Review article

Laughter-inducing therapies: Systematic review and meta-analysis

C. Natalie van der Wal[a,b,c,*], Robin N. Kok[b,d,e]

[a] University of Leeds, Leeds University Business School, Centre for Decision Research, Leeds, United Kingdom
[b] Vrije Universiteit, Faculty of Behavioural and Movement Sciences, Department Clinical, Neuro & Developmental Psychology, Section Clinical Psychology, Amsterdam, Netherlands
[c] Vrije Universiteit, Department of Computer Science, Section Artificial Intelligence, Amsterdam, Netherlands
[d] University of Southern Denmark, Faculty of Health Sciences, Institute for Psychology, Odense, Denmark
[e] Centre for Innovative Medical Technology, Odense University Hospital, Odense, Denmark

ABSTRACT

Rationale: Laughter-inducing therapies are being applied more regularly in the last decade, and the number of scientific reports of their beneficial effects is growing. Laughter-inducing therapies could be cost-effective treatments for different populations as a complementary or main therapy. A systematic review and meta-analysis has not yet been performed on these therapies for different populations and outcomes, but is needed to examine their potential benefits. This research aims to broadly describe the field of laughter-inducing therapies, and to estimate their effect on mental and physical health for a broad range of populations and conditions.
Method: A systematic review of the field was undertaken, followed by a meta-analysis of RCTs and quasi-experimental studies. The systematic review included intervention studies, one-session therapies, lab studies and narrative reviews to provide a broad overview of the field. The meta-analysis included RCTs or quasi-experimental studies that assessed multi-session laughter or humor therapies compared to a control group, performed on people of any age, healthy or with a mental or physical condition. English and non-English articles were searched using PubMed, Web of Science, EBSCO and EMBASE. Search terms included laugh(ing), laughter, humo (u)r, program, therapy, yoga, exercise, intervention, method, unconditional, spontaneous, simulated, forced. Studies were classified as using humor ('spontaneous' laughter) or not using humor ('simulated' laughter).
Results: This systematic review and meta-analysis suggests that (1) 'simulated' (non-humorous) laughter is more effective than 'spontaneous' (humorous) laughter, and (2) laughter-inducing therapies can improve depression. However, overall study quality was low, with substantial risk of bias in all studies. With rising health care costs and the increasing elderly population, there is a potential for low-cost, simple interventions that can be administered by staff with minimal training. Laughter-inducing therapies show a promise as an addition to main therapies, but more methodologically rigorous research is needed to provide evidence for this promise.

1. Introduction

There is some evidence that laughter has physical, emotional, and social benefits (Bennett et al., 2014; Mora-Ripoll, 2011; Yim, 2016). However, scientific research is still in an early stage when it comes to empirically determining the therapeutic value of laughter. Different physiological and psychological effects of laughter have been anecdotally reported, e.g. decreasing pain, strengthening immune function, mitigating stress and improving social support (Bennett and Lengacher, 2006; Martin, 2001; Mora-Ripoll, 2011). Laughter is presumed to decrease levels of stress hormones, and theorized to buffer the effects of stress on the immune system and thus elevate our mood (Bennett and Lengacher, 2009). Current literature broadly distinguishes between

'simulated' versus 'spontaneous' laughter (Mora-Ripoll, 2011; Yim, 2016). Spontaneous laughter is laughter triggered by a stimulus (e.g., a joke). This typically causes contractions around the eye sockets (Duchenne laughter/smile or 'humorous laughter'; Ekman et al., 1990), and is often perceived as a 'genuine' laugh. Simulated laughter is voluntarily and consciously triggered by oneself (self-induced, a 'fake' laugh, or non-humorous laughter), in a controlled environment, for no specific reason, and by definition is not caused by humor or other stimuli. The term forced laughter has also been coined in the literature as meaning simulated laughter (Foley et al., 2002), but is better known as pathological, uncontrollable laughter (Black, 1982). In this article, we use the terms 'spontaneous' for humorous, stimulus-induced laughter and 'simulated' for non-humorous, self-induced laughter.

---

* Corresponding author. University of Leeds, Leeds University Business School, Centre for Decision Research, Leeds, United Kingdom.
E-mail addresses: c.n.vanderwal@leeds.ac.uk (C.N. van der Wal), rkok@health.sdu.dk (R.N. Kok).

Following the saying that laughter is the best medicine, in the past twenty years, 'spontaneous' (humorous) and 'simulated' (non-humorous) laughter have been applied in humor or laughter therapies in a wide range of settings to improve health and well-being. These therapies have been tried in a range of settings, from children to the elderly, and with a broad range of targeted outcomes such as mental health, cancer, diabetes, migraine, and other chronic conditions. Despite many studies on myriad forms of laughter therapies for different patient populations or healthy individuals; as yet there has been no comprehensive systematic review or meta-analysis to assess whether these applications are effective or not. In this study, we focus on laughter-inducing therapies that are compared to control groups (no treatment or attention control). We also investigate the outcome differences based on how laughter was induced, either through 'spontaneous' or simulated' laughter.

Laughter-inducing therapies come in different formats. Humor therapies typically include laughter exercises with humor such as humorous videos or clowns (Brodaty et al., 2014; Low et al., 2013; Szabo, 2003; Tan et al., 2007). Laughter therapy without using humor typically includes exercises such as clapping, dancing, and vocalizing laughter-like sounds like "hoho-hahaha", but can also include elements not involving laughter, such as breathing and relaxation exercises (Yim, 2016). Laughter Yoga is a specific example of non-humorous laughter-inducing therapy; is encouraged to be done in groups, involving laughter exercises, clapping, and yoga exercises like breathing and relaxation exercises. (Cokolic et al., 2013; Farifeth et al., 2014; Miles et al., 2016; Nagendra et al., 2007; Yazdani et al., 2014). Overall, most of these studies report an improvement of physical or mental health, as compared to a control group. The control group typically receives usual care, no intervention, or an attention control condition where e.g. plants are watered or crafts projects are made (Cai et al., 2014; George and Jacob, 2014; Jung et al., 2009; Kim et al., 2015; Low et al., 2013). Improvements such as decreased depression, pain, and stress hormones, and improved mood and life satisfaction, were reported and will be discussed briefly.

Previous reviews have attempted to summarize the field. McCreaddie and Wiggins (2008) reviewed the direct and indirect links between humor and health, specifically in nursing applications. They found that research designs were lacking methodologically, and were unable to provide strong evidence for humor having a positive effect on health. Furthermore, most included studies were correlational and lacked the appropriate randomized controlled design to infer causality.

Mora-Ripoll (2011) conducted a narrative literature review of both 'simulated' and 'spontaneous' laughter therapies, and concluded that there is some evidence that 'simulated' laughter (non-humorous laughter) has positive effects on health compared to control groups (waiting list or receiving no intervention), other experimental groups (exercise therapy), or not compared to another group (interventional study). He also noted that there are practically no contraindications of laughter, and that few adverse effects were reported. These conclusions were based on a combination of randomized, interventional, observational, and non-randomized studies.

Bennett et al. (2014) conducted a narrative review of laughter and humor therapy specifically for patients undergoing dialysis. They concluded that laughter and humor therapies have positive effects compared to control groups (no intervention or active control) on immunity, pain, sleep quality, respiratory function, depression, and anxiety; all of which are relevant for patients undergoing dialysis. Bennett and colleagues concluded that non-humorous laughter therapies, such as Laughter Yoga, are suitable for dialysis patients. However, they noted that it is unclear if these health benefits are sustained in the long-term. Again, these conclusions were based on a combination of both randomized and non-randomized studies.

Gonot-Schoupinsky and Garip (2018) conducted a systematic review of laughter and humor interventions for adults over 60 years old. They concluded that these interventions appear to enhance well-being,

but that there is insufficient evidence that the laughter itself is causing the enhancement, as there are a range of confounding factors. They note that participant laughter has to be isolated and measured to build evidence for these interventions.

Overall, all four reviews concluded that sufficient evidence exists to suggest health benefits of laughter and humor, but that clear guidelines for laughter and humor therapies are necessary. In two reviews (Bennett et al., 2014; Mora-Ripoll, 2011), 'simulated' (non-humorous) laughter was mentioned as specifically interesting for laughter therapies, due to its applicability to many different populations and settings.

Against a backdrop of increasing healthcare expenditure, there is a potential for simple, cheap, broadly applicable, and easily implementable therapies such as laughter therapies as a cost-effective addition to regular healthcare. Therefore, there is a need for a systematic review and meta-analysis of laughter-inducing therapies to explore their potential efficacy and effectiveness. Moreover, the existing (narrative) reviews are due for an update with evidence from recently published laughter-inducing interventions.

The goal of this article is to provide a systematic review of the literature on laughter-inducing therapies. Specifically, we aimed to summarize the state of the literature on laughter-inducing interventions for a broad range of (mental) health outcomes. In this work, we focus on laughter-inducing therapies that are compared to control groups (no treatment or attention control), and, following previous literature, looking specifically at how laughter was induced: using humor or not using humor. Sufficient outcome data were retrieved to perform a meta-analysis of randomized and quasi-experimental laughter-inducing therapy trials.

2. Method

2.1. Search strategy

All searches in PubMed, EMBASE, Web of Science (WoS) and EBSCO were performed in October 2016 and continuously updated until December 2017 using the following queries (for WoS and EBSCO): TITLE: (laugh* OR humo$r) AND TITLE: (program* OR therapy OR yoga OR exercise OR intervention OR method OR unconditional OR spontaneous OR simulated OR forced) NOT TITLE: ("aqueous humor" OR "aqueous humor" OR "vitreous humor" OR "vitreous humor"). Exact search terms can be found in Online Supplement Appendix I. Initial searches for laughter-inducing therapies in the grey literature were performed from July until October 2016 to determine a minimum of 10 benchmark articles for the complete search in the databases. Google Scholar was searched with the search terms: Laughter, Laughter yoga, Unconditional laughter, Simulated laughter, Laughter therapy, Laugh, Laughing, Laugh therapy, Review. Eligible studies on the first 10 Google Scholar pages were added. References from eligible studies and review articles were checked for articles to add. Authors of included articles and experts in the field were contacted for additions. Journals in which eligible studies were published were hand-searched for more studies. This resulted in a list of 13 benchmark studies for the complete search in the databases mentioned above, as well as providing additional search terms (Beckman et al., 2007; Bennett et al., 2003, 2014, Chang et al., 2013, Cho and Oh, 2011, De La Fuente and Gonzalez, 2010, Dolgoff-Kaspar et al., 2012, Foley et al., 2002, Hirosaki et al., 2013, Hsieh et al., 2015, Ko and Youn, 2011, Mora-Ripoll, 2011, Nagendra et al., 2007, Neuhoff and Schaefer, 2002, Raja and Sundari, 2014., Sakai et al., 2013, Shahidi et al., 2011, Yazdani et al., 2014, Weinberg et al., 2014). Of all included studies for the systematic review, one article was in Spanish, one in German and 31 in Korean. The articles were translated with Google Translate. One of the authors reads German and Spanish, verified the translations against the original text and found that the translations were sufficient to extract the necessary data. The translations from Korean to English were first checked against the provided English language abstract. The authors and a Korean

translator looked at the translations to see if the data could be extracted. Even though the grammar was suboptimally translated by Google Translate, we concluded that sufficient data could be extracted as most tables with statistics were understandable and the translator clarified mistranslations from Google Translate.

2.2. Eligibility criteria

We included studies in the systematic review with (a): people of any age, healthy or with a mental or physical health condition; undergoing (b) laughter-inducing therapy compared to (c): all forms of control or comparison groups (e.g. waiting list control or control group therapy), as assessed by (d) all mental and physical health-related outcomes. Case-control studies were excluded. From this selection, included studies in the meta-analysis (e): had a control group and (f) used multiple intervention sessions.

2.3. Study selection

Eligibility assessment was performed independently in a blinded standardized manner by 2 reviewers (CNvdW and RK), screening first on a title basis, and then on an abstract and full-text basis. In the title and abstract phase, disagreement on whether to include a paper meant this paper would be assessed in more detail at the abstract or full-text level. Final disagreements between reviewers on full-text basis were resolved by discussion until consensus was reached (initial kappa at full-text screening = 0.62; final kappa after discussion = 1). Disagreements were about whether articles should be included in the systematic review if it included (1) more background information on laughter, but no report from an experiment, or (2) experiments with just one laughter session. We agreed to exclude explorative, correlational and one-session studies.

2.4. Data extraction

Study information on the publication language, study design, participants, type of intervention, study outcomes, number of sessions, overall duration of treatment, and humorous/non-humorous therapy, was extracted by one reviewer (CNvdW) and independently cross-checked by another reviewer (RK) using a predefined, standardized data extraction spreadsheet. Inconsistencies in data extraction were resolved between the reviewers by referring to the source study, and consensus was reached between both reviewers.

2.4.1. Therapy classification

Therapies were classified based on the way laughter was induced in the therapy, as mentioned in the primary study's methods section. Interventions were classified as 'using humor', 'not using humor' or 'unknown'. The classifications were operationalized as following: an intervention was classified as 'using humor' ('spontaneous laughter') if some humorous stimulus was mentioned as being part of the therapy, such as humor, jokes or humorous videos. Interventions were classified 'not using humor' ('simulated laughter') if it was specifically mentioned that only Laughter Yoga (which is non-humorous per definition) or non-humorous laughter was used, or when all elements of the interventions were clearly mentioned and none of them involved humor. 'Unknown' interventions were classified as such when the content of the 'laughter therapy' was not specifically mentioned, and thus could not be definitively classified as 'humorous' or 'non-humorous' laughter. When interventions combined both non-humorous and humorous laughter, these interventions were classified 'humorous', to indicate the effects of non-humorous laughter therapies versus a therapy that includes only humorous laughter or a mix of laughter including humorous laughter. In case of disagreement between the reviewers, the text was revisited for a stimulus inducing laughter to determine if humor was used or not. All disagreements were due to overlooking this information and after careful inspection of the text, all disagreements were resolved.

