 | 11 | 0.069 | 0.0 | 0.09 | — — | 0.125 |
Other | 24 | 2720 | 27.8 | 23 | 0.225 | 17.2 | 0.08 | 0.00 – 0.16 | 0.045 |
Adjusted for publication bias | (33) | — | — | — | — | — | 0.01 | — — | — |

Intervention Modality

Face-to-face | 22 | 1886 | 27.4 | 21 | 0.057 | 23.5 | 0.11 | 0.01 – 0.22 | 0.037 |
Adjusted for publication bias | (30) | — | — | — | — | — | <0.01 | — — | — |
Webphone | 7 | 789 | 2.4 | 6 | 0.084 | 0.0 | 0.03 | — — | 0.637 |
Combined | 9 | 1356 | 6.0 | 8 | 0.645 | 0.0 | 0.09 | — — | 0.063 |

Intervention recipient

Caregiver only | 8 | 676 | 2.3 | 7 | 0.042 | 0.0 | 0.13 | — — | 0.064 |
Dyad/group | 28 | 3144 | 33.0 | 27 | 0.197 | 18.2 | 0.08 | 0.00 – 0.15 | 0.050 |
Adjusted for publication bias | (39) | — | — | — | — | — | —0.01 | — — | — |

Intervention format

Individual | 28 | 3315 | 30.5 | 27 | 0.291 | 11.5 | 0.07 | — — | 0.059 |
Group | 8 | 505 | 4.2 | 7 | 0.760 | 0.0 | 0.16 | 0.01 – 0.32 | 0.045 |
Adjusted for publication bias | (9) | — | — | — | — | — | —0.14 | — — | — |

Patient disease stage

Early stage | 6 | 255 | 2.7 | 5 | 0.074 | 0.0 | 0.08 | — — | 0.396 |
Late stage | 9 | 1364 | 18.5 | 8 | 0.018 | 56.8 | 0.05 | — — | 0.509 |
Mixed | 14 | 1888 | 7.9 | 13 | 0.849 | 0.0 | 0.09 | — — | 0.046 |
Adjusted for publication bias | (18) | — | — | — | — | — | 0.07 | — — | — |
Survivors | 3 | 128 | 1.7 | 2 | 0.435 | 0.0 | 0.31 | 0.05 – 0.56 | 0.021 |

Note. $K$ and $N$ do not necessarily add up because of exclusion of non-independent samples/studies from the comparison analyses. Statistically significant results are highlighted in bold.

*Possible publication bias was examined with funnel plots and Egger’s test, followed by imputation of missing studies [39]. ($K=K$ number of imputed studies).

$^b$Q-statistic: p-values < 0.1 taken to suggest heterogeneity. $^c$Hedge’s g; Standardized mean difference, adjusting for small sample bias. A positive value indicates an effect size in the hypothesized direction, i.e. improvement following CBT. To ensure independency, if a study reported results for more than one measure, effect sizes were combined (mean), ensuring that only one ES per study was used in the calculation. Conventions: small (<0.3); medium (0.5); large (>0.8).

In case of statistically significant effect sizes, it was planned to examine the robustness of findings by calculating the Failsafe $N$ (number of non-significant studies that would bring the $p$-value to non-significant ($p>0.05$)) [40].
significance. Effect sizes for generic quality of life (g=0.04; K=10, p=0.868) and mastery (g=0.07; K=20, p=0.138) were non-significant. The results of the heterogeneity tests indicated no statistically significant systematic differences between effects on interpersonal (p=0.139, \(I^2=0.0\)) or physical well-being (p=0.139, \(I^2=27.2\)). However, there was statistically significant heterogeneity of the effects on psychological well-being (p=0.050, \(I^2=31.4\)), mastery outcomes (p=0.075, \(I^2=33.3\)), and the generic quality of life outcomes (p=0.001, \(I^2=70.4\)).

Setting the effect size to 0 in the four instances of a missing effect size may be too conservative. We therefore calculated the mean effect size excluding the four studies (g=0.09) and repeated the main analysis imputing this effect size. This did not change the overall effect (g=0.08).

Furthermore, in order to evaluate the influence of possible outliers, the standard deviation for Hedge’s g across outcomes was first estimated (SD=0.20). A search for outliers above or below two standard deviations from the pooled effect size (range: -0.32–0.40) was conducted. Only one study fell outside of this range (g=0.82; [43]). The effect size for this study was winsorized by replacing it with the upper value of the range (0.40), thereby retaining the study with an attenuated influence [44]. Re-analyzing the pooled effect size across outcomes with the winsorized effect size revealed an effect of similar magnitude (g=0.07, 95% CI [0.12–0.13]). Consequently, this effect size was not adjusted in the following analyses.

