Abstract
This systematic review and meta-analysis aimed to evaluate the effectiveness of psychological interventions in improving quality of life for head and neck cancer patients. Five databases were systematically searched in July 2016. Studies were included if they reported original empirical data from intervention studies utilising psychological approaches (excluding psychoeducational-only interventions) and provided data on quality of life outcomes. Six studies, involving 185 participants, fulfilled eligibility criteria. Study designs included a case study, single-group designs, non-randomised controlled trials and one randomised controlled trial. Meta-analysis of two studies did not provide support for the effectiveness of psychological intervention improving total quality of life scores (or subscales) compared to control groups at end of intervention. Intervention studies evaluating psychological interventions for patients with head and neck cancer have produced insufficient data to support their effectiveness for improving quality of life. This review further highlights the limited evidence base within this area. Existing studies are based on small samples and are inconsistent regarding: intervention type, duration, and intensity; follow-up measurement periods; and methodological quality. Further research, addressing these limitations, is required for more definitive conclusions to be drawn about the effectiveness of psychological interventions with this population.

Keywords: head and neck cancer, psychological intervention, quality of life, systematic review, meta-analysis
Introduction
Head and neck cancer is the seventh most prevalent cancer worldwide, accounting for 4.8% of newly diagnosed cancers in 2012 (Ferlay et al., 2015). Head and neck cancer comprises cancers of various sites, including lip, oral cavity, nasopharynx, pharynx and larynx. Vital structures, located within these sites, are responsible for essential functions such as eating, swallowing, breathing, and speech. Head and neck cancer is often diagnosed late, thus treatment is often intrusive including surgery, radiotherapy and/or chemotherapy and can have severe side effects (Neilson et al., 2010; Reich et al., 2014). Head and neck cancer has been described as the most “emotionally traumatic” of all cancers due to its detrimental impact on appearance and fundamental functions (Björklund, Sarvimäki, & Berg, 2010).

Previously, survival was the exclusive aim of cancer treatment, however, quality of life is increasingly recognised as an important outcome, and has been linked to predicting survival (Morton, 2012; Osthus, Aarstad, Olofsson, & Aarstad, 2011). An increase in the number of papers published regarding quality of life in head and neck cancer patients demonstrates growing recognition of the importance and relevance of this outcome within this population (Rogers, Ahad, & Murphy, 2007).

Quality of life is a difficult concept to define and quantify as it includes many aspects and is inherently based in the subjective experience of individuals. The World Health Organisation (WHO, 1995) defines quality of life as “individuals’ perception of their position in life in the context of the culture and value systems in which they live and in relation to their goals, expectations, standards and concerns”. It has also been defined as the ‘difference’ between an individual’s present experience and their hopes and expectations at a particular time point (Calman, 1984). Since it is likely that our expectations are formed largely from previous experience, if changes in experience or function occur, as is often the case with head and neck cancer patients, it is unsurprising that quality of life may be diminished for these individuals (Morton, 2012).

Head and neck cancer bears unique challenges as it imposes changes in lifestyle and everyday functioning including physical function, interpersonal relationships and social functioning (Björklund et al., 2010; Semple, Dunwoody, George Kernohan, McCaughan, & Sullivan, 2008). These enduring difficulties tend to impinge upon other functions and are reported to be a constant reminder of the changes and challenges head and neck cancer patients must face and overcome daily. Lang, France, Williams, Humphris, and Wells (2013) presented a metasynthesis of the lived experience of head and neck cancer, evidencing ways in which these cancers may challenge and erode quality of life. For example, core themes arising from this metasynthesis included ‘disruption to daily life’ (physical functioning, financial stability, social life, emotions and family life), ‘diminished self’ (losses associated with sense of self, confidence, self-esteem/image, anticipated healthy future), and ‘finding a path’ (regarding how the individual may see the future as either different or diminished).

A recent review identified numerous predictive psychological factors for lower quality of life within head and neck cancer patients, including: depressive symptoms, fear of recurrence, anxiety, and both active and passive coping strategies (Dunne et al., 2016). Depression has also been reported to increase significantly over the duration of radiotherapy treatment, in association with a significant decline in quality of life (Neilson et al., 2010). Furthermore, suicide rates are four times higher in head and neck cancer patients than in the general population and other cancer populations,
which may be due to the “disproportionate impact” that head and neck cancer and its treatment have on quality of life (Zeller, 2006). These high suicide rates highlight the importance of improving not only survival following head and neck cancer diagnosis and treatment, but also quality of life. Resulting from these observations, strong statements have been made about the importance of improving quality of life within head and neck cancer patients; it is “the single most important construct” (Humphris, 2008, p. 109) in facilitating adjustment to head and neck cancer.