2.5. Outcome measures

All mental and physical health outcome measures were included in the systematic review to give a full overview of the field. This was done for two reasons: 1. the absence of a systematic review so far; 2. the relatively low number of studies allowed for this task to be undertaken. In the meta-analysis, only mental health outcomes were used due to the diversity of physiological outcomes, which would have made pooling of results difficult to interpret.

2.6. Analytic approach and data synthesis

Due to the methodological differences in the data set, a two-step approach to data synthesis was used. Firstly, a systematic review was performed on all included studies to provide an overview of the field (see Sections 3.2-3.5). We included quantitative results in this overview as well, such as effect sizes and mean differences with 95% confidence intervals. Adding quantitative results prevents a simple vote-counting procedure, which could leave important information out about the replicability of results, effect sizes, and difference in mean values between the experimental and control groups. Moreover, it adds contextual information on clinical relevance in the case of statistically significant outcomes, for example in cases where outcomes are statistically significant but the difference is so small as to be clinically negligible. The Cohen's *d* effect sizes are corrected by using the pooled pretest standard deviation for weighting the differences of the pre-post-means, d_{ppc2}, (Morris, 2008) or by computing Hedges' *g* for both groups and to subtract them afterwards to correct for different sample sizes and pre-test values, d_{Korr}, (Klauer, 2001).

Secondly, for a subset of quasi-experimental studies and randomized controlled trials, a meta-analysis was performed to estimate separate pooled effect sizes for both study designs. Randomized and quasi-experimental studies were not pooled and analyzed as separate data sets (see Section 3.6) (Verde and Ohmann, 2015). For the randomized studies, the outcome measures were the standardized effect sizes between treatment and control groups as calculated using Hedges' *g*, which resembles Cohen's *d* but attempts to correct the effect size for small studies (Hedges and Olkin, 1985). Studies with insufficient information for calculating effect sizes using Comprehensive Meta-analysis were excluded if sufficient additional information could not be retrieved from article appendices, additional publications from the same data set, or by contacting study authors. Subgroup analyses were conducted using the procedures as implemented in Comprehensive Meta-Analysis, using a mixed effects model (pooled within groups using a random effects model; but differences between groups were tested using a fixed-effects model). We calculated the I^2 and T^2 statistics to assess relative and absolute heterogeneity of effect sizes within subgroups. A common interpretation is that an I^2 value of 0% indicates no observed heterogeneity, and higher values indicate higher observed within-subgroup heterogeneity (25% - low, 50% - moderate, 75% - high; Higgins et al., 2003). A higher observed statistical heterogeneity indicates a higher proportion of observed variance, which can indicate underlying differences between the pooled studies. This makes interpreting the pooled effect size difficult, as it is difficult to distinguish the observed effect size from the true population effect size (Borenstein et al., 2017). Comprehensive Meta-Analysis version 2.2.057 was used for all analyses.

2.7. Assessment of risk of bias

RK and CNvdW independently rated the risk of bias in each study using the Cochrane Collaboration's Risk of Bias Assessment Tool. Studies were rated for adequate sequence allocation, concealment of allocation, selective outcome reporting, masking of assessors, and whether incomplete data were addressed (Higgins et al., 2011). Differences in the assessment of bias were resolved through discussion.
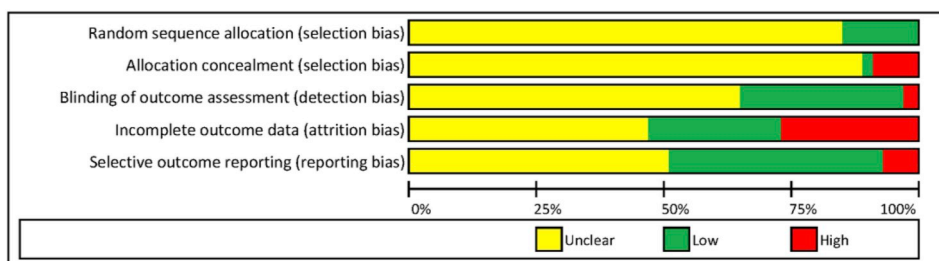
Fig. 1. Risk of bias assessment.

Since treatment allocation was obvious for most studies, blinding of participants was not assessed. Additionally, blinding of outcome assessment was not considered a major risk of bias as most outcomes were self-report measures; nevertheless, this was assessed. Risk of bias is presented graphically in Fig. 1 and per-study results can be found in Online Supplement Table 1. In the meta-analysis, publication bias was tested by inspecting a funnel plot of primary outcome measures, and the 'trim and fill'-procedure was used to correct for publication bias by imputing studies presumed missing due to unpublished negative or null findings (Duval and Tweedie, 2000). Egger's intercept test (Egger et al., 1997) and Begg and Muzamdar's tests (Begg and Mazumdar, 1994) were performed as additional tests of publication bias.

3. Results

3.1. Risk of bias

Overall, reporting quality was low or very low, and crucial information to assess risk of bias was often missing from the studies. Therefore, most studies were at unclear risk of bias for the five outcomes that were assessed. Since none of the assessed studies were pre-registered in a trial registry, we could not evaluate whether outcomes were switched or selectively reported; hence the risk of bias assessment for these elements was performed using only information available in the study. Thus, if a study reported using an outcome measure in the methods section but failed to report the actual outcome of this measure in the results section, this would be rated as being at a high risk of bias. This is a rather lenient interpretation of the selective outcome reporting criterion as used in the Cochrane Collaboration's Risk of Bias Assessment Tool (Higgins et al., 2011). Fig. 1 graphically depicts the overall risk of bias within studies, and the individual assessments are presented in Online Supplement Table 1.

3.2. Systematic review

We screened 98 potentially relevant articles, and finally included 86 studies in our systematic review and 29 in our meta-analysis. Reasons for exclusion from the systematic review were: no intervention ($k=7$), full-text unavailable ($k = 3$), or the text was not translatable using Google Translate ($k = 3$). Reasons for exclusion from the meta-analysis were: the intervention had only 1 session ($k = 12$), there was no control group ($k = 14$), mixed-therapies intervention (1), or is a review article ($k = 3$). The process of study selection is shown in Fig. 2. Therefore, the systematic review includes low quality studies, such as studies without a control group, a small sample size, or a very high risk of bias, but gives a more complete overview. The meta-analysis only includes quasi-experimental studies and randomized controlled trials.

3.3. Study characteristics

The 86 articles in this systematic review can be categorized into three review articles and 83 studies on laughter therapy with humor, without humor, or from which it is unknown if it is with or without

humor. From all 83 articles about interventions, 14 were randomized controlled trials and 41 quasi-experimental pre-test post-test design studies including a control group. The other 31 studies were either pilot studies, field studies without a control group, or had an unknown study design. These study details are presented in Online Supplement Table 2.

Studies originated from the Americas (USA, Colombia), Asia and the Middle East (China, Iran, India, Hong Kong, Korea, Taiwan, Thailand, Japan), Europe (Germany, Slovenia, Switzerland, United Kingdom), Africa (South Africa) and Australia. Outcome measures varied broadly, from mental health and well-being (e.g. agitation, anxiety, cognitive function, coping responses, depression, laughter, life satisfaction, mood, pain, quality of life, resilience, self-efficacy, self-esteem, stress) to physical health (e.g. blood glucose level, blood pressure, body weight, fatigue, heart rate, immune function, insomnia, pulmonary function, sleep quality). Laughter was induced through: clapping, dancing, facial muscle exercises, laughing with a clown, laughter exercises (e.g. laughing 'big' and 'small', giving each other applause, smiling), laughter yoga, and watching (self-selected) humorous videos. Detailed study characteristics, such as study aim, outcome measures, assessment points, limitations, and results are presented in Online Supplement Table 3. The study results will be discussed below.

3.4. Systematic review: broad overview of all included studies

When reviewing the literature, we found three broad main subgroups. The first subgroup is based on three most common outcome measures reported in the studies: depression, anxiety and stress. The second subgroup is based on the elderly population that was used most often in the studies. The third subgroup consists of populations or outcome measures that otherwise do not come to light in the meta-analysis or systematic review, but we feel are important when giving a full overview of the field.

3.4.1. Subgroup 1: most common outcomes - depression, stress, and anxiety

Depression. In 31 different studies, depression was an outcome measure (see Table 1). In 26 studies, depression significantly decreased due to laughter-inducing therapy. Besides the statistical significance, we compared effect sizes and mean differences for all studies. The average corrected effect size for all studies is $d_{\mathrm{dppc2}} = 0.85$ ($d_{\mathrm{Korr}} = 0.80$). The average effect size for randomized controlled trials only is $d_{\mathrm{dppc2}} = 0.57$ ($d_{\mathrm{Korr}} = 0.63$). We also made a distinction between the way laughter was induced in the therapy; either with or without using humor. The average effect size for humorous therapies is: $d_{\mathrm{dppc2}} = 0.43$ ($d_{\mathrm{Korr}} = 0.40$), and for non-humorous therapies: $d_{\mathrm{dppc2}} = 1.14$ ($d_{\mathrm{Korr}} = 1.187$). After removal of an extreme outlier (Kim, 2010), for non-humorous therapies, this dropped to $d_{\mathrm{dppc2}} = 0.73$ and $d_{\mathrm{Korr}} = 0.78$. Overall, all studies show a medium effect size and non-humorous therapies show an effect size twice as large as humorous therapies.

We then looked at 'replicability' of results. This meant, we plot the mean differences and the 95% confidence intervals of studies that measured depression with the Beck Depression Inventory (BDI) and the Geriatric Depression Scale (GDS) respectively (see Figs. 3 and 4). These

Table 1
Effect sizes and mean differences for studies with depression as outcome.