Moderator analyses

Pooled effect sizes can be found in Table 1, and results from meta-regression-based moderation analyses are displayed in Table 2. The association between continuous moderators and the magnitude of the effect is expressed in unstandardized regression coefficients (B).

Concerning study quality characteristics, the difference in effect size magnitude depended on study design (p=0.028), showing that the effect size was larger for OTs (g=0.21) than for RCTs (g=0.04). Of the RCTs, two studies had two different control conditions and were therefore excluded from the analyses concerning possible difference in effect size magnitude between active and non-active control groups. Results revealed a non-significant difference (p=0.059) between active (g=0.03) and non-active (g=0.07) control groups. The modified Jadad-score was not associated with the magnitude of the effect (B < -0.01, p=0.874).

Exploring the role of caregiver characteristics showed that only the percentage of women was a significant moderator of the effect size when evaluated separately. However, when evaluated together, both age (B = -0.02, p=0.046) and percentage of women (B=0.01, p=0.001) were significantly associated with the magnitude of the effects, with younger age and more female participants both being associated with larger effects. Because age became significant when evaluated together with the percentage of women, the age x women interaction term was explored post hoc, which did not reach statistical significance (B < -0.01, p=0.731). It was explored if participants in studies with more women were younger, but the opposite was true as age and percentage of women was positive correlated (r=0.43, p=0.027). Finally, five different regression models were tested, in which percentage of age and gender predicted the five individual outcomes. Higher percentage of female participants was associated with larger effects on physical well-being (B < 0.01, p=0.029) and mastery (B=0.01, p=0.001), but was not associated with psychological well-being (B < -0.01, p=0.616), interpersonal well-being (B < 0.01, p=0.493), or generic quality of life (B=0.01, p=0.184).

A number of intervention characteristics were also evaluated as moderators. Studies stating to be mainly oriented towards a cognitive-behavioral framework (K = 12)...

---

**Table 2. Results from meta-regression-based moderation analyses**

<table>
<thead>
<tr>
<th>Variable</th>
<th>Unadjusted model(^a)</th>
<th>Adjusted model(^b)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>B</td>
<td>SE</td>
</tr>
<tr>
<td>Study quality characteristics(^c)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Design (RCT vs. OT)</td>
<td>-0.17</td>
<td>0.08</td>
</tr>
<tr>
<td>Control type (Active vs. non-active)</td>
<td>-0.15</td>
<td>0.08</td>
</tr>
<tr>
<td>JADAD</td>
<td>0.00</td>
<td>0.02</td>
</tr>
<tr>
<td>Caregiver characteristics</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age</td>
<td>-0.01</td>
<td>0.01</td>
</tr>
<tr>
<td>% women</td>
<td>&lt;0.01</td>
<td>&lt;0.01</td>
</tr>
<tr>
<td>Interaction(^d)</td>
<td>-0.00</td>
<td>&lt;0.01</td>
</tr>
<tr>
<td>Intervention characteristics</td>
<td></td>
<td></td>
</tr>
<tr>
<td># sessions</td>
<td>-0.01</td>
<td>0.01</td>
</tr>
<tr>
<td>Treatment duration</td>
<td>-0.01</td>
<td>&lt;0.01</td>
</tr>
<tr>
<td># components</td>
<td>0.02</td>
<td>0.02</td>
</tr>
<tr>
<td>CBT (CBT vs. other)</td>
<td>0.01</td>
<td>0.08</td>
</tr>
<tr>
<td>Recipient (IC vs. group/dyad)</td>
<td>0.05</td>
<td>0.09</td>
</tr>
<tr>
<td>Modality (face-to-face vs. web/phone)</td>
<td>0.04</td>
<td>0.08</td>
</tr>
<tr>
<td>Format (individual vs. group)</td>
<td>-0.09</td>
<td>0.09</td>
</tr>
<tr>
<td>Patient characteristics</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Stage</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mixed (vs. early)</td>
<td>0.01</td>
<td>0.12</td>
</tr>
<tr>
<td>Late (vs. early)</td>
<td>-0.05</td>
<td>0.12</td>
</tr>
<tr>
<td>Survivor (vs. early)</td>
<td>0.21</td>
<td>0.18</td>
</tr>
<tr>
<td>Time since diagnosis</td>
<td>0.02</td>
<td>0.04</td>
</tr>
</tbody>
</table>

Note. Statistically significant p-values are in bold. K = number of studies in adjusted model.

\(^a\)Variables were explored individually in unadjusted models.

\(^b\)Variables within the same group of characteristics were explored together in adjusted models.

\(^c\)The three variables concerning study quality were not explored in a combined model due an overlap between variables.