There is a need for psychological input into teams which manage head and neck cancer (Humphris, 2008; Morton, 2012). Shiraz, Rahtz, Bhui, Hutchison, and Korszun (2014) reported an association between high levels of psychological distress and low quality of life, and asserted that 40% of individuals showing this pattern would accept psychological support, but the majority had not actively pursued it. It is clear therefore that psychological intervention is needed, however the evidence base is currently inconclusive regarding the efficacy of psychological interventions for improving quality of life for head and neck cancer patients.

To date, two systematic reviews have synthesised the evidence regarding psychosocial interventions for head and neck cancer (Luckett, Britton, Clover, & Rankin, 2011; Semple et al., 2013) and have reported that it was not possible to come to strong conclusions regarding the effectiveness of these interventions. One review limited inclusion to randomised and quasi-randomised controlled trials reporting psychological distress and/or quality of life (Semple et al., 2013). Whilst randomised designs may represent an evidential ‘gold standard’, restriction to these designs may be considered premature, given the current (inchoate) state of the literature. Within this review, data for only 26% of participants were extractable for analysis. The second review (Luckett et al., 2011) also reviewed psychosocial interventions for use with head and neck cancer patients, but the authors considered that the data were too heterogeneous to support meta-analysis. A notable limitation of previous reviews is that both included interventions based on psychoeducation without other psychological components; such a broad definition of ‘psychological intervention’ may limit interpretability of findings and potentially assuage aggregate estimates of effect-size (by conflating low-intensity didactic interventions with high-intensity psychotherapeutic interventions). Consistent with this argument, an extensive review of mixed cancer studies reported moderate positive effect-sizes of individual psychotherapy for quality of life, maintained throughout follow-up periods, whereas psychoeducational interventions only produced small effects on quality of life (Faller et al., 2013). Moreover, a narrative review suggested that interventions using specific psychological (e.g. cognitive and behavioural) techniques, delivered by a professional with an understanding of psychological distress, may be more beneficial for people with head and neck cancer than an educational and emotionally supportive programme (Semple, Sullivan, Dunwoody, & Kernohan, 2004).

Objective and rationale for current review
This review aimed to assess whether psychological interventions are effective in improving quality of life in head and neck cancer patients. The review built upon the evidence provided in previous reviews by reducing the potential heterogeneity of data collected: through (1) focusing on one primary outcome (quality of life), and (2) being more selective about the ‘psychological interventions’ included. Conversely, in recognition of the incipient state of research in this area, this review was relatively inclusive with respect to study design.
Methods

Inclusion/exclusion criteria

Publications were considered for inclusion if they met the following criteria:

1. Delivery of psychological interventions to people with head and neck cancer;
2. Original empirical data presented (e.g. not a review, prospective, or descriptive paper);
3. Quality of life reported as an outcome measure (no limit on whether this was a primary or secondary outcome);
4. Sourced from peer-reviewed publications (i.e. not grey literature; dissertations, conference/poster abstracts);
5. Available in English.

However, studies were excluded if:

1. Outcomes related to a mixed group of cancer patients and head and neck cancer results were not available separately;
2. The intervention was predominantly psychoeducation. Whilst psychoeducation is a common component of broader psychological interventions (such as cognitive-behavioural therapy; CBT; Blagys & Hilsenroth, 2002), standalone psychoeducation was not considered to be equitably comparable to these more comprehensive (multifaceted and psychological theory-based) interventions. Empirically, as discussed in the introduction, psychoeducation-only interventions have been shown to produce smaller effect-sizes (relative to more comprehensive interventions) in other cancer populations: considering psychoeducation to have parity with broader psychological interventions (like CBT) was thus considered likely to under-estimate their efficacy. Thus, this review followed the precedent of other reviews of psychological interventions (e.g., Olthuis, Watt, Bailey, Hayden, & Stewart, 2016): excluding psychoeducation unless delivered as part of a broader psychological intervention (meeting criteria outlined in the checklist below).
3. Psychological intervention was used but data were not separable from other intervention modalities (e.g. multidisciplinary/stepped care).