Study	O[a]	T[b]	R[c]	Pre-test intervention	Pre-test control	Post-test intervention	Post-test control	Follow-up intervention	Follow-up control	p[d]	d_corrected[e] Pre-post	d_corrected Pre-flw-up	Mean difference [95% CI][f]
Kim (2010)[g]	BDI	NH	–	1.45 ± 0.13 (n = 12)	1.41 ± 0.1 (n = 12)	0.78 ± 0.13 (n = 12)	1.44 ± 0.16 (n = 12)	0.92 ± 0.17 (n = 12)	1.39 ± 0.11 (n = 12)	< .001[#]	5.316, 5.255[§]	3.873[–], 3.59[§]	–0.7 [–0.82, –0.58]
Park et al. (2013)	BDI	NH	–	19.29 ± 6.65 (n = 17)	16.3 ± 6.02 (n = 20)	14 ± 7.65 (n = 17)	14.15 ± 5.72 (n = 20)	NA	NA	.122*	0.493[–], 0.496[§]	NA	–3.14 [–7.61, 1.33]
Cho & Oh (2011)	BDI	NH	–	14.06 ± 7.93 (n = 16)	11.05 ± 5.56 (n = 21)	10.5 ± 8.01 (n = 16)	10.52 ± 5.67 (n = 21)	NA	NA	.176*	0.450[–], 0.454[§]	NA	–3.03 [–7.07, 1.64]
Cai et al. (2014)	BDI	H	+	36.6 ± 8.3 (n = 15)	38.9 ± 9 (n = 15)	30.7 ± 8.7 (n = 15)	36.9 ± 7.7 (n = 15)	NA	NA	< .005	0.446[–], 0.489[§]	NA	–6.2 [–12.15, –0.25]
Kim et al. (2010)	CES-D	NH	–	15.4 ± 10.89 (n = 30)	14 ± 10.47 (n = 31)	11.33 ± 8.24 (n = 30)	13.26 ± 12.02 (n = 31)	NA	NA	.611[#]	0.310[–], 0.318[§]	NA	–3.33 [–8.55, 1.89]
Cha et al. (2012)	CES-D	NH	–	45 ± 8.8 (n = 28)	45.3 ± 6.75 (n = 29)	37.5 ± 6.03 (n = 28)	42.9 ± 8.27 (n = 29)	NA	NA	< .001[#]	0.649[–], 0.706[§]	NA	–5.1 [–8.89, 1.31]
Low et al. (2014)[h]	CSDD	H	+	8.5 ± 6.1 (n = 189)	7.8 ± 5.6 (n = 209)	6.9 ± 5.2 (n = 197)	6.5 ± 4.5 (n = 197)	6.4 ± 4.8 (n = 162)	6.3 ± 5.4 (n = 181)	.01[#], .89[#]	–0.051[–], –0.037[§]	–0.103[–], –0.1[§]	0.4 [–0.61, 1.41]
George & Jacob (2014)	GDS	?	–	16.7 ± 3.63 (n = 30)	? (n = 30)	11.97 ± 3.55 (n = 30)	17.1 ± 3.22 (n = 30)	13.9 ± 3.52 (n = 30)	19.43 ± 2.68 (n = 30)	< .001*	1.514[&]	1.768[&]	–5.13 [–6.87, –3.39]
Hirosaki et al. (2013)	GDS	H	+	8.7 ± 6.8 (n = 14)	8.3 ± 6.3 (n = 13)	7.1 ± 5.4 (n = 14)	8.8 ± 7.4 (n = 12)	6.3 ± 4.5 (n = 14)	9 ± 7.4 (n = 12)	.211[#]	0.316[–], 0.327[§]	0.467[–], 0.511[§]	–1.7 [–6.81, 3.41]
Hirsch et al. (2010)	GDS	H	–	7.65 ± 3.47 (n = 52)	7.89 ± 3.47 (n = 38)	4.75 ± 3.63 (n = 52)	5.45 ± 3.63 (n = 38)	NA	NA	.001[#]	0.132[–], 0.124[§]	NA	–0.46 [–2, 1.08]
Hsieh et al. (2015)	GDS	NH	–	4.91 ± 4.49 (n = 32)	5.69 ± 4.35 (n = 30)	3.39 ± 3.14 (n = 32)	6.37 ± 3.95 (n = 32)	NA	NA	.001[#]	0.495[–], 0.659[§]	NA	–2.2 [–4.01, –0.39]
Jung et al. (2009)[i]	GDS	?	+	7.98 ± 3.58 (n = 48)	8.08 ± 3.95 (n = 61)	6.94 ± 3.19 (n = 48)	8.43 ± 3.44 (n = 61)	NA	NA	.027[#], .006*	0.366[–], 0.421[§]	NA	–1.39 [–2.65, –0.13]
Konradt et al. (2013)	GDS	H	–	8.73 ± 2.81 (n = 49)	8.56 ± 3.22 (n = 50)	5.39 ± 3.71 (n = 49)	6.12 ± 3.99 (n = 50)	NA	NA	.000[#]	0.297[–], 0.246[§]	NA	–0.9 [–2.44, 0.64]
Lee & Eun (2011)	GDS	H	–	10.4 ± 3.15 (n = 20)	8.45 ± 3.17 (n = 20)	7.4 ± 2.96 (n = 20)	10.45 ± 3.61 (n = 20)	NA	NA	.000[#]	1.571[–], 1.541[§]	NA	–5 [–7.07, –2.93]
Lee & Ji (2011)	GDS	NH	-	?	?	?	?	NA	NA	?	?	NA	?
Lee et al. (2013)	GDS	NH	–	16.9 ± 7.1 (n = 24)	15 ± 7.1 (n = 25)	13.5 ± 7 (n = 24)	15.8 ± 7.9 (n = 25)	NA	NA	< .05[#]	0.588[–], 0.575[§]	NA	–4.2 [–8.43, 0.03]
Park (2013)	GDS	H	–	8.45 ± 2.29 (n = 42)	9.26 ± 2.32 (n = 47)	5.06 ± 3.29 (n = 42)	7.01 ± 3.02 (n = 47)	NA	NA	< .001[#], .02*	0.346[–], 0.127[§]	NA	–1.14 [–2.47, 0.19]
Shahidi et al., 2011	GDS	NH	–	16 ± 5.3 (n = 20)	15.2 ± 3.9 (n = 20)	10 ± 6.9 (n = 20)	15.2 ± 6.1 (n = 20)	NA	NA	< .01[#], .001*	1.28[–], 0.97[§]	NA	–6 [–10.09, –1.91]
Walter et al. (2007)[j]	GDS	H	–	5.3 ± 5.3 (n = 10)	2.5 ± 2 (n = 10)	?	?	?	?	?	?	NA	?
Walter et al. (2007)	GDS	H	–	6.7 ± 3.3 (n = 10)	7.6 ± 5.3 (n = 10)	?	?	?	?	.005[#]	?	NA	?
Yazdani et al. (2014)	GHQ	NH	–	4.7 ± 5.3 (n = 19)	4.3 ± 3.9 (n = 19)	2.2 ± 2.9 (n = 19)	5.2 ± 2.8 (n = 19)	1.7 ± 2.4 (n = 19)	5.1 ± 3.3 (n = 19)	.001[#]	0.725[–], 1.138[§]	0.811[–], 1.246[§]	–3.5 [–5.33, –1.67]
Kim et al. (2009)	HAD	NH	+	7.32 ± 3.2 (n = 31)	6.59 ± 2.96 (n = 29)	4.74 ± 3.33 (n = 31)	6.55 ± 3.26 (n = 29)	NA	NA	< .01[#]	0.819[–], 0.786[§]	NA	–1.81 [–3.5, –0.12]
Kim and Lee (2012)	MHS	NH	+	3.33 ± 1 (n = 10)	3.2 ± 0.84 (n = 10)	2.1 ± 0.79 (n = 10)	3.33 ± 0.77 (n = 10)	NA	NA	< .05[#]	1.449[–], 1.718[§]	NA	–1.23 [–1.92, –0.54]
Lee & Sohn (2010)	SCL-90R	H	–	2 ± 0.64 (n = 15)	1.9 ± 0.42 (n = 16)	1.7 ± 0.42 (n = 15)	1.8 ± 0.54 (n = 16)	NA	NA	.892[#]	0.366[–], 0.391[§]	NA	–0.2 [–0.54, 0.14]
Park (2010)	SCL-90R	H	–	48.44 ± 11.57 (n = 25)	?	41.64 ± 7.75 (n = 25)	?	NA	NA	.00[#]	?	NA	?
Sahai-Srivastava et al. (2014)	PHQ-9	NH	+	14.5 ± ? (n = 19)	17.5 ± ? (?)	6.75 ± ? (n = 14)	12.5 ± ? (?)	NA	NA	?	?	NA	–5.75 [?]
Song et al. (2013)	POMS	?	–	2.45 ± 1.2 (n = 24)	1.18 ± 0.64 (n = 24)	1.2 ± 0.69 (n = 24)	1.14 ± 0.65 (n = 24)	NA	NA	.000[#]	1.287[–], 1.268[§]	NA	–1.21 [–1.59, –0.83]
Koo & Kim (2013)	?	?	–	?	?	?	?	NA	NA	< .01	?	NA	?

(continued on next page)

Table 1 (continued)

Study	O^a	T^b	R^c	Pre-test		Post-test		Follow-up		p^d	$d_{corrected}^e$	$d_{corrected}$	Mean difference [95% CI]^f
				intervention	control	intervention	control	intervention	control		Pre-post	Pre-flw-up	
Lee & Young (2011)	?	?	–	?	?	?	?	NA	NA	<.05	?	NA	?
Yu & Kim (2009)	?	?	–	?	?	?	?	NA	NA	.028	?	NA	?

a Outcome.
b Therapy type: H = Humor, NH = non-humor, ? = unknown.
c + = RCT, - = non-RCT.
d # = t-test post-test; * = ANCOVA or ANOVA interaction.
e ~ = d_{ppc2}, § = d_{Korr}, & = d_{Cohen}.
f For RCT's mean diff at post-test, other designs difference between mean diff pre-post.
g Unusual results, but not technically impossible.
h and Brodaty et al. (2014)..
i and Ko and Youn (2011)..
j Two different populations tested.

plots show the results between different studies over time. One can then analyze if the outcomes show the same trend. Kim (2010) is a clear outlier (see Fig. 3). Further, there is no clear replication of the significant Cai et al. (2014) outcome, in the other two studies with non-significant outcomes (Park et al., 2013; Cho and Oh, 2011). From this plot, we cannot conclude whether laughter-inducing therapy has an effect on depression. For GDS as outcome, a reduction in depression is likely, as seven out of 10 studies have their 95% confidence interval below zero (see Fig. 4). Only four out of 10 studies have an expected reduction of the GDS outcome of four or more points (GDS outcome range is [0,30]). The other six studies show a reduction of zero to two points. Even though the reduction is significant, it might not be clinically relevant. Only a small to medium reduction in the GDS score can be expected based on this plot. There is no clear reduction for CES-D and CSDD scores (see Table 2). For the other results there are not enough similar outcomes measures to plot a comparison.

Stress. In 19 different studies, stress was an outcome measure (see Table 2). In 18 studies, stress significantly decreased due to laughter-inducing therapy. Stress was measured subjectively as perceived stress or objectively as a cortisol level. The average effect size of all studies is: $d_{dppc2} = 0.58$ ($d_{Korr} = 0.60$). The average effect size for randomized controlled trials only: $d_{dppc2} = 0.51$ ($d_{Korr} = 0.56$). We could not ascertain the difference in effect sizes between humorous and non-humorous therapies, as there were only non-humorous therapies for this outcome measure. The average effect size for non-humorous therapies only: $d_{dppc2} = 0.66$ ($d_{Korr} = 0.55$).

Next, we looked at the possible 'replication' of stress outcomes. The confidence intervals of the perceived stress measures show evidence of a reduction in stress (see Table 2). The upper bounds of four out of five studies stay below zero, but because they are different measures we cannot conclude replicability of these results. The mean differences for cortisol measurements indicate there is no evidence that laughter-inducing therapy will be beneficial for cortisol (see Fig. 5). The majority of mean differences (6/7) show a positive (statistically significant) effect, but the effects are between zero and negative two, and the upper bound of the confidence intervals cross zero in all cases. Although most studies report statistically significant reductions in cortisol levels, these differences are unlikely to be of meaningful clinical benefit.

Anxiety. In 15 different studies, anxiety was an outcome measure (see Table 3). In 14 studies, anxiety significantly decreased after laughter-inducing therapy. The average effect size for all studies is: $d_{dppc2} = 0.81$ ($d_{Korr} = 0.92$) and for randomized controlled trials: $d_{dppc2} = 0.98$ ($d_{Korr} = 1.04$). A further distinction between humorous and non-humorous therapies could be made. The average effect size for humorous therapies is: $d_{dppc2} = 0.51$ ($d_{Korr} = 0.53$) and for non-humorous therapies is: $d_{dppc2} = 1.00$ ($d_{Korr} = 1.19$). Similarly, as for the studies with depression as an outcome, the non-humorous studies have an effect size twice as large as the humorous studies.

The possible replication of anxiety outcomes is plotted in Fig. 6. The results are inconclusive. All mean differences show a positive effect, but only two upper bounds of the confidence intervals stay below zero. Again, the vote counting does not translate to the analysis of the mean differences.

3.4.2. Subgroup 2: elderly people

In 21 different studies, a positive effect of laughter or humor on the mental and physical health of the elderly was found. Most studies classified 'elderly' as 65 years or older (Bains et al., 2014; Brodaty et al., 2014; Cha and Hong, 2013; Hsieh et al., 2015; Jung et al., 2009; Kim and Lee, 2012; Ko and Youn, 2011; Lee and Eun, 2011; Lee et al., 2013; Park, 2013; Song et al., 2013; Tse et al., 2010; Walter et al., 2007), some studies as 60 years or older (George and Jacob, 2014; Hirosaki et al., 2013; Konradt et al., 2013; Lee and Sohn, 2010, Shahidi et al., 2011), one study in this subgroup studied middle-aged women between 40 and 60 years old (Cha and Hong, 2013), and two studies had unspecified samples (Lee and Young, 2011; Song et al., 2011). In general,
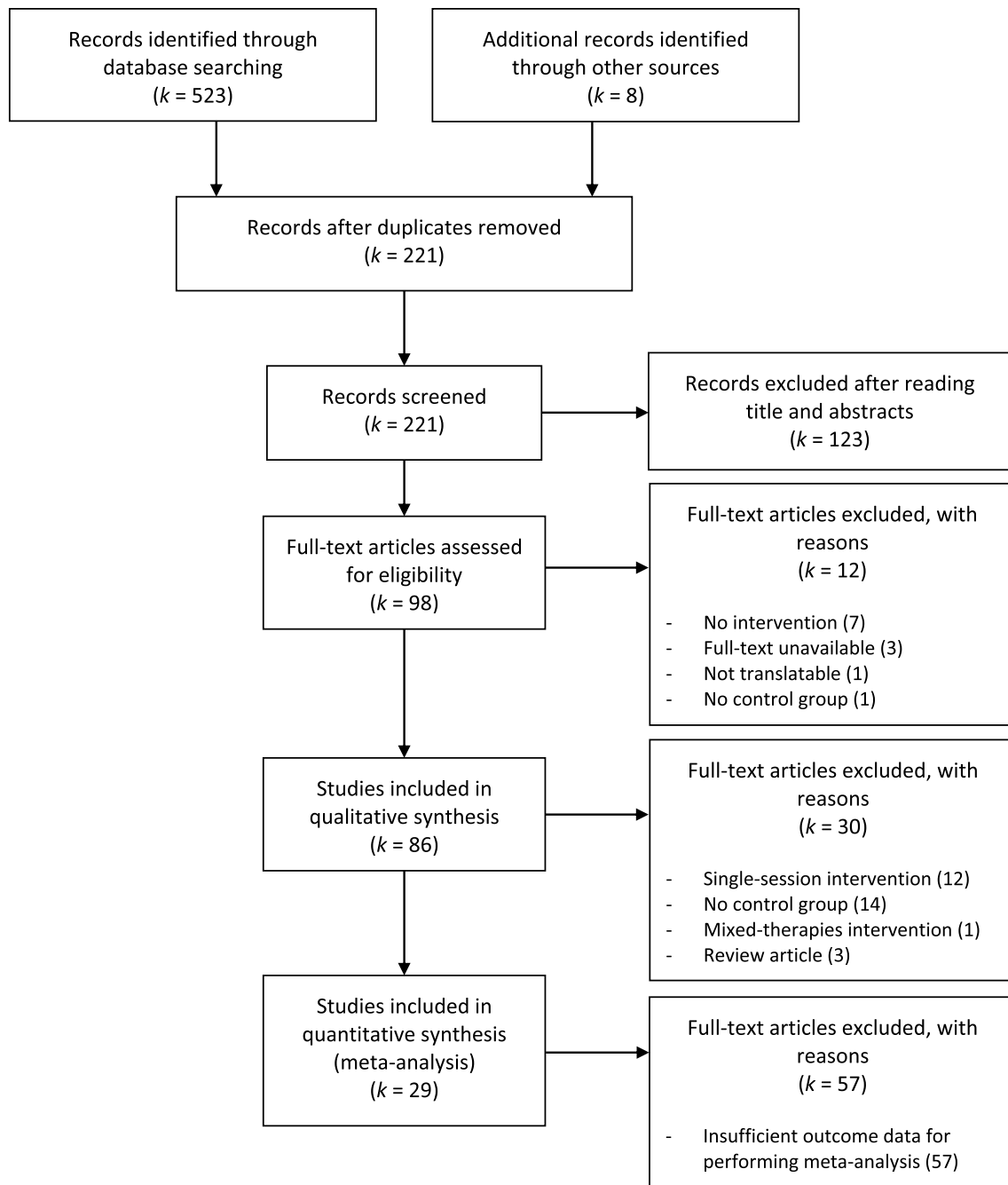
```
┌─────────────────────────────┐      ┌─────────────────────────────┐
│  Records identified through │      │ Additional records identified│
│    database searching       │      │   through other sources     │
│        (k = 523)            │      │         (k = 8)             │
└─────────────────────────────┘      └─────────────────────────────┘
              │                                    │
              └──────────────┬─────────────────────┘
                             ▼
              ┌─────────────────────────────┐
              │ Records after duplicates    │
              │         removed             │
              │        (k = 221)            │
              └─────────────────────────────┘
                             │
                             ▼
              ┌─────────────────────┐        ┌─────────────────────────────┐
              │  Records screened   │───────▶│ Records excluded after reading│
              │     (k = 221)       │        │     title and abstracts     │
              └─────────────────────┘        │        (k = 123)            │
                             │               └─────────────────────────────┘
                             ▼
              ┌─────────────────────┐        ┌─────────────────────────────┐
              │ Full-text articles  │        │ Full-text articles excluded, │
              │  assessed for       │───────▶│       with reasons          │
              │   eligibility       │        │        (k = 12)             │
              │     (k = 98)        │        │                             │
              └─────────────────────┘        │  -  No intervention (7)     │
                             │               │  -  Full-text unavailable (3)│
                             │               │  -  Not translatable (1)    │
                             │               │  -  No control group (1)    │
                             │               └─────────────────────────────┘
                             ▼
              ┌─────────────────────┐        ┌─────────────────────────────┐
              │ Studies included in │        │ Full-text articles excluded, │
              │ qualitative synthesis│──────▶│       with reasons          │
              │     (k = 86)        │        │        (k = 30)             │
              └─────────────────────┘        │                             │
                             │               │ -  Single-session intervention (12)│
                             │               │ -  No control group (14)    │
                             │               │ -  Mixed-therapies intervention (1)│
                             │               │ -  Review article (3)       │
                             │               └─────────────────────────────┘
                             ▼
              ┌─────────────────────┐        ┌─────────────────────────────┐
              │ Studies included in │        │ Full-text articles excluded, │
              │ quantitative synthesis│─────▶│       with reasons          │
              │   (meta-analysis)   │        │        (k = 57)             │
              │     (k = 29)        │        │                             │
              └─────────────────────┘        │ -  Insufficient outcome data for│
                                             │    performing meta-analysis (57)│
                                             └─────────────────────────────┘
```