\(^d\)Two models were tested concerning caregiver characteristics, one with and one without the interaction term. Results for age and number of women refer to the model without the interaction term.
obtained a small, non-significant effect size ($g = 0.09$, $p = 0.125$). An effect size for studies not claiming to adhere to a cognitive-behavioral framework ($K = 24$) obtained an effect size of similar magnitude ($g = 0.08$, $p = 0.045$), and the between-study difference did not reach statistical significance ($p = 0.132$). Concerning delivery mode, face-to-face therapy ($g = 0.11$, $p = 0.037$) obtained a numerically larger effect size than therapy delivered over the phone or the Internet ($g = 0.03$, $p = 0.637$), but the difference was not statistically significant ($p = 0.516$). Studies using combined delivery modes also obtained a small, non-significant effect size ($g = 0.09$, $p = 0.063$). The effects of studies providing therapy for the IC only ($g = 0.12$, $p = 0.064$) or for the couples/dyads ($g = 0.08$, $p = 0.050$) were both of a small magnitude and did not differ between studies ($p = 0.272$). The same was true for treatment delivered individually ($g = 0.07$, $p = 0.059$) or in groups ($g = 0.16$, $p = 0.045$), where no between-study difference was detected ($p = 0.309$). Number of treatment sessions ($B = -0.01$, $p = 0.484$), treatment duration ($B = -0.01$, $p = 0.154$), and number of CBT components ($B = 0.02$, $p = 0.454$) were not associated with the magnitude of the effects.

Patient characteristics were also explored as potential moderators. Patient disease stage did not moderate the effect. ICs of survivors obtained a numerically larger effect ($g = 0.31$, $p = 0.021$) than ICs of patients with early stage cancer ($g = 0.08$, $p = 0.396$), late stage ($g = 0.05$, $p = 0.509$), and mixed stages ($g = 0.09$, $p = 0.046$). Only the effect for ICs of patients with mixed stages and survivors reached statistical significance. Time since diagnosis was not statistically significantly associated with the magnitude of the effect ($B = 0.02$, $p = 0.431$).

Retrospective power analyses showed that the statistical power to detect significant associations between the individual proposed moderators and the effect size varied between 0.1 and 0.9.

Publication bias

For all statistically significant results, the risk of publication bias was evaluated. For nine of the 23 analyses, effect sizes were asymmetrically distributed, as determined by a significant Egger’s test, indicating possible publication bias. Adjusted effect sizes can be found in Table 1. Furthermore, all analyses failed to meet the criterion for the fail-safe $N$, indicating a lack of robustness of the results.

Discussion

Overall, the effect of CBTs for ICs was negligible. Although the effect across all outcome types was statistically significant, and the between study variance was largely homogenous, the robustness of the effect was poor, as indicated by a small failsafe number. The largest and statistically significant effects were found for psychological well-being, physical well-being, and interpersonal well-being. However, the robustness was also poor for these effects. There could be several reasons for these findings.

First, the magnitude of the effect could be associated with study quality characteristics. Comparing studies that used a randomized controlled versus an open design revealed that RCTs obtained a non-significant and smaller effect than OTs. Because an RCT design is a more rigorous test of an intervention’s effect, it may be less surprising that the effect was smaller in the RCTs. No moderating effect was found for the remaining study quality variables, including control type and overall quality of the study (modified Jadad score). This is in line with other findings, where quality scores have generally been found to be poor predictors of study results [32,45].

Second, the evaluated studies, although considered CBTs, varied in main theoretical framework, and it is possible that interventions that did not state to be mainly oriented towards a cognitive-behavioral framework might not be a true CBT. However, the most commonly employed treatment components were coping-skills training, problem-solving techniques, cognitive restructuring, the use of structured home-work, and relaxation techniques, which are all frequently used techniques across CBTs [e.g. 46]. Furthermore, self-claimed therapeutic framework did not moderate the effect. Other intervention characteristics could also be hypothesized to moderate the effect. For instance, one could argue that the flexibility characterizing phone and web interventions would be very suitable for the caregiver population who, because of caregiver responsibilities, may not be able to leave the home. However, the effect of CBTs delivered over the phone or web was non-significant and numerically smaller than that obtained for face-to-face CBTs, albeit not significantly different. The finding that treatment recipient did not moderate the effect is somewhat surprising because it has often been proposed that coping with cancer occurs within an interpersonal system, that is, the affected couple engages in dyadic coping [e.g. 47,48]. However, the present data suggest that – across outcomes – the effect of CBTs does not differ between interventions for the IC only or the couple.

Third, as mentioned in the introduction, there is great variation in levels of distress within this population, and some ICs’ distress levels may have been rather low to begin with, leaving little room for improvement. Related to this, attrition rates could have affected the outcome in the sense that the most distressed ICs may have dropped out of treatment. However, when available, the $N$ used in the calculation was the $N$ in the final analysis for each outcome, thereby adjusting the analyses for this potential issue.
Finally, in the traditional cognitive treatment model, the client’s emotional struggles are primarily viewed as a result of erroneous and maladaptive cognitions [49]. However, when working with ICs of patients with cancer, there is often validity to the negative thoughts, and traditional cognitive methods may therefore be experienced by the client as invalidating [50]. This could be one explanation for the larger effect size detected for survivors, where the situation is characterized by a factually better outlook.