A checklist was developed to structure whether an intervention was deemed “psychological” or not. The checklist comprised of 3 items:

- Was a recognised psychological program/protocol used for the delivery of the intervention? (if answered “yes” and other criteria were met the study was included)
- Does the intervention clearly state the psychological principles and/or theories informing the intervention? (e.g. based on cognitive behavioural therapy, cognitive theory, learning theory; if answered “yes” the following question was considered)
  - Is the intervention based upon more than psychoeducation and non-directive professional support? (if this was also answered “yes” and other criteria were satisfied, the study was included).

No limitations were placed upon study design, year of publication, or target of intervention.

Search methods

Systematic searches were conducted across Medline, PsycINFO, Embase, CINAHL and Web of Science databases in July 2016. Search strategies were modelled upon, and adapted from, those used by previous similar reviews (Luckett et al., 2011; Semple et al., 2013). The search strategy used for Medline is presented in Figure 1.
‘Quality of life’ was not included as a search term as some articles were not indexed with this heading if it was not the primary outcome. This was instead included within eligibility criteria. Reference lists of relevant articles, including reviews, were searched in order to identify any additional studies which may meet eligibility criteria.

Study selection
Following the removal of duplicates, articles were initially screened by title and abstract for their eligibility for inclusion by the first author, though when this did not provide sufficient data to determine eligibility, full texts were obtained and examined. Grey literature (dissertations, poster and conference abstracts) was excluded. A summary of reasons for exclusion are presented in Figure 2.

Data extraction
The first author extracted data from each article using a pre-designed form which included:

- General: Author, year of publication, journal, country and publication title.
- Participants: Age, gender, site and stage of cancer, cancer treatment received, inclusion/exclusion criteria, sample size, sample size at intervention completion and follow-up.
- Methodology: Design of the study, including whether it was a pilot or feasibility study (methodology quality is discussed below).
- Intervention: Intervention description, mode of delivery, professional delivering intervention and training received, target of intervention, number of sessions, duration, intervention integrity and characteristics of control intervention.
- Outcome: Quality of life measure, other outcome measures, time points of outcomes/follow-up, key findings, quality of life results including raw scores, means, standard deviation and effect size if reported.

Quality appraisal
Methodological quality was rated by the first author using the “Checklist for Measuring Study Quality” (Downs & Black, 1998). This quality appraisal tool was specifically designed for use with healthcare intervention studies, and is not limited to the appraisal of randomised controlled trials. Quality appraisal is based upon five subscales relating to the study quality and reporting, external validity, internal validity (separated into bias and confounding subscales) and power. Items within these subscales are rated as “1” (“Yes”) or “0” (“No” or “Unable to determine”). Most subscales of the checklist were reported to demonstrate high reliability in terms of internal consistency (Kuder-Richardson formula 20; KR-20 > .69), test-retest reliability (r = .69–.90) and inter-rater reliability (r > .70); exceptionally, the external validity subscale demonstrated poor reliability (KR-20 = .54, r = .37, and r = -.14 respectively).

Data synthesis
Reported data for quality of life measures were continuous. Results were synthesised according to study type. Extracted data from controlled studies were entered into the Cochrane Collaboration’s Review Manager (RevMan; The Nordic Cochrane Centre, 2014). Means and standard deviations were extracted where available from published data. Where this was not directly reported, standard deviation was calculated from standard errors using 95% confidence intervals. A random effects model was used for all meta-analyses, and standardised mean difference rather than mean difference was used when studies utilised different quality of life measures. Where insufficient published data precluded meta-analysis, studies were excluded and compared.
narratively. Forest plots were created for meta-analysis comparisons. Heterogeneity was assessed through RevMan using I². Despite efforts to limit heterogeneity in some respects (e.g., definition of ‘psychological intervention’, focus on a single outcome) it was anticipated that some heterogeneity would remain due to the purposive absence of restrictions on study design and target of intervention.

Where multiple post-intervention quality of life scores were available, the time points which were most similar in timeframe from the end of intervention were used to compare across studies.