Fig. 2. Flow diagram.

reported effects were: improved sleep quality, improved mood, increased life satisfaction, decreased depression and decreased pain. For most of these outcome measures, there are between one and three studies that investigated the effect. We chose to focus on the three most common outcome measures: stress, depression, and anxiety.

Depression. Eight different studies reported that laughter-inducing therapy significantly decreased depression in the elderly; two used 'spontaneous' (humorous) laughter, three used 'simulated' (non-humorous) laughter and three are unknown. (George and Jacob, 2014; Hsieh et al., 2015; Jung et al., 2009; Ko and Youn, 2011; Konradt et al., 2013; Lee and Eun, 2011; Lee et al., 2013; Lee and Young, 2011, Shahidi et al., 2011). In three studies, humor therapy did not have a significant effect on depression (Hirosaki et al., 2013; Low et al., 2013, 2014; Park, 2013). Two of these studies are the largest studies in this review (Low et al., 2013, 2014). Based on vote-counting only, most studies show that laughter-inducing therapy

improves depression in the elderly (eight out of 11 studies). Next, Table 1 shows average effect sizes for humorous versus non-humorous therapies. For humorous therapies, the average effect size is: $d_{dppc2} = 0.79$ ($d_{Korr} = 0.73$), while for non-humorous therapies it is: $d_{dppc2} = 0.53$ ($d_{Korr} = 0.52$). The positive effect of laughter-inducing therapy on depression in the elderly and the larger effect size for non-humorous therapies are in accordance with the results for all studies and populations in subgroup 1.

Stress. The effect of laughter on cortisol is mixed. One study reported decreased cortisol in the elderly, which is reported to be beneficial for memory (Bains et al., 2014) versus two studies that did not (Cha and Hong, 2013; Hsieh et al., 2015). This is in accordance with the inconclusive cortisol results over all studies in subgroup 1.

Anxiety. Only one study determined laughter-inducing therapy to have no effect on anxiety in the elderly (Low et al., 2014). No comparison could be made with other studies.

Table 2
Effect sizes and mean differences for studies with stress as outcome.

Study	O^a	T^b	R^c	Pre-test		Post-test		p^d	$d_{corrected}$ [e] Pre-Post	Mean difference [95% CI] [f]
				intervention	control	intervention	control			
Kim et al. (2009)	BEPSI	NH	+	9.45 ± 2.19 (n = 31)	8.93 ± 2.28 (n = 29)	7.77 ± 1.86 (n = 31)	9.28 ± 1.89 (n = 29)	< .01*	0.905~ / 1.038§	−0.03 [−2.47, −0.55]
Nagendra et al. (2007)	PSS	NH	+	17.79 ± 6.89 (n = 61)	19.33 ± 5.04 (n = 57)	15.80 ± 5.28 (n = 61)	18.14 ± 5.78 (n = 57)	< .05*	0.132~ / 0.17§	−2.34 [−4.37, −0.31]
Song et al. (2011, 2013)	POMS scale	?	−	2.28 ± 1.21 (n = 24)	1.43 ± 0.78 (n = 24)	1.22 ± 0.73 (n = 24)	1.45 ± 0.77 (n = 24)	< .001#	1.055~ / 1.132§	−1.08 [−1.66, −0.50]
Farifeth et al. (2014)	QSC-R23	NH	+	8.9 ± 4.4 (n = 23)	8.5 ± 3.8 (n = 14)	6.1 ± 4.2 (n = 23)	8.4 ± 4.4 (n = 14)	.03#	0.64~ / 0.633§	−2.3 [−5.02, 0.42]
Choi et al. (2010)	VAS	NH	−	4.8 ± 2.82 (n = 20)	4.83 ± 2.83 (n = 20)	2.95 ± 1.88 (n = 20)	4.35 ± 2.72 (n = 20)	.02*	1.905~ / 1.6§	−2.33 [−3.84, −0.82]
Jung and Park (2012)	VAS	?	−	71.84 ± 14.13 (n = 38)	?	50.26 ± 15.68 (n = 38)	?	.000*	?	?
Chuchuen et al. (2015)	SPST-60	?	−	? (n = 20)	? (n = 20)	? (n = 20)	? (n = 20)	> .05*	?	?
Chung (2011)	?	?	?	? (n = 20)	? (n = 20)	? (n = 20)	? (n = 20)	.019*	?	?
Lee and Ji (2011)	STAI	NH	−	?	?	2 ± 0.47 (n = 53)	?	> .05* / < .001*	?	?
Mesidor (2009)	PSS	H	−	? (n = 30)	?	? (n = 30)	?	?	?	?
Park et al. (2007)	VAS	?	+	? (n = 11)	? (n = 18)	? (n = 11)	? (n = 18)	< .05*	?	?
Bennett et al. (2003)	SACL	H	−	?	?	?	?	.008*	?	?
Jung and Park (2012)	?	?	−	? (n = 38)	?	? (n = 38)	?	?	?	?
Koo and Kim (2013)	?	?	?	?	?	?	?	< .05*	?	?
Szabo (2003)	SEES	H	−	? (n = 39)	? (n = 39)	? (n = 39)	? (n = 39)	< .001#	?	?
Oh et al. (2011)	SRS	?	?	? (n = 18)	? (n = 17)	? (n = 18)	? (n = 17)	> .05*	?	?
Kim et al. (2012)	Stress coping	?	?	? (n = 11)	? (n = 10)	? (n = 11)	? (n = 10)	?	?	?
Chang et al. (2013)	Salivary cortisol	NH	+	48.35 ± 12.53 (n = 34)	36.37 ± 14.53 (n = 33)	38.51 ± 13.13 (n = 34)	35.87 ± 14.01 (n = 33)	.058#	0.686~ / 0.689§	−0.88 [−7.47, 5.71]
Cha and Hong (2013)	cortisol	NH	−	10.1 ± 4 (n = 27)	10.3 ± 6.2 (n = 27)	8.6 ± 3.8 (n = 27)	9.6 ± 4.4 (n = 27)	.158*	0.153~ / 0.205§	−0.8 [−3.02, 1.42]
Choi et al. (2010)	Serum cortisol	NH	−	13.51 ± 9.89 (n = 20)	10.88 ± 6.94 (n = 20)	12.56 ± 9.37 (n = 20)	11.71 ± 6 (n = 20)	.498*	0.207~ / 0.2§	−1.78 [−6.72, 3.16]
Han et al. (2011)	Serum cortisol	NH	−	7.86 ± 3.46 (n = 23)	7.97 ± 3.01 (n = 17)	6.2 ± 2.75 (n = 23)	6.63 ± 2.53 (n = 17)	.613*	0.097~ / 0.128§	−0.32 [−1.99, 1.35]
Hsieh et al. (2015)	Salivary cortisol	NH	−	0.24 ± 0.19 (n = 32)	0.19 ± 0.2 (n = 30)	0.18 ± 0.16 (n = 32)	0.27 ± 0.13 (n = 30)	.009*, .003*	0.715~ / 0.874§	−0.14 [−0.21, −0.07]
Nagendra et al. (2007)	Salivary cortisol	NH	+	0.25 ± 0.14 (n = 32)	0.24 ± 0.16 (n = 15)	0.18 ± 0.11 (n = 32)	0.20 ± 0.12 (n = 15)	< .001*	0.204~ / 0.245§	−0.01 [−0.08, 0.06]
Oh et al. (2011)	Serum cortisol	?	?	? (n = 18)	? (n = 18)	? (n = 17)	? (n = 17)	< .05*	?	?
Park et al. (2007)	Serum cortisol	?	+	? (n = 11)	? (n = 18)	? (n = 11)	? (n = 18)	> .05*	?	?
Shin et al. (2011)	Breast milk cortisol	NH	−	22.85 ± 16.57 (n = 33)	22.41 ± 12.07 (n = 34)	16.14 ± 10.19 (n = 33)	19.3 ± 14.56 (n = 34)	.601*, .021*	0.247~ / 0.282§	−3.6 [−9.68, 2.48]

[a] Outcome.
[b] Therapy type: H = Humor, NH = non-humor, ? = unknown.
[c] + = RCT, - = non-RCT.
[d] # = t-test post-test; * = ANCOVA or ANOVA interaction.
[e] ~ = d_{ppc2}, § = d_{korr}.
[f] For RCT's mean diff at post-test, other designs difference between mean diff pre-post.

Table 3
Effect sizes and mean differences for studies with anxiety as outcome.

Study	O[a]	T[b]	R[c]	Pre-test intervention	Pre-test control	Post-test control	Follow-up intervention	d_corrected[e] intervention	d_corrected[e] control	p[d]	Mean difference [95% CI][f]	Pre-Post	Pre-flw-up
Cai et al. (2014)	STAI	H	+	$45.2 \pm 4.1(n=15)$	$43.8 \pm 5.1(n=15)$	$43.8 \pm 6(n=15)$	$40.6 \pm 3.6(n=15)$	NA	NA	<.005*	−2.6 [−5.80, 0.60]	$0.886{\sim}0.997$§	NA
Park et al. (2013)	STAI	NH	−	$44.47 \pm 8.06(n=17)$	$43.55 \pm 11.22(n=20)$	$44.05 \pm 10.07(n=20)$	$36.17 \pm 7.58(n=17)$	NA	NA	.002*	−7.8 [−13.98, −1.63]	$0.841{\sim}0.804$§	NA
Szabo et al. (2005)	STAI	H	−	$35.8 \pm 4.5(n=20)$	$36.3 \pm 9.9(n=20)$	$38.2 \pm 9.5(n=20)$	$33 \pm 5(n=20)$	NA	NA	<.02*	−0.9 [−5.82, 4.02]	$0.107{\sim}0.085$§	NA
You and Choi (2012)	STAI	NH	−	$56.53 \pm 15.88(n=17)$	$59.13 \pm 8.87(n=16)$	$60.19 \pm 11.38(n=16)$	$41.41 \pm 9.1(n=17)$	NA	NA	.001*	−14.06 [−23.56, −4.56]	$1.004{\sim}1.708$§	NA
Jung and Park (2012)	STAI	?	−	$46.92 \pm 8.96(n=38)$	?(n=38)	?(n=38)	$39.66 \pm 8.02(n=38)$	NA	NA	.000#	?	?	NA
Han et al. (2011)	STAI	NH	−	$43.47 \pm 11.99(n=23)$	$44.53 \pm 0.12(n=17)$	$48.82 \pm 10.03(n=17)$	$34.39 \pm 12.63(n=23)$	NA	NA	.006#	−4.79 [−10.02, 0.44]	$0.424{\sim}0.578$§	NA
Chung (2011)	STAI	?	−	?(n=20)	?(n=20)	?(n=20)	?(n=20)	NA	NA	.046#	?	?	NA
Szabo (2003)	STAI	H	−	?(n=39)	?(n=39)	?(n=39)	?(n=39)	NA	NA	<.01*	?	?	NA
Yu and Kim (2009)	STAI	?	−	?	?	?	?	NA	NA	.046#	?	?	NA
Yazdani et al. (2014)	GHQ	NH	+	$7.1 \pm 5.3(n=19)$	$6.8 \pm 4.5(n=19)$	$6.1 \pm 4.3(n=19)$	$3.7 \pm 3.19(n=19)$	$4.1 \pm 4.1(n=19)$	$6.9 \pm 5.4(n=19)$	.04#, .02#	−4.1 [−6.61, −1.59]	?	$0.782{\sim}0.791$§
Kim et al. (2009)	HAD	NH	+	$7.9 \pm 3.27(n=31)$	$6.83 \pm 3.07(n=29)$	$7.24 \pm 2.81(n=29)$	$6.16 \pm 2.63(n=31)$	NA	NA	.04#	−0.67 [−2.14, 0.80]	$0.433{\sim}0.451$§	NA
Kim and Lee (2012)	MHS	NH	−	$3.13 \pm 0.59(n=10)$	$3.14 \pm 0.54(n=10)$	$3.14 \pm 0.6(n=10)$	$1.63 \pm 0.61(n=10)$	NA	NA	<.001#	−1.51 [−2.02, −1]	$2.481{\sim}2.604$§	NA
Lee and Sohn (2010)	SCL-90	H	−	$1.3 \pm 0.45(n=15)$	$1.4 \pm 0.34(n=16)$	$1.5 \pm 0.39(n=16)$	$1.6 \pm 0.59(n=15)$	NA	NA	.475#	0.4 [0.05, 0.75]	$0.943{\sim}0.895$§	NA
Low et al. (2013)	BEAM	H	+	$43.6 \pm 93.8(n=209)$	$32.5 \pm 70.2(n=189)$	$35.2 \pm 92(n=189)$	$30.1 \pm 73(n=209)$	$14.2 \pm 45.8(n=209)$	$16.7 \pm 56.8(n=209)$	<.001#, .414*	−2.4 [−16.65, 11.85]	$0.116{\sim}0.124$§	$0.117{\sim}0.139$§
Park (2010)	SCL-90	H	−	$48.8 \pm 10.33(n=25)$	NA	NA	$43.2 \pm 9.07(n=25)$	NA	NA	.000#	?	?	NA

[a] Outcome.