The present findings could encourage future investigations of psychological interventions within a cognitive-behavioral framework to look beyond traditional methods. We believe that recent advances in the basic affective sciences may elucidate caregiver treatment. These advances emphasize functional and motivational aspects of emotional responses, and how these motivations can be in conflict and lead to distress [51,52]. A variety of motivational conflicts may arise as a result of taking responsibility for the care of a patient with cancer. A prominent conflict for ICs is balancing self-care and care for the patient, in which the IC may, for example, feel the need to leave the house or the hospital and at the same time experience fear and guilt in thinking about leaving the patient. The IC may also wish to have positive experiences with the patient while concurrently feeling distress about the anticipated loss of their loved one [53]. Such conflicts can manifest themselves as intense and painful emotions that may be difficult to handle. For instance, when faced with intense or painful emotions, individuals with anxiety and depression – emotional distress characteristic of ICs – often use emotion regulation strategies aimed at diminishing or avoiding these emotions, resulting in increased negative emotions and exacerbated stress-responses [54–56].

One could therefore argue that it would be relevant to offer ICs psychotherapy specifically aimed at handling motivational conflicts and emotion regulation. Indeed, these emotional processes are a core feature of the ‘third wave’ behavior therapies as well as other contemporary CBTs [e.g. 57–59], in which realistic reappraisals are granted no or a minor role. Congruently, the authors of the present paper are currently running a trial investigating Emotion Regulation Therapy for ICs [58,60] (ClinicalTrials.gov identifier: NCT02322905), an approach that draws directly from this basic affect science framework and offers specific interventions aimed at improving motivational awareness and emotion regulatory ability.

In addition to the search for other, more effective interventions for ICs, attention should also be paid to caregiver characteristics that may be associated with better treatment response. In the present study, the percentage of female participants was positively associated with the effect of CBT, and – when considered together with percentage of women – age was negatively associated with the magnitude of the effect. In further exploring the effect of gender, it was found that the number of female participants was associated with larger effects concerning physical well-being and mastery. Identifying characteristics associated with better treatment response may help target different treatments to different subgroups of ICs.

Limitations of the present review include a broad definition of CBT, leading to the inclusion of a diverse set of therapies. However, the definition of CBTs reflects the large variety in what historically have been and currently are considered cognitive behavioral therapies [e.g. 46], holding central elements hypothesized to drive the treatment effect in both traditional and contemporary CBTs [61,62]. Furthermore, the inclusion criteria regarding the IC population varied between studies, limiting the internal validity. In some cases the IC was the primary caregiver responsible for many elements of physical and emotional care, whereas other studies defined ICs as an intimate partner without regard to actual caregiving demands. Similarly, some studies enrolled the ‘closest relative’, a category that included individuals in varying relationships with the patient (e.g. spouse or child). However, despite including studies representing variation in intervention type and ICs, results indicated relative homogeneity among studies. Additional data were requested from 15 authors, and data were only received from five. A number of proposed moderator variables were explored, but only few were significantly associated with the effect, which may in part be because of low statistical power. This, combined with the exploratory nature of the analyses, suggests that the results should therefore be interpreted with caution. Furthermore, the temporal burden experienced by ICs may have limited their adherence to homework assignments, thereby mitigating treatment response. Finally, four of the five outcome groups were categorized as ‘well-being’. However, for the psychological well-being outcome group, the vast majority of measures concerned distress, for instance anxiety and depression. The lack of distress does not necessarily constitute psychological well-being. There is now emerging evidence for the potential rewards of providing care, including gaining meaning in life and increased appreciation of others [e.g. 63,64]. The effect of CBTs on such growth or gains remains uninvestigated.

Conclusions

Overall, the results indicated that CBTs for ICs had a negligible effect. A number of moderators were explored of which only few reached statistical significance suggesting that CBTs may be more effective for younger, female ICs. However, when compared to a control
group in randomized designs, the effect of CBTs did not reach statistical significance. A number of reasons could be hypothesized to underlie this negligible effect, and it is suggested that future studies move beyond traditional CBT methods and orient towards recent advances in the basic affective sciences in order to better understand and treat the emotional struggles experienced by ICs.

**Conflict of interest**

The authors declare that there are no conflicts of interest.

**Funding resources**

This work was supported by the Danish Cancer Society.

**References**


Supporting information

Additional supporting information may be found in the online version of this article at the publisher’s web site.