Where individual data was presented (case study design), reliable change indices were calculated according to Jacobson and Truax (1991) and based upon estimates of reliability and standard deviation described by List et al. (1996) who reported upon the performance of 151 head and neck cancer patients. Effect sizes (Cohen’s d) were extracted or calculated for single-group pre-post designs and compared narratively. Meta-analysis and comparisons were planned for total quality of life scores and for individual subscales.

Results regarding additional outcomes, such as psychological distress, were also extracted and described narratively, though it is acknowledged that some studies investigating effects on distress have been excluded due to their omission of quality of life measures. As such, these data were extracted to contextualise findings for contemporaneous quality of life outcomes and were not intended to provide a comprehensive review and analysis of effects of psychological interventions on psychological distress. In view of this, a ‘vote counting’ approach to additional outcomes was considered sufficient: vote counting is defined by the Cochrane Handbook (Higgins & Green, 2011) as comparing number of studies finding positive effects to number of studies finding negative effects, thus enabling the question “is there any evidence of an effect?” to be answered.

**Results**

*Results of the search/study selection*

Database searches retrieved 1718 articles and reference list searches returned 20 articles of interest. This number reduced to 1435 following deletion of duplicates. Following title, abstract, and full-text screens, six articles, including 185 participants, met eligibility criteria. Figure 2 details the process of identifying and excluding studies, summarising reasons for exclusion of full-text articles. With the exception of the case study (n = 1), sample sizes ranged from 19 to 55 participants.

*Study characteristics*

Table 1 contains summaries of the sample and intervention type, delivery, duration, adherence and integrity checks.

Five studies cite cognitive-behavioural principles as the therapeutic framework from which interventions were built, however this was operationalised differently across studies. Hammerlid, Persson, Sullivan, and Westin (1999) delivered their intervention with a psychologist-led group which employed cognitive-behavioural techniques though did not describe these. Semple, Dunwoody, Kernohan, and McCaughan (2009) and Kilbourn et al. (2013) also stated the use of cognitive-behavioural principles “with an emphasis on social skills training” and additional psychoeducation respectively. Kangas, Milross, and Bryant (2014) provided a detailed session-by-session plan of the cognitive-behavioural therapy program, describing the techniques used by the facilitating clinical psychologists; this was also the intervention provided

Control groups were utilised within three studies (Hammerlid et al., 1999; Kangas et al., 2013; Semple et al., 2009) with differing allocation strategies and provisions ranging from simple collection of data with no additional input (Hammerlid et al., 1999) to “non-directive supportive counselling” including psychoeducation and unstructured problem-solving (Kangas et al., 2013).

**Methodological features**

Studies were rated on each of the subscales presented by Downs and Black (1998); an overview is provided in Table 2.

Generally, lower scores were achieved by studies without control groups (Kangas et al., 2014; Kilbourn et al., 2013; Pollard et al., 2016). All studies scored consistently low on external validity due to the omission of data regarding representativeness of samples from the population source. Studies generally reported participant characteristics, but did not comment upon representativeness, with the exception of Semple et al. (2009). However, Downs and Black (1998) acknowledged the development needs of the external validity subscale since it has the poorest reliability and validity of the five subscales. Nevertheless, whilst this tool suggests low external validity for all studies, the papers often recognised the limited generalisability of their findings, stating the need for larger trials, since they were mostly pilot or feasibility studies (Hammerlid et al., 1999; Kangas et al., 2013; Kilbourn et al., 2013; Pollard et al., 2016; Semple et al., 2009) and therefore limited in scope.

Only one study (Kangas et al., 2013) addressed power analysis for their sample size; the necessary sample size was met (n = 14 per group), however at follow-up only 10 participants remained within the control group. Statistical analyses were completed on an “intent to treat” basis in order to minimise biasing effects of noncompliance and withdrawal. The remaining studies did not acknowledge the sample size needed for sufficient power, likely due to being pilot or feasibility studies, and as such have scored “0” within the quality appraisal checklist.

**Key findings**

An overview of the methodology, control group characteristics, outcome measures, key reported results and effect sizes (where calculable) are presented in Table 3.