[b] Therapy type: H = Humor, NH = non-humor, ? = unknown.

[c] + = RCT, - = non-RCT.

[d] # = t-test post-test; * = ANCOVA or ANOVA interaction.

[e] ~ = d_{ppc2}, § = d_{Korr}.

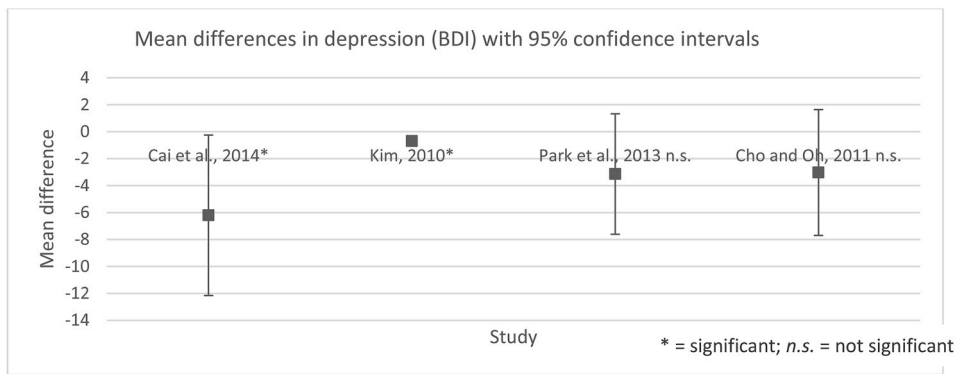[f] For RCTs mean diff at post-test, other designs difference between mean diff pre-post.

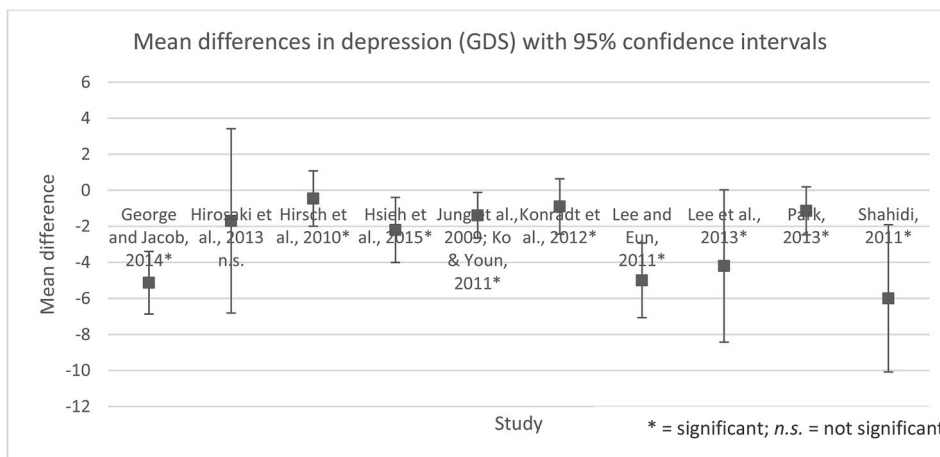**Fig. 3.** Mean differences in depression (BDI) with 95% confidence intervals.

Fig. 4. Mean differences in depression (GDS) with 95% confidence intervals.

3.4.3. Subgroup 3: other important populations and outcome measures

Cancer patients. Eight studies reported that laughter therapy had a significant positive effect on the mental and physical health of cancer patients. More specifically, laughter therapy decreased perceived stress and improved mood (Choi et al., 2010; Farifeth et al., 2014; Kim et al., 2009, 2015). Laughter therapy also decreased anxiety (You and Choi, 2012; Han et al., 2011; Kim et al., 2009). Mixed results were found for pain and immunological response. Laughter therapy decreased pain in one study (You and Choi, 2012), but not in another one (Choi et al., 2010). The immunological response was improved in one study (Sakai

et al., 2013), but no effect on immunity was found in Cho and Oh (2011). Improvement of mood and perceived stress in cancer patients seems convincing; with three out of four studies being a RCT and half of the criteria show mostly a low risk of bias versus an unclear risk of bias in the other half. For pain and immunological response, no clear indication can be given at this point.

Healthy adults. Thirteen studies reported a positive effect of laughter inducing-therapy on mental, social or physical health in (healthy) adults. Seven of these studies used 'spontaneous' laughter (Bennett et al., 2003; Berk et al., 2014; Buchowski et al., 2007; Lowis, 1997; Lee
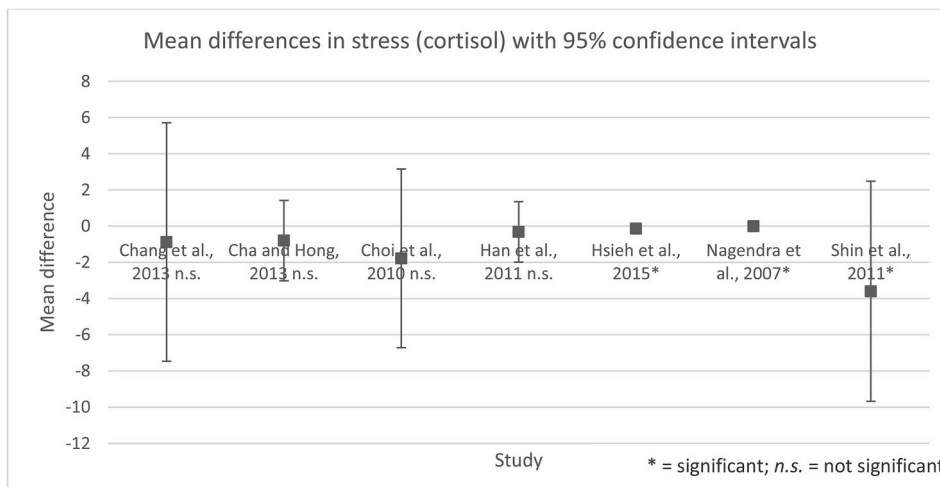

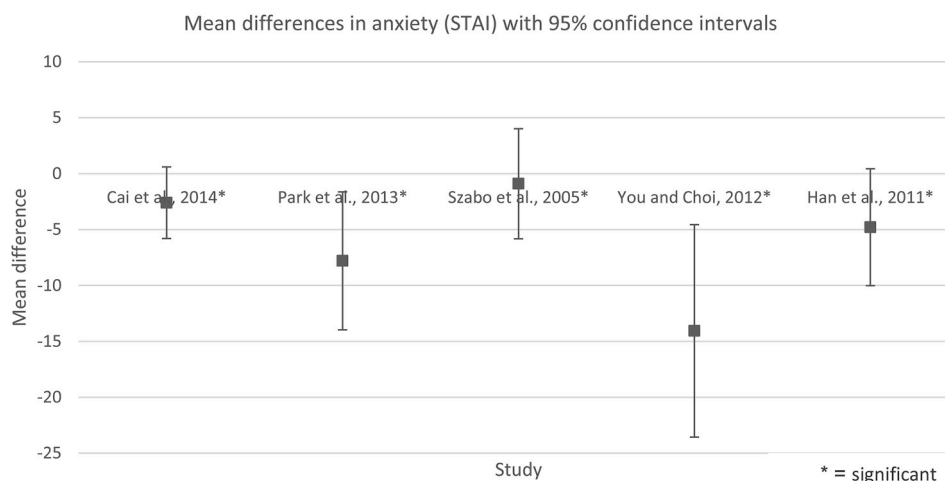**Fig. 5.** Mean differences in stress (cortisol) with 95% confidence intervals.

Fig. 6. Mean differences in anxiety (STAI) with 95% confidence intervals.

and Ji, 2011; Park, 2010; Szabo, 2003; Szabo et al., 2005), three studies used 'simulated' laughter (Nagendra et al., 2007; Wagner et al., 2014; Yazdani et al., 2014) and two studies were unclear (Jung and Park, 2012; Oh et al., 2011). Findings report that laughter or humor can reduce stress and increase "Natural Killer cell activity" (Bennett et al., 2003), improve mental health (Park, 2010; Yazdani et al., 2014), increase coping humor (Lowis, 1997), have EEG correlates in beta and gamma bands (Berk et al., 2014), increase energy expenditure, activate abdominal and back-lifting muscles (Buchowski et al., 2007; Wagner et al., 2014) and improve mood and reduce anxiety or reduce stress (Szabo, 2003; Szabo et al., 2005; Jung and Park, 2012; Lee and Ji, 2011; Nagendra et al., 2007). Although Oh et al. (2011) found laughter therapy can decrease serum cortisol, there was no effect on coping and stress response in their study. The positive findings on stress reduction and energy expenditure can be convincing, given they come from lab studies and have low risk of bias; however, these studies used small sample sizes. The other results are hampered by medium to high risk of bias, mainly coming from non-randomization. Three studies determined the effect of humor on employees and found it can increase self-efficacy in employees (Beckman et al., 2007), increase enjoyment at work and reduces employees' stress levels (Chenoweth et al., 2014) or increase positive mood and reduce blood pressure (Nagendra et al., 2007). These results are not very convincing because they either lack a control group (Beckman et al., 2007), had a high drop-out rate (Chenoweth et al., 2014) or found no significant changes in stress with a stronger study design (Nagendra et al., 2007).

Children and teenagers. Four studies reported that laughter significantly improved the social health of children and teenagers. More specifically, laughter improved social support and life satisfaction (Kim et al., 2012), improved self-efficacy and social competence (Koo, 2010), reduced stress and depression (Koo and Kim, 2013), and self-esteem and coping skills (Choi and Cho, 2011) of children and teenagers. For all four studies it is unknown what type of laughter was used. These results should be interpreted with caution due to the unclear risk of bias in three studies and the small sample size and non-randomization in the other study.

Physical health: pain. Nine studies determined the effect of laughter-inducing therapy on pain. (Choi et al., 2010; Herschenhorn, 1995; Kessler et al., 2010, 2012; Kim et al., 2010; Ko and Youn, 2011; Lee and Eun, 2011; Tse et al., 2010; You and Choi, 2012; Yu and Kim, 2009). Six out of nine studies reported a decrease of pain after laughter-inducing therapy. Moreover, studies reported that laughter can improve fatigue in patients with arthrosis (Kim et al., 2010) or decrease anxiety, depression, pulse rate and blood pressure in military patients with low back pain (Yu and Kim, 2009). One study did not find significant effects on quality of life, headache, fatigue, or general health in females with

chronic migraines (Sahai-Srivastava et al., 2014). The relationship between laughter intensity and pain in rheumatoid women has also been investigated, but no results are reported in the article (Herschenhorn, 1995). Again, most of these findings are unconvincing due to the high risk of bias or small sample size.

Female health and wellbeing: postpartum stress & infertility. Three studies reported that laughter therapy can decreases infertility stress and anxiety and increases laughter (Jung and Park, 2012), improve the immunoresponse in postpartum women (Ryu et al., 2015) and decrease fatigue and serum cortisol levels in breastfeeding postpartum women (Shin et al., 2011). These studies are at high or unclear risk of bias and are therefore not very convincing.

3.5. Results summary systematic review

In conclusion, there is a reasonably convincing trend indicating 'simulated' (non-humorous) laughter having a more positive effect on depression and anxiety compared to 'spontaneous' (humorous) laughter. This is based on effect sizes twice as large for 'simulated' laughter than 'spontaneous' laughter-inducing therapies. Both for subgroup 1 (all populations), as well as in subgroup 2 (the elderly). Furthermore, laughter-inducing therapies seem to improve depression, and perceived stress. This is based on the replication of a decrease in depression (measured with GDS) and perceived stress (measured with differing instruments). The majority of the lower bounds of these confidence intervals stay below zero, indicating a positive effect, though this may not be clinically relevant. For cortisol levels, the results remain inconclusive. Finally, laughter-inducing therapy also seems to improve mood, perceived stress and depression in cancer patients, infertile women, adults and children. This conclusion should be interpreted with caution, as the systematic review included a high number of low quality studies (no control, small N or no results in the article due to conference abstract) and studies with a high risk of bias. Therefore, although a summary of these studies suggests an overall positive effect of these therapies, the methodological shortcomings of these studies make it hard to critically interpret the positive results. As most of the included studies in the systematic review used nonrandomized study designs, had very small sample sizes or were otherwise at high risk of bias, many of these positive results could be spurious findings or other statistical artefacts such as regression to the mean or nonspecific factors.

As a practical significance, laughter-inducing therapies seem to be applicable in a wide range of settings and for many different populations, including severe or terminally ill, disabled, healthy, employees, elderly, adults and children. Moreover, there are almost no contraindications, which makes it safe for many people to try. It seems to be feasible for practitioners to teach laughter-inducing therapies after

Table 4
Meta-analysis results.