Quality of life was measured using the Functional Assessment of Cancer Therapy-General scale (FACT-G; Cella et al., 1993) in four studies (Kangas et al., 2013; Kangas et al., 2014; Kilbourn et al., 2013; Pollard et al., 2016), three of which also provided data from the head and neck cancer specific subscale (FACT-H&N; Kangas et al., 2014; Kilbourn et al., 2013; Pollard et al., 2016). FACT-G is composed of four subscales (emotional, physical, functional and social/family wellbeing) which require ratings to statements on a 5-point Likert scale ranging from “Not at all” (0) to “Very much” (4). The head and neck cancer subscale is rated upon the same scale and includes 12 statements related to common side effects of head and neck cancer including difficulties with swallowing, breathing, voice and pain. Hammerlid et al. (1999) used the European Organization for Research and Treatment of Cancer Quality of Life Questionnaire-Core 30 (EORTC QLQ-C30), which is a well validated measure, and the head and neck cancer specific scale (EORTC QLQ-H&N37; Bjordal et al., 1994), where statements, separated into subscales are rated on a Likert scale. Semple et al. (2009) utilised the University of Washington Quality of Life
Questionnaire version 4 (UWQoLv4; Rogers et al., 2002). Higher scores indicate better quality of life in all measures.

Quality of Life
Total quality of life scores were available from three controlled studies (Hammerlid et al., 1999; Kangas et al., 2013; Semple et al., 2009). However, Hammerlid et al. (1999) only provided mean scores and no estimates of variability, it was therefore not possible to include this study within the meta-analysis; it is presented in Figure 3 and excluded thereafter. There was no observable effect of psychological intervention on total quality of life within the meta-analysis (standardised mean difference (SMD) 0.03, 95% confidence interval (CI) -0.62 to .68), as indicated by the confidence interval including 0. However, sample sizes are small and heterogeneity is moderate (p = .13, I² = 56%), therefore these results should be taken tentatively (confidence intervals are wide, and do not preclude the possibility of moderate effect sizes in either direction). Additionally it was only possible to pool data from one time point (at intervention completion) as follow-up periods were too different (3 month follow-up vs 6 and 12 month follow-up); it is possible that later measures may have shown a greater difference between groups, as was suggested within the study papers.

Since the studies (Kangas et al., 2013; Semple et al., 2009) included within meta-analysis used different measures for quality of life, the subscales from UWQoLv4 were categorised using the subscales from FACT-G. Averages of means and standard deviations were used from the UWQoLv4 subscales. Only three subscales were comparable; Functional Wellbeing (including Activity and Recreation subscales from UWQoLv4), Emotional Wellbeing (including Mood and Anxiety subscales from UWQoLv4) and Physical Wellbeing (including Pain, Appearance, Swallowing, Chewing, Speech, Shoulder, Taste and Saliva from UWQoLv4).

As can be seen in Figure 4 and Figure 6, results do not support psychological interventions being beneficial for functional or physical wellbeing as assessed by quality of life measures at the end of intervention (SMD -0.27, 95% CI -1.50 to .96 and SMD -0.13, 95% CI -.55 to .29 respectively). Heterogeneity for functional wellbeing subscale was also very high (p = .006, I² = 87%). However, whilst not reaching significance, there was a slight trend towards intervention groups having improved emotional wellbeing compared to controls at the end of intervention (SMD -0.28, 95% CI -.71 to .14). Thus suggesting that subscale analysis may be advantageous when considering quality of life since any small effects between scales may be masked when examining total quality of life scores.

Hammerlid et al. (1999) reported that their intervention group showed improvement on six symptoms and functions at 12 month follow-up, based on a score change of >10 being "possibly" clinically relevant. This was compared with the control group who showed minimal change on most subscales, and only clinically relevant change on one symptom (hoarseness). From these results it appears that, unlike the studies within the meta-analysis, Hammerlid et al’s (1999) intervention produced greater beneficial effects compared to a control group on several domains (largest changes occurred in emotional and social functioning and overall quality of life scores), however, without additional statistical testing, this result must be taken cautiously.

Kangas et al’s (2014) case study showed significant and reliable improvement at several time points on three scales; for emotional wellbeing at the end of intervention and at 6-months follow-up (Reliable Change Indices (RCI) = 2.65, p = .01 and RCI = 3.43, p < .01 respectively), where a reliable change index over 1.96 reflects a
statistically reliable change (i.e. change that is not attributable to measurement error or chance variation). Functional wellbeing showed a statistically reliable change at 6- and 12-month follow-up (RCI = 2.76, p = .01 at both time points). Total FACT-G scores also showed significant and reliable improvement at 6- and 12-month follow-up (RCI = 4.16, p < .01 and RCI = 2.55, p = .02 respectively). The head and neck cancer subscale approached reliable change (RCI = 1.94, p = .06) at 12-month follow-up. However it has been reported that studies of single cases are likely to overestimate intervention effects when compared to group designs (Beretvas & Chung, 2008).

This is evident within the current review as the two single-group design studies showed only small effect sizes from pre- to post-intervention on quality of life total and subscale scores – with one study demonstrating effects in the direction of deterioration. Kilbourn et al. (2013) reported small, negative effect sizes at the end of intervention for all subscales on FACT-G ($d = -0.08$ to $-0.35$), though did not report on total quality of life. Slightly larger declines were observed within physical and functional wellbeing. These show a similar pattern to the studies included within the meta-analysis. The FACT-H&N score showed a moderate (defined by an effect size of 0.50 or greater) negative effect following intervention ($d = -0.66$). Pollard et al. (2016) reported the opposite trend, showing a large effect size (where $d \geq .80$) for total quality of life and physical wellbeing ($d = 1.06$, $d = 0.83$ respectively) and a moderate effect size for functional wellbeing ($d = .60$). Emotional and social wellbeing showed negligible change ($d = 0.02$, $d = -0.04$ respectively). However, when effect size was recalculated (partial eta squared) after controlling for pre-intervention mindfulness, the effect sizes showed the same pattern as the previous studies; favourable results for emotional and social wellbeing over physical and functional wellbeing. These results must be considered cautiously as they are based on small sample sizes and do not have control comparisons.

**Other Outcome Measures**

A variety of other outcomes were also reported, cited as primary or secondary findings, including anxiety, depression, post-traumatic stress symptoms, mindfulness, social impairment, social support and pain. A summary of the positive effects can be seen in Table 4 (negative effects are not reported).

All studies reporting anxiety showed a positive effect of intervention. Two of these improvements (40%) were present at the end of intervention (Pollard et al., 2016; Semple et al., 2009), and three (60%) were observed at 12-month follow-up (Hammerlid et al., 1999; Kangas et al., 2013; Kangas et al., 2014).

Positive effects were reported in all studies investigating depression, though this was negligible in one study due to the control group also showing improvement in depression scores (Hammerlid et al., 1999). Four of the five studies investigating depression (80%) reported positive effects of the intervention at post-intervention. Beneficial effects compared to controls were only observed at 12-month follow-up in Kangas et al’s (2013) study, and Pollard et al. (2016) only reported positive effects once pre-intervention mindfulness was controlled for.

All three studies measuring post-traumatic stress symptoms reported positive effects of intervention (Kangas et al., 2013; Kangas et al., 2014; Kilbourn et al., 2013). There was minimal difference in mindfulness at post-intervention when compared to pre-intervention, though scores correlated with the amount of meditation practice per week (Pollard et al., 2016). There was a small trend for improvement in social support
in one study (Semple et al., 2009). One study investigating pain and social support found no evidence and a small unfavourable change respectively (Kilbourn et al., 2013).

Discussion
Summary of main results
Six studies investigating psychological interventions for head and neck cancer patients met inclusion criteria for the current review. Evidence provided by the studies was not sufficient to support the effectiveness of psychological intervention upon total quality of life scores within this population. Indeed, one constraint of this review included the focus on quality of life, which may have contributed to the lack of findings (see limitations section). However there was a small trend towards improvement upon emotional wellbeing, physical wellbeing, and functioning wellbeing subscales. Disparity of results between subscales must be considered when investigating overall quality of life as such differences may cancel each other out, resulting in the appearance of no change. It was also observed that where changes in quality of life scores did occur, these tended to be at a later point within the cancer journey; 6-12 months follow-up rather than at the end of psychological intervention. Since four studies actively selected participants undergoing concurrent cancer treatment (most often radiotherapy), this suggests that a degree of adjustment and reduced side-effects from cancer treatment may be necessary before any beneficial effects of psychological intervention may be observed. Where improvements in quality of life were seen, these were most often preceded by a decrease in psychological distress.