Study	N	g	95% CI of g	p	Q (df)	I^2	T^2	$p_{between}$
All studies								
Quasi-experimental studies	20	0.72	0.49, 0.96	< .001	86 (19)	78	0.28	
Randomized controlled trials	9	0.50	0.15, 0.85	< .01	22 (8)	63	0.08	
								0.30
Type of laughter therapy								
Randomized controlled trials								
No humor	5	0.62	0.17, 1.04	< 0.01	18 (4)	77	0.24	
Humor	3	0.33	−0.17, 0.83	0.40	2 (2)	0	0.00	
								0.46
Quasi-experimental studies								
No humor	18	0.83	0.56, 1.09	< 0.001	69 (17)	75	0.27	
Humor	2	0.05	−0.68, 0.79	0.89	0 (1)	0	0.00	
								0.05
Mental health outcomes								
Randomized controlled trials								
All depression outcomes	7	0.48	0.21, 0.75	< 0.001	14 (6)	55	0.06	
Validated depression outcomes[a]	5	0.65	0.27, 1.04	0.001	7 (4)	43	0.08	
All anxiety outcomes	3	1.05	−0.10, 2.21	0.07	17 (2)	89	0.89	
Quasi-experimental studies								
All depression outcomes	15	0.74	0.43, 1.06	< 0.001	66 (14)	79	0.29	
Validated depression outcomes[b]	15	0.75	0.45, 1.06	< 0.001	62 (14)	78	0.27	
All anxiety outcomes	5	0.59	−0.11, 1.29	< 0.001	23 (4)	83	0.52	

[a] BDI, GDS, MHS-depression.
[b] BDI, BDI-II, CES-D, GDS, HADS-depression, MHS-depression, POMS-depression, SCL-90-depression.

following a laughter teacher training and when experienced in giving group therapy or other type of group sessions (e.g. improvisation, singing, dancing, sports, yoga, etc.) as the group dynamics seem to be important to stimulate and manage.

3.6. Meta-analysis

In total, 29 studies (Flow diagram 2) reported sufficient information, such as outcome statistics and number of participants, to be eligible for the meta-analysis, with a total number of $n = 1986$ participants ($n = 976$ in the intervention group and 1010 in the control group; on average $n = 68$ participants per study, range $n = [20, 398]$). Of these, a total of $n = 894$ participants were available from randomized controlled trial designs ($n = 430$ in the intervention group and $n = 464$ in the control group).

Of all mental health outcomes, depression, anxiety, and stress were reported most often in the included studies. Depression outcomes included both validated and commonly accepted measures of depression (e.g., BDI and GDS), but also ad-hoc self-report measures. We therefore report depressions outcomes both at a pooled level, and as a subgroup present results only from validated measures (see Table 4). For anxiety outcomes, too few validated measures were used to make this distinction. Too few studies - randomized or quasi-experimental - were

available to pool effect sizes for stress-specific or quality of life outcomes.

Results are presented for depression and anxiety outcomes for all types of laughter-inducing therapies. Randomized versus quasi-experimental studies will be compared and integrated with the results found in the systematic review (qualitative synthesis). See Figs. 7 and 8 for forest plots of these results and Table 4 for detailed quantitative results.

3.6.1. Depression and anxiety outcomes for all types of laughter-inducing therapy

Randomized controlled trials. Seven comparisons from randomized controlled trials were available for combined depression outcomes. The mean depression effect size for randomized studies was $g = 0.48$ (95% $CI = 0.21, 0.75; I^2 = 55, T^2 = 0.06$), which increased slightly when considering only validated measures ($g = 0.65$, 95% $CI = 0.27, 0.04$). Heterogeneity in this subgroup analysis was medium ($I^2 = 43\%$, $T^2 = 0.08$, see Fig. 7). A separate analysis for all anxiety outcomes, had only 3 comparisons available, yielding $g = 1.05$ (95% $CI = -0.10$, 2.21) with high heterogeneity ($I^2 = 89\%$, $T^2 = 0.89$, see Table 4).

Quasi-experimental studies. For quasi-experimental studies 15 comparisons were available, yielding a combined depression effect size of $g = 0.74$ (95% $CI = 0.43, 1.06$). As these studies included mostly validated measure, this result was relatively robust when we calculated
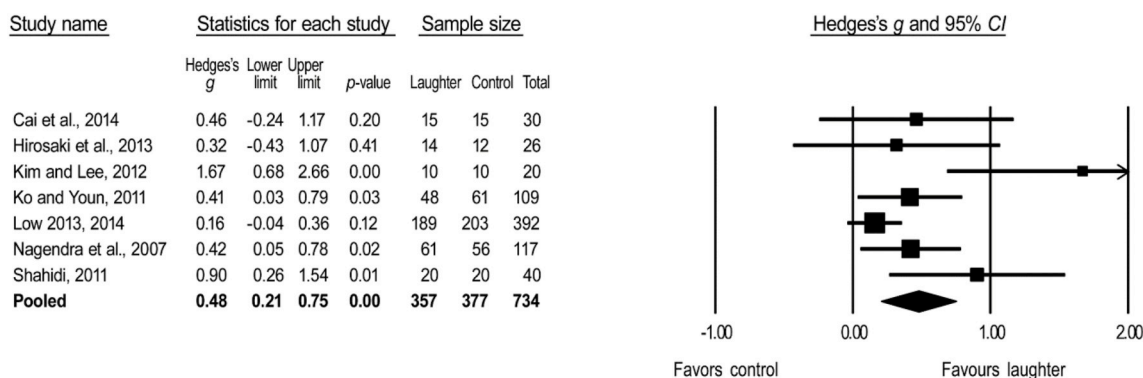

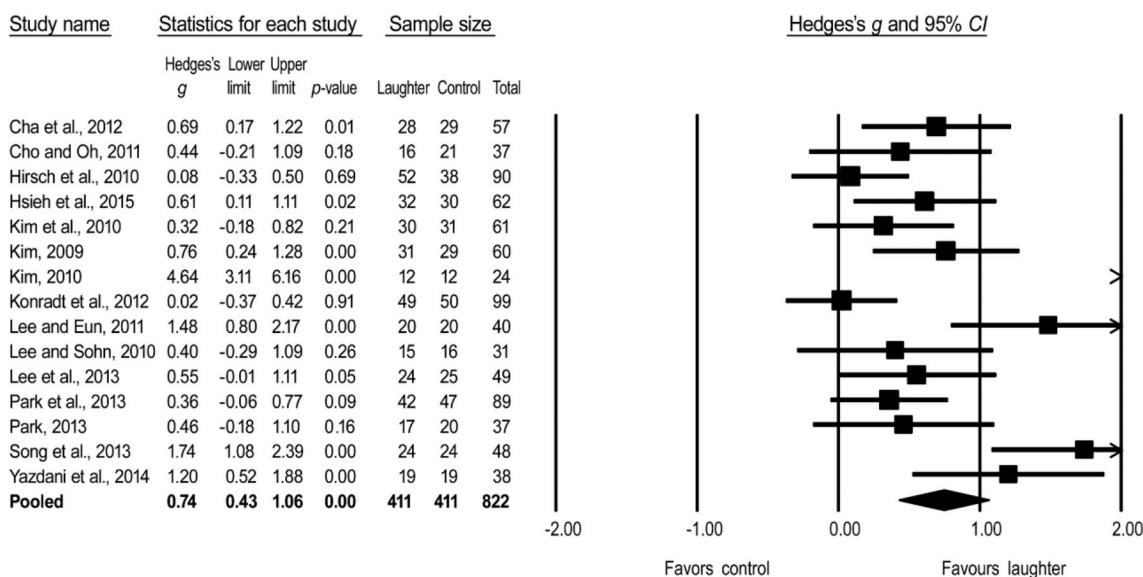**Fig. 7.** Forest plot of RCT studies with depression as outcome.

Study name	Statistics for each study				Sample size			Hedges's g and 95% CI
	Hedges's g	Lower limit	Upper limit	p-value	Laughter	Control	Total	
Cha et al., 2012	0.69	0.17	1.22	0.01	28	29	57	
Cho and Oh, 2011	0.44	-0.21	1.09	0.18	16	21	37	
Hirsch et al., 2010	0.08	-0.33	0.50	0.69	52	38	90	
Hsieh et al., 2015	0.61	0.11	1.11	0.02	32	30	62	
Kim et al., 2010	0.32	-0.18	0.82	0.21	30	31	61	
Kim, 2009	0.76	0.24	1.28	0.00	31	29	60	
Kim, 2010	4.64	3.11	6.16	0.00	12	12	24	
Konradt et al., 2012	0.02	-0.37	0.42	0.91	49	50	99	
Lee and Eun, 2011	1.48	0.80	2.17	0.00	20	20	40	
Lee and Sohn, 2010	0.40	-0.29	1.09	0.26	15	16	31	
Lee et al., 2013	0.55	-0.01	1.11	0.05	24	25	49	
Park et al., 2013	0.36	-0.06	0.77	0.09	42	47	89	
Park, 2013	0.46	-0.18	1.10	0.16	17	20	37	
Song et al., 2013	1.74	1.08	2.39	0.00	24	24	48	
Yazdani et al., 2014	1.20	0.52	1.88	0.00	19	19	38	
Pooled	**0.74**	**0.43**	**1.06**	**0.00**	**411**	**411**	**822**	

Favors control · Favours laughter

Fig. 8. Forest plot of quasi-experimental studies with depression as outcome.

effect sizes for depression as for validated outcomes $g = 0.75$ (95% $CI = 0.45, 1.06$) Heterogeneity was high in quasi-experimental studies ($I^2 = 79\%$, $T^2 = 0.30$, see Fig. 8). For all anxiety outcomes, 5 comparisons were available, yielding $g = 0.59$ (95% $CI = -0.11, 1.29$), again with high heterogeneity ($I^2 = 83\%$, $T^2 = 0.52$, see Table 4).

3.7. Results summary meta-analysis

The results show that overall, laughter interventions show medium to large between-group effect sizes for depression and anxiety outcomes, which were the only outcomes that were reported often enough to be pooled meaningfully. The results are similar to those in the systematic review. Although the effect sizes are comparable to, e.g., those found for various psychotherapeutic interventions for depression - see e.g. (Barth et al., 2013) - this does not mean that laughter therapies are equally as effective as established therapies as these effect sizes cannot be compared directly.

3.8. Publication bias

Publication bias was assessed using Duval and Tweedie's trim and fill procedure. The procedure indicated considerable publication bias, imputing 10 missing studies. This lowered the pooled effect size from $g = 0.47$ (95% $CI = 0.13, 0.81$) to $g = 0.28$ (95% $CI = 0.08, 0.53$). However, in the presence of high heterogeneity as in this sample, the trim-and-fill procedure may give unreliable results (Peters et al., 2007). Additionally, both Egger's test (intercept = 3.67, $p < 0.0001$) and Begg and Mazumdar's test (Kendall's Tau = 0.51, $p < 0.0001$) were significant, also indicating publication bias. Therefore, the results of publication bias analysis must be interpreted with caution as they warrant no conclusion to the (non)existence of publication bias.

4. Discussion

4.1. Strengths and limitations

To our knowledge, this review is the first to systematically retrieve and assess literature on laughter interventions. Systematic reviews and meta-analysis are by necessity limited by the quality of the included primary studies. We tried to overcome limitations by including quantitative measures in the systematic review to prevent categorizing results as either statistically significant or not. Vote-counting could show

a positive image, while taking sample sizes, effect sizes and mean differences into account can show a different picture and provide a meaningful context with information on clinical relevance. Therefore we included quantitative measures when reported in the article.

In this review, limitations are the low quality of studies and high risk of bias in the included studies. Many studies in the meta-analysis had a very low sample size (average $n = 68$) with one distinctive outlier, a multi-site randomized controlled trial (Low et al., 2013; Low et al., 2014, $n = 398$). Twelve studies had 20 or fewer participants per condition, and they could perhaps more accurately be described as pilot studies unsuitable for reliable effect size estimations (Kraemer et al., 2006). This limits the conclusions drawn from the systematic review, where generally favorable results were found; and in the meta-analytic results, this is reflected in the large confidence intervals around the estimated pooled effect size and the high heterogeneity. Moreover, the results of the systematic review rely for the largest part on non-randomized studies. This means that the results of these studies – mostly statistically significant – could be the result of nonspecific treatment effects, spurious findings, regression to the mean or non-corrected multiple testing, suboptimal analytic strategies, and all the other inherent limitations of nonrandomized studies.

None of the included trials were pre-registered in a publicly accessible trial registry, as required by the declaration of Helsinki and recommended by the ICMJE guidelines, which made it impossible to check if outcome measures were omitted or switched. Some studies had mixed therapies, which make it difficult to conclude whether the results found are attributable to the laughter or humor part of the therapy, or to nonspecific treatment effects. Dismantling studies could provide more insight into this. Moreover, a number of outliers in the meta-analysis reported improbably high effect sizes, which could have resulted in the distortion of the effect size estimates (e.g., Kim, 2010; $g = 4.19$; however, the very small sample size ($n = 24$) attenuates its effects on the analyses). Conversely, the study of Low et al. (2013) in the meta-analysis could be considered a negative outlier, as its effect size was very low but its sample size very high. Almost all pooled outcomes showed high heterogeneity, even in subgroup analyses. This can be explained by a large variability between studies, especially the 'simulated' (non-humorous) laughter studies. There were many non-English journal articles of studies performed in Asian countries, which could lead to a language bias in the results, although much data could be extracted from non-English papers using Google Translate. We would like to suggest to make English reports available as well. Finally,

even though great care was put into search terms and covering the field of laughter and humor therapies, we decided to include the databases most commonly used for meta-analysis research of health outcomes: WOS, EBSCO, EMBASE, and PUBMED. Additionally, there is lack of shared vocabulary among researchers, which leads to a wide range of phrases, terms, and definitions used; in turn this suggests a research field that is yet to reach maturity.