There was large variability between the studies regarding design, intervention type, delivery, and target. Most interventions were based upon cognitive-behavioural principles though these were operationalised in different ways and combined with additional theories and approaches. Due to the variability within the studies and small sample sizes, the findings from this review should be taken tentatively; nevertheless, the findings reflect a need to develop a stronger evidence-base in relation to the effectiveness of psychological interventions for improving the quality of life in patients with head and neck cancer.

Applicability of evidence
The available evidence does not support generalizable results regarding the efficacy and effectiveness of psychological interventions for quality of life for people with head and neck cancer. This is due to small sample sizes and a lack of extractable data published within studies. Pooling data for meta-analysis was only possible for two studies (representing 48.1% of the total sample). The large variability within psychological interventions and the corresponding differences in delivery (including delivery mode and professional delivering the intervention), study design and omission of sample representativeness information contributed to the limited generalisability of results. Intensity and duration of interventions varied widely between studies, from 2 telephone sessions (Kilbourn et al., 2013) to 10-month group psychotherapy (Hammerlid et al., 1999). Control comparisons also varied from no additional input (Hammerlid et al., 1999) to non-directive supportive counselling which included six weekly 90 minute support sessions with a psychologist (Kangas et al., 2013). It is therefore unsurprising that studies may have found differing levels of change in quality of life.

The focus of interventions also varied between studies, and whilst it is acknowledged that a primary focus on quality of life was not specified within inclusion criteria, the
difference in intervention focus may impact upon quality of life results. Furthermore, only two studies specified psychological distress/dysfunction as a necessary criterion for participants to be offered interventions. It has previously been recommended that routine screening for psychological distress within cancer patients may be beneficial for interventions to specifically target the areas of difficulty identified and that the “best evidence” for these interventions is likely to be provided by studies selecting participants based on increased levels of psychological distress (Faller et al., 2013; Jacobsen & Jim, 2008; Mitchell et al., 2012).

A particular issue for the applicability of results is the lack of consistent follow-up data. Greater differences were reported at follow-up, suggesting the importance of obtaining these outcomes since the greatest effects may be observed after a passage of time. A review of mixed cancer samples reported that effects on quality of life were generally seen at medium and long-term follow-up rather than at the end of intervention, though these results were also tentative (Newell, Sanson-Fisher, & Savolainen, 2002). Four studies within the current review selected patients undergoing concurrent cancer treatment (Kangas et al., 2013; Kangas et al., 2014; Kilbourn et al., 2013; Pollard et al., 2016), therefore at the end of intervention, side-effects from this treatment were still present. Only three studies provided follow-up data for a period greater than 3 months post-intervention (Hammerlid et al., 1999; Kangas et al., 2013; Kangas et al., 2014). Nevertheless, the available follow-up data suggests that, once physical side-effects of radiotherapy/cancer treatment have reduced, the effects of the intervention may be more apparent - though without controlled trials this suggestion remains somewhat speculative. It would also be advantageous to investigate the effects of psychological interventions for patients who have completed treatment since resulting physical sequela may persist for a prolonged period of time (Howren, Christensen, Karnell, & Funk, 2013).

Quality of evidence
The quality of study design varied – one randomised controlled trial; two non-randomised controlled trials; two pre-test post-test single group designs and one case study. The lack of control groups within some studies substantially reduced their quality. Where present, the characteristics of the control groups varied widely, thus making comparison difficult. Allocation to groups within non-randomised studies left them vulnerable to bias as one allowed self-selection to intervention or control group (resulting in greater psychological distress for intervention group at baseline; Semple et al., 2009), and the other based allocation on distance from the hospital (Hammerlid et al., 1999). However, Kangas et al. (2013) note that randomisation to group may result in dissatisfaction with a control intervention and subsequently bias attrition rates as some participants withdrew from the study stating the reason as wanting to be in the cognitive-behavioural intervention group. There are also substantial pragmatic issues with blinding investigators to group allocation when considering psychological interventions.