4.2. Practical implications

Laughter-inducing therapies could be a valuable complementary or in some cases even main therapy in different settings, but the lack of high-quality studies currently preclude recommending its use in clinical practice on anything other than a complementary, patient-preference basis. The results of the systematic review suggest that laughter therapies are acceptable in a wide range of settings, and for a broad range of patient groups; this is a good starting point for further research. From the included studies it can be concluded that it requires a trainer that can give the therapy in group sessions. The participants and staff can practice laughter on their own as well, between sessions. From multiple studies, it also seems acceptable to use in serious or terminal conditions, such as cancer patients waiting for chemotherapy or terminally ill patients (Farifeth et al., 2014; Kessler et al., 2010, 2012; Kim et al., 2015). Laughter-inducing therapy can also be applied in populations with reduced mobility as it can be done lying or sitting (watching humorous videos or doing laughter exercises). Especially non-humorous laughter could be used well in elderly or cognitively impaired populations, as this type does not rely on, e.g., verbal skills such as wordplay.

Laughter has shown to produce effects on many levels: emotional, psychological, behavioral, and biological. Laughter-inducing therapies show promise as an addition to main therapies or medication, and should be investigated further, especially as an adjunct therapy for somatic diseases with a psychological component such as depression or anxiety. Furthermore, to our knowledge, there are few imaginable contraindications, which makes it universally safe to practice. The social aspect of laughing together seems to be an advantage to alleviate societal problems such as loneliness, bullying at schools, or aggression in general. We can imagine the intensity and therefore effectiveness of laughter during sessions might depend on the experience of the teacher or therapist, and group dynamics. However, more research is needed to establish the importance of a possible 'therapeutic alliance' as is commonly measured in psychotherapeutic interventions.

We encourage practitioners to experiment with and test laughter-inducing therapies, and suggest co-creation activities to help develop and protocolize these therapies. For example, in elderly care homes, there could be a discussion with the residents to assess their preferences. Our results show that humor is not necessary, and that 'laughing about nothing' seems to work as well, so taste or sense of humor is not relevant, although some might initially feel embarrassed laughing about 'nothing' in a group.

In the context of other rising health care costs and the increasing elderly population, there is a potential for low-cost, simple interventions that can be administered by staff with minimal training. Unfortunately, the seemingly great potential of laughter-inducing therapies has not yet materialized, perhaps partly due to a lack of evidence of effectiveness. Should effectiveness be shown, how can laughter be implemented in healthcare? We recommend 'simulated' (non-humorous) laughter instead of 'spontaneous' (humorous) laughter, as the systematic review shows a more consistent effect and a higher effect size for 'simulated' (non-humorous) laughter. Furthermore, 'simulated' laughter does not need any cognition (e.g. no necessity of understanding a joke or funny story) and could be performed seated or in a hospital bed. Laughter-inducing therapies can be performed in group and private settings, preferably with a teacher experienced in group dynamics and laughter exercises.

4.3. Theoretical and research implications

Although the evidence for laughter-inducing therapies is growing, methodological rigor is still lacking and the quality of evidence is either low or very low. Future research should focus on adequately powered, pre-registered randomized controlled interventions with large samples, longer therapy durations (e.g. 10 or more sessions over more than 5 weeks) and with follow-up measurements. Also, an intervention where the level or 'dosage' of laughter can be measured should be created. Some efforts in this direction have been made, for example, Mora-Ripoll (2011) suggests the diaphragm electromyogram as an exact measurement of laughter. We also see possibilities for a web- or smartphone-based intervention where laughter is measured with facial recognition and voice analysis using integrated smartphone sensors. Great advances have been made recently in automatically recognizing (genuine) laughter through social signals processing and face and voice recognition (Dibeklioğlu et al., 2015; Dupont et al., 2016), and this could provide a useful measure of whether the posited positive effects of the therapies are indeed mediated by laughter itself.

We propose that diverse laughter and humor trainers come together to find a consensus on what guidelines for laughter-inducing therapies should be created, and whether a protocol for a 'standard' laughter therapy treatment can be established. Such a protocol should include information on which exercises should be included in training, how many sessions are needed to derive benefit, how long these sessions should be, and other basic information. Furthermore, it is unclear what the necessary qualification for teachers or therapists would be, and what the 'minimal' dose per session should be. Standardization of treatments would make it easier to independently reproduce results and to systematically research and improve the therapies. This can then be validated in large-scale randomized controlled trials.

5. Conclusions

The aim was to systematically estimate the efficacy or effectiveness of laughter-inducing therapies. This was done by conducting a systematic review and meta-analysis on the effects of 'spontaneous' (humorous) and 'simulated' (non-humorous) laughter-inducing therapies on physical and mental health outcomes. Firstly, the systematic review found there was a reasonably convincing trend indicating 'simulated' (non-humorous) laughter has a more positive effect on depression and anxiety compared to 'spontaneous' (humorous) laughter. This was based on effect sizes twice as large for non-humorous therapies than humorous therapies. Secondly, in the systematic review as well as the meta-analysis, a similar pattern was found for the laughter-inducing studies and their effect on depression. Both found a positive effect on depression, with similar medium effect sizes ranging between: Hedges' g [0.48, 0.65] or d_{ppc2} [0.51, 0.58]. The systematic review also found a convincing replication of depression outcomes (measured with the GDS). Furthermore, laughter-inducing therapies seem to improve perceived stress. The systematic review showed that cortisol levels and pain could potentially be reduced by laughter-inducing therapy, but results remain inconclusive. Although a summary of these studies suggests an overall positive effect of these therapies, the methodological shortcomings of these studies make it hard to critically interpret the summary positive results. The systematic review included a high number of low-quality studies (no control, small n or no results in conference abstracts), and studies with a high risk of bias.

Future directions for this research field are: performing randomized, pre-registered, controlled trials of some standardized form of laughter therapy with sufficient sample sizes; ideally testing the cost-effectiveness of laughter-inducing therapies in clinical settings with trained staff, developing interventions where the precise 'dosage' of laughter can be measured and managed and investigating the most effective (and minimal) number and recurrence of therapy sessions.

Author contributions

Study conception and design: NW. Acquisition of data: NW and RK. Analysis and interpretation of qualitative data: NW. Analysis and interpretation of quantitative data: NW and RK. Drafting of manuscript: NW and RK. Critical revision: NW and RK.

Funding

Acknowledgements

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.socscimed.2019.02.018.

References

Bains, G., Berk, L., Daher, N., Lohman, E., Petrofsky, J., Schwab, E., Deshpande, P., 2014. Effectiveness of humor on short-term memory function and cortisol levels in age matched elderly and diabetic subjects vs. control group (684.4). FASEB J. 28 (1 Supplement) 684-4.

Barth, J., Munder, T., Gerger, H., Nüesch, E., Trelle, S., Znoj, H., et al., 2013. Comparative efficacy of seven psychotherapeutic interventions for patients with depression: a network meta-analysis. PLoS Med. 10 (5), e1001454. https://doi.org/10.1371/journal.pmed.1001454.

Beckman, H., Regier, N., Young, J., 2007. Effect of workplace laughter groups on personal efficacy beliefs. J. Prim. Prev. 28 (2), 167–182.

Begg, C.B., Mazumdar, M., 1994. Operating Characteristics of a Rank Correlation Test for Publication Bias. Biometrics, pp. 1088–1101.

Bennett, M.P., Zeller, J.M., Rosenberg, L., McCann, J., 2003. The effect of mirthful laughter on stress and natural killer cell activity. Altern. Ther. Health Med. 9 (2), 38.

Bennett, M.P., Lengacher, C.A., 2006. Humor and laughter may influence health. I. History and background. Evid. Based Complement Altern. Med. 3 (1), 61–63.

Bennett, M.P., Lengacher, C., 2009. Humor and laughter may influence health IV. humor and immune function. Evid. Based Complement Altern. Med. 6 (2), 159–164.

Bennett, P.N., Parsons, T., Ben-Moshe, R., Weinberg, M., Neal, M., Gilbert, K., et al., 2014. Laughter and humor therapy in dialysis. Semin. Dial. 27 (5), 488–493. https://doi.org/10.1111/sdi.12194.

Berk, L., Alphonso, C., Thakker, N., Nelson, B., 2014. Humor similar to meditation enhances EEG power spectral density of gamma wave band activity (31-40Hz) and synchrony. FASEB (Fed. Am. Soc. Exp. Biol.) J. 28 (1).

Black, D.W., 1982. Pathological laughter: a review of the literature. J. Nerv. Ment. Dis. 170 (2), 67–71.

Borenstein, M., Higgins, J., Hedges, L.V., Rothstein, H.R., 2017. Basics of meta-analysis: I2 is not an absolute measure of heterogeneity. Res. Synth. Methods 8 (1), 5–18.

Brodaty, H., Low, L.F., Liu, Z., Fletcher, J., Roast, J., Goodenough, B., Chenoweth, L., 2014. Successful ingredients in the SMILE study: resident, staff, and management factors influence the effects of humor therapy in residential aged care. Am. J. Geriatr. Psychiatry 22 (12), 1427–1437.

Buchowski, M., Majchrzak, K., Bloomqist, K., Byrne, D., Chen, K., Bachorowski, J., 2007. Energy expenditure of genuine laughter. Obes. Res. 12 A134-A134.

Cai, C., Yu, L., Rong, L., Zhong, H., 2014. Effectiveness of humor intervention for patients with schizophrenia: a randomized controlled trial. J. Psychiatr. Res. https://doi.org/10.1016/j.jpsychires.2014.09.010.

Cha, M.Y., Na, Y.K., Hong, H.S., 2012. An effect of optimism, self-esteem and depression on laughter therapy of menopausal women. Korean Journal of Women Health Nursing 18 (4).

Cha, M.Y., Hong, H.S., 2013. Effects on the laughter score, cortisol and immunoglobulin of laughter therapy in middle aged women. Journal of Korean Biological Nursing Science 15 (4), 230–236.

Chang, C., Tsai, G., Hsieh, C.-J., 2013. Psychological, immunological and physiological effects of a Laughing Qigong Program (LQP) on adolescents. Complement. Ther. Med. 21 (6), 660–668. https://doi.org/10.1016/j.ctim.2013.09.004.

Chenoweth, L., Low, L.F., Goodenough, B., Liu, Z., Brodaty, H., Casey, A.N., et al., 2014. Potential benefits to staff from humor therapy with nursing home residents. J. Gerontol. Nurs. 40 (2), 47–52.

Cho, E.A., Oh, H.E., 2011. Effects of laughter therapy on depression, quality of life, resilience and immune responses in breast cancer survivors. Journal of Korean Academy of Nursing 41 (3), 285–293.

Choi, J., Kim, K., Cha, S., Pyo, H., Kim, Y., 2010. Effects of laughter therapy on mood, pain, and stress of mastectomy patients. Journal of Korean Clinical Nursing Research 16, 83–93.

Choi, S.O., Cho, S.Y., 2011. The development and effect analysis of a laughter therapy program to, improve the self-esteem and stress-coping skills of children using a community children's center. Journal of Family Welfare, The Korean Association of Family Welfare 16 (16), 97–117.

Chuchuen, U., Pampiansil, P., Busarakumtragul, P., 2015. Effects of laughing training on stress levels in Thai private office workers. J. Med. Assoc. Thail. 98, S130–S134.

Chung, H.J., 2011. The effect of laughter therapy on infertility stress and anxiety of women receiving in vitro fertilization. Fertil. Steril. 96 (3), S195–S196.

Cokolic, M., Stangler Herodez, S., Sternad, S., Krebs, S., 2013. The inhibitory effect of laughter yoga on the increase in postprandial blood glucose in type 2 diabetic patients. Diabetol. Croat. 42 (2), 54–58.

De La Fuente, M.M., González, C.M., 2010. Laughter therapy for chronic skeletal muscular pain. Rev. Enfermeria (Barc., Spain) 33 (6), 43–44.

Dibeklioğlu, H., Salah, A.A., Gevers, T., 2015. Recognition of genuine smiles. IEEE Trans. Multimed. 17 (3), 279–294.

Dolgoff-Kaspar, R., Baldwin, A., Johnson, M.S., Edling, N., Sethi, G.K., 2012. Effect of laughter yoga on mood and heart rate variability in patients awaiting organ transplantation: a pilot study. Altern. Ther. Health Med. 18 (5), 61–66.

Dupont, S., Çakmak, H., Curran, W., Dutoit, T., Hofmann, J., McKeown, G., et al., 2016. Laughter research: a review of the ILHAIRE project. In: Toward Robotic Socially Believable Behaving Systems-Volume I. Springer International Publishing, pp. 147–181.

Duval, S., Tweedie, R., 2000. Trim and fill: a simple funnel-plot–based method of testing and adjusting for publication bias in meta-analysis. Biometrics 56 (2), 455–463.

Egger, M., Smith, G.D., Schneider, M., Minder, C., 1997. Bias in meta-analysis detected by a simple, graphical test. Br. Med. J. 315 (7109), 629–634.

Ekman, P., Davidson, R.J., Friesen, W.V., 1990. The Duchenne smile: emotional expression and brain physiology: II. J. Personal. Soc. Psychol. 58 (2), 342.

Farifeth, S., Mohammadi-Aria, A., Kiamanesh, A., Mofid, B., 2014. The impact of laughter yoga on the stress of cancer patients before chemotherapy. Iran. J. Cancer Prev. 7 (4), 179–183.

Foley, E., Matheis, R., Schaefer, C., 2002. Effect of forced laughter on mood. Psychol. Rep. 90 (1) 184-184.

George, J.R., Jacob, V., 2014. A study to assess the effectiveness of laughter therapy on depression among elderly people in selected old age homes at mangalore. International Journal of Nursing Education 6 (1), 152.

Gonot-Schoupinsky, F.N., Garip, G., 2018. Laughter and humour interventions for well-being in older adults: a systematic review and intervention classification. Complement. Ther. Med. 38, 85–91.

Han, H.J., Park, A., Kim, H.S., Moon, H., Park, Y.H., 2011. The effects of laughter therapy on stress responses in patients with preoperative breast cancer. Journal of Korean Oncology Nursing 11 (2), 93–100.

Hedges, L.V., Olkin, I., 1985. Statistical Methods for Meta-Analysis. Academic Press, New York.

Herschenhorn, S., 1995. The Impact of Focused Laughter Therapy on Chronic Pain in Young Women with Rheumatoid Arthritis.

Higgins, J.P., Thompson, S.G., Deeks, J.J., Altman, D.G., 2003. Measuring inconsistency in meta-analyses. Bmj 327 (7414), 557–560.

Higgins, J.P.T., Altman, D.G., Gøtzsche, P.C., Jüni, P., Moher, D., Oxman, A.D., et al., 2011. The Cochrane Collaboration's tool for assessing risk of bias in randomised trials. Br. Med. J. 343, d5928.

Hirosaki, M., Ohira, T., Kajiura, M., Kiyama, M., Kitamura, A., Sato, S., Iso, H., 2013. Effects of a laughter and exercise program on physiological and psychological health among community-dwelling elderly in Japan: randomized controlled trial. Geriatr. Gerontol. Int. 13 (1), 152–160.

Hirsch, R.D., Junglas, K., Konradt, B., Jonitz, M.F., 2010. Humor therapy in the depressed elderly: results of an empirical study. Zeitschrift fur Gerontologie und Geriatrie 43 (1), 42–52.

Hsieh, C.J., Chang, C., Tsai, G., Wu, H.F., 2015. Empirical study of the influence of a Laughing Qigong Program on long-term care residents. Geriatr. Gerontol. Int. 15 (2), 165–173.

Jung, H.W., Youn, C.H., Cho, N.R., Lee, M.K., Lee, J.B., 2009. The effect of laughter therapy on sleep in the community-dwelling elderly. Korean Journal of Family Medicine 30 (7), 511–518.

Jung, K.H., Park, S.Y., 2012. The effect of laughter program on the stress of nursing college students in clinical practice. The Korean Journal of Health Service Management 6 (1), 163–172.

Kessler, A.G., Dreher, U., Mueller, M., Fischer, G., Wiedemann, G.J., 2010. Humor therapy as a new concept in palliative care. J. Clin. Oncol. 28 (15).

Kessler, A., Mueller, M., Dreher, U., Wiedemann, G., 2012. Humor Therapy" in palliative care - a new self-contained concept. Onkologie 35 198-198.

Kim, S.H., Kim, Y.H., Kim, H.J., Lee, S.H., Yu, S.O., 2009. The effect of laughter therapy on depression, anxiety, and stress in patients with breast cancer undergoing radiotherapy. Journal of Korean Oncology Nursing 9 (2), 155–162.

Kim, S.H.S., 2010. The effectiveness of a laughter therapy group counseling program for anxiety, depression, and low self-esteem in mothers. Korean Journal of Child Studies 31 (2), 197–207.

Kim, K.S., Yu, J.A., Kim, J.A., Lee, Y.S., Lee, I.O., 2010. Effects of Tai-Chi exercise and self-help management program applying laughter therapy in patients with osteoarthritis. Journal of muscle and joint health 17 (1), 68–78.

Kim, A.H., Choi, W.S., Park, I.J., 2012. The effects of the group Art therapy with laughing activities on the stress-coping ability, social support and life-satisfaction level of children in after-school-daycare programs. Korean Journal of Art Therapy 19 (1), 1–24.

Kim, J.-A., Lee, E.H., 2012. A study on the effectiveness of a laughter therapy program for the mental health of the elderly. Korean journal of family welfare. Asian Oncology Nursing 9 (2), 155–162 17(1), 85-104.

Kim, S.H., Kook, J.R., Kwon, M., Son, M.H., Ahn, S.D., Kim, Y.H., 2015. The effects of laughter therapy on mood state and self-esteem in cancer patients undergoing radiation therapy: a randomized controlled trial. J. Altern. Complement. Med. 21 (4), 217–222.

Klauer, K.J., 2001. Handbuch Kognitives Training. Göttingen: Hofgrefe.

Ko, H.J., Youn, C.H., 2011. Effects of laughter therapy on depression, cognition and sleep among the community-dwelling elderly. Geriatr. Gerontol. Int. 11 (3), 267–274.

Konradt, B., Hirsch, R.D., Jonitz, M.F., Junglas, K., 2013. Evaluation of a standardized humor group in a clinical setting: a feasibility study for older patients with depression. Int. J. Geriatr. Psychiatry 28 (8), 850–857.

Koo, J.E., 2010. The influence of the laughter gymnastic program on elementary student's self-efficacy and social competence. Journal of Exercise Rehabilitation 6 (2), 3–12.

Koo, N.H., Kim, H.S., 2013. The effect of laughing exercise on stress and depression in teenagers. The Korean Journal of Sport 11 (4), 425–438.

Kraemer, H.C., Mintz, J., Noda, A., Tinklenberg, J., Yesavage, J.A., 2006. Caution regarding the use of pilot studies to guide power calculations for study proposals. Arch. Gen. Psychiatr. 63 (5), 484–489.

Lee, Y.M., Sohn, J.N., 2010. The effects of laughter therapy on anger, anger expressions and mental status after oil spill in victimized community residence. Journal of Korean Academy of Psychiatric and Mental Health Nursing 19 (2), 186–195.

Lee, K.I., Eun, Y., 2011. Effect of laugher therapy on pain, depression and sleep with elderly patients in long term care facility. Journal of muscle and joint health 18 (1), 28–38.

Lee, K.J., Ji, Y.H., 2011. A study on the utilization of laughter therapy program for the relief and treatment of the job stress (PTSD) of police officers. A collection of Police Science 6 (1), 69–93.

Lee, S.Y., Young, Y., 2011. Effect of laughing therapy on sleep, depression and self-esteem of elderly women in senior home. Journal of Regional Studies 19 (4), 211–225.

Lee, H.K., Byeon, D.H., Park, Y.S., Kim, J.S., Gil, J.H., 2013. Effects of the laughter therapy on blood pressure, depression and quality of life in rural elderly women. Journal of the Korea Academia-Industrial cooperation Society 14 (4), 1810–1819.

Low, L.F., Brodaty, H., Goodenough, B., Spitzer, P., Bell, J.P., Fleming, R., et al., 2013. The Sydney Multisite Intervention of LaughterBosses and ElderClowns (SMILE) study: cluster randomised trial of humour therapy in nursing homes. BMJ Open 3 (1).

Low, L.F., Goodenough, B., Fletcher, J., Xu, K., Casey, A.N., Chenoweth, L., et al., 2014. The effects of humor therapy on nursing home residents measured using observational methods: the SMILE cluster randomized trial. J. Am. Med. Dir. Assoc. 15 (8), 564–569.

Lowis, M.J., 1997. A Humor Workshop program to aid coping with life stress. Mank. Q. 38 (1–2), 25–38.

Martin, R.A., 2001. Humor, laughter, and physical health: methodological issues and research findings. Psychol. Bull. 127 (4), 504.

McCreaddie, M., Wiggins, S., 2008. The purpose and function of humour in health, health care and nursing: a narrative review. J. Adv. Nurs. 61 (6), 584–595.

Mesidor, M., 2009. Poster 38: evaluation of a laughter-based intervention for individuals with serious mental illnesses. Arch. PM&R (Phys. Med. Rehabil.) 90 (10) e23-e23.

Miles, C., Tait, E., Schure, M.B., Hollis, M., 2016. Effect of laughter yoga on psychological well-being and physiological measures. Adv. Mind Body Med 30 (1), 12–20.

Mora-Ripoll, R., 2011. Potential health benefits of simulated laughter: a narrative review of the literature and recommendations for future research. Complement. Ther. Med. 19 (3), 170–177.

Morris, S.B., 2008. Estimating effect sizes from pretest-posttest-control group designs. Organ. Res. Methods 11 (2), 364–386.

Nagendra, H.R., Chaya, M.S., Nagarathna, R., Kataria, M., Manjunath, M.K., Raghavendra, R.M., 2007. The Efficacy of Laughter Yoga on IT Professionals to Overcome Professional Stress. Laughter Yoga International, Mumbai, India.

Neuhoff, C.C., Schaefer, C., 2002. Effects of laughing, smiling, and howling on mood. Psychol. Rep. 91 (3f), 1079–1080.

Oh, E., Gang, G., Hand, O., Woo, M., Lee, M., Kim, S., 2011. The effects of laughter therapy on stress response and coping for specialized burn hospital nurses. The Korean Journal of Stress Research 19 (4), 323–331.

Park, H., Kim, S., Heo, E., Yang, D., 2007. Effect of laughter therapy on stress and health-related quality of life in hemodialysis patients. Value Health 10 (6). https://doi.org/10.1016/s1098-3015(10)65132-7. A313-A313.

Park, H.W., 2010. The effect of laughter therapy program on mental health. The Korean Journal of Stress Research 18 (3), 287–294.

Park, E., 2013. Effects of visiting laughter therapy on depression and insomnia among the vulnerable elderly. Journal of Korean Academy of Community Health Nursing 24 (2), 205–213.

Park, S.Y., Kwak, E.A., Heo, J.E., Shin, E.K., Kim, J.H., Lee, J.Y., Lee, J.H., 2013. The effects of laughter therapy on the anxiety and depression of inflammatory bowel disease patients. Journal of Korean Clinical Nursing Research 19 (1), 105–114.

Peters, J.L., Sutton, A.J., Jones, D.R., Abrams, K.R., Rushton, L., 2007. Performance of the trim and fill method in the presence of publication bias and between-study heterogeneity. Stat. Med. 26, 4544–4562. https://doi.org/10.1002/sim.2889.

Raja, A., Sundari, F.J., 2014. Comparison of the efficacy of laughter therapy and breathing exercises on pulmonary function among smokers. Asian Journal of Nursing Education and Research 4 (1), 105.

Ryu, Kyung Hee, Shin, Hye Sook, Yang, Eun Young, 2015. Effects of laughter therapy on immune responses in postpartum women. J. Altern. Complement. Med. https://doi.org/10.1089/acm.2015.0053.

Sahai-Srivastava, S., Joyce, A., Christie, C.C., Ugurlu, C., Manh, C., Kimblin Licht, H., 2014. Pilot study of laughter yoga therapy in treatment of chronic migraine. Headache 54, 30.

Sakai, Y., Takayanagi, K., Ohno, M., Inose, R., Fujiwara, H., 2013. A trial of improvement of immunity in cancer patients by laughter therapy. Jpn. Hosp.: *the journal of the Japan Hospital Association* (32), 53–59.

Shahidi, M., Mojtahed, A., Moddabbernia, A., Mojtahed, M., Shafiabady, A., Delavar, A., Honari, H., 2011. Laughter yoga versus group exercise program in elderly depressed women: a randomized controlled trial. Int. J. Geriatr. Psychiatry 26 (3), 322–327. https://doi.org/10.1002/gps.2545.

Shin, H.S., Ryu, K.H., Song, Y.A., 2011. Effects of laughter therapy on postpartum fatigue and stress responses of postpartum women. Journal of Korean Academy of Nursing 41 (3), 294–301.

Song, M., Park, K., Park, H., 2011. The effect of laughter-therapy on mood and life satisfaction in older adults staying at nursing facilities. Gerontol. 51 476-476. Journals Dept, 2001, Evans RD, Cary, NC27513 USA: Oxford Univ Press Inc.

Song, M.S., Park, K.M., Park, H., 2013. The effects of laughter-therapy on moods and life satisfaction in the elderly staying at care facilities in South Korea. Journal of Korean Gerontological Nursing 15 (1), 75–83.

Szabo, A., 2003. The acute effects of humor and exercise on mood and anxiety. J. Leis. Res. 35 (2), 152–162.

Szabo, A., Ainsworth, S.E., Danks, P.K., 2005. Experimental comparison of the psychological benefits of aerobic exercise, humor, and music. Humor-International Journal of Humor Research 18 (3), 235–246. https://doi.org/10.1515/humr.2005.18.3.235.

Tan, S.A., Tan, L.G., Lukman, S.T., Berk, L.S., 2007. Humor, as an adjunct therapy in cardiac rehabilitation, attenuates catecholamines and myocardial infarction recurrence. Adv. Mind Body Med. 22 (3–4), 8–12.

Tse, Mimi M.Y., Lo, Anna P.K., Cheng, Tracy L.Y., Chan, Eva K.K., Chan, Annie H.Y., Chung, Helena S.W., 2010. Humor therapy: relieving chronic pain and enhancing happiness for older adults. J. Aging Res 2010, 343574 9 pages. https://doi.org/10.4061/2010/343574.

Verde, P.E., Ohmann, C., 2015. Combining randomized and nonrandomized evidence in clinical research: a review of methods and applications. Res. Synth. Methods 6 (1), 45–62. http://doi.org/10.1002/jrsm.1122.

Wagner, H., Rehmes, U., Kohle, D., Puta, C., 2014. Laughing: a demanding exercise for trunk muscles. J. Mot. Behav. 46 (1), 33–37.

Walter, M., Hanni, B., Haug, M., Amrhein, I., Krebs-Roubicek, E., Muller-Spahn, F., Savaskan, E., 2007. Humour therapy in patients with late-life depression or Alzheimer's disease: a pilot study. Int. J. Geriatr. Psychiatry 22 (1), 77–83.

Weinberg, M.K., Hammond, T.G., Cummins, R.A., 2014. The impact of laughter yoga on subjective wellbeing: a pilot study. Eur. J. Humour Res. 1 (4), 25–34.

Yazdani, M., Esmaeilzadeh, M., Pahlavanzadeh, S., Khaledi, F., 2014. The effect of laughter Yoga on general health among nursing students. Iran. J. Nurs. Midwifery Res. 19 (1), 36–40.

Yim, J., 2016. Therapeutic benefits of laughter in mental health: a theoretical review. Tohoku J. Exp. Med. 239 (3), 243–249.

You, E.S., Choi, Y.H., 2012. The effects of laughter therapy on pain and anxiety among patients with postmastectomy. Journal of East-West Nursing Research 18 (1), 47–52.

Yu, J.A., Kim, K.S., 2009. Effects of laughter therapy on stress response and pain of military personnel with low back pain in hospital. Journal of muscle and joint health 16 (1), 36–45.