The statistical power of the studies included within this review was also limited. One study supplied an estimate of required sample size for the study to be sufficiently powered, but, whilst the necessary number per group was recruited, attrition rates resulted in the study being underpowered (Kangas et al., 2013). The remaining studies did not acknowledge power, though all stated that they were either pilot or feasibility studies. Difficulty with recruitment left the studies vulnerable to bias; it was reported that between 11.5% and 78% of participants invited to participate declined. Attrition rates were also high within most studies. Semple et al. (2009) reported the lowest attrition rate of 9.26%, however within the remaining studies (Hammerlid et al.,
1999; Kangas et al., 2013; Kilbourn et al., 2013; Pollard et al., 2016) almost half of the participants dropped out between consent and study completion (38.5% - 48.6%).

Only two studies addressed the integrity of the intervention delivered, in one case this was open to bias as adherence to intervention and session content was rated by the therapist on a checklist and was not externally validated (Pollard et al., 2016). Kangas et al. (2013) randomly monitored 25% of sessions to ensure intervention integrity and fidelity.

Limitations of this review
There are several limitations to this current review which may have introduced bias within the process. Owing to the large number of studies returned through initial searches, grey literature was excluded, thus allowing the possibility of publication bias.

A focus upon quality of life, which allowed for some meta-analysis, may have overlooked other relevant data and therefore introduced bias in results and reporting. The a priori interest in quality of life was considered apt, given evidence for the myriad ways in which head and neck cancer (and its treatment) can limit functional wellbeing (Lang et al., 2013) – moreover, quality of life can be considered a common endpoint of other outcomes of interest (such as distress; e.g., Dunne et al., 2016). However, quality of life represents a more distal outcome with a less established evidence-base versus e.g., psychological interventions for distress in cancer and non-cancer populations. The possibility that a selective focus on quality of life versus distress under-estimates the efficacy of reviewed interventions is considered further in the next sub-section.

Furthermore, the current review stipulated that only studies investigating psychological interventions should be included. Criteria were developed to define psychological intervention and to distinguish this from psychosocial interventions, however these terms can often be used interchangeably and a distinction may be superficial. This is further compounded by limited information regarding interventions, therefore decisions of inclusion or exclusion may also be subjective. Associated with this, only the first author reviewed titles, abstracts and full-texts to determine eligibility, although attempts were made to limit this potential bias by consulting with the second author if uncertainty about inclusion arose. Eligible studies may have been missed within searches, however, broad search terms were used and terms known to limit results were excluded from the search strategy. Additional manual searches of relevant reference lists were also conducted in order to minimise this risk.

This review also limited papers to English; it is apparent from a previous review that some Chinese studies may have met other eligibility criteria. However, the authors of this review (Semple et al., 2013) acknowledged that the inclusion of these papers may have limited the generalisability of results due to the diversity in culture regarding health and emotional wellbeing. Therefore, whilst it may have been more complete to include other languages, it may also have further added to the heterogeneity of studies and results.

Lastly, the findings of this review are based on a small number of studies – which limits confidence in their generalisability (overall estimates to date will be highly sensitive to new data). This is a limitation that reflects the state of available evidence, and one that could be readily addressed through further empirical investigation.
Implications for practice and research

This review is in agreement with other systematic reviews (Luckett et al., 2011; Semple et al., 2013) identified within this area which also conclude that at present there is limited-to-no evidence supporting psychosocial/psychological interventions for improving quality of life for patients with head and neck cancer. This reflects the limited evidence available within this population, especially when compared with reviews of other cancer types. For example, a review into breast cancer and psychological interventions identified 28 randomised controlled trials spanning 3940 participants (Jassim, Whitford, Hickey, & Carter, 2015). Moreover, the evidence from Jassim et al. (2015) indicated that one comprehensive psychological intervention – CBT – had favourable effects for both reducing distress and improving quality of life in breast cancer patients (although evidence for the latter was notably weaker). Extrapolating from Jassim et al. (2015), we might expect CBT to demonstrate similar efficacy in head and neck cancer once investigated to a similar extent. It may also be the case (based on the differential strengths of evidence in Jassim et al., 2015) that CBT is more apt to reduce distress versus improve quality of life in head and neck cancer – and distress reduction is arguably the more direct (proximal) target within the CBT model (Gaudiano, 2008). However, the meta-analytic findings of Semple et al. (2013) did not support this distinction: finding no evidence to favour psychological interventions for improving distress outcomes in head and neck cancer – although available studies were limited in number and quality (e.g., failure to screen for clinical distress during recruitment, potentiating suppressive ‘floor effects’ on outcomes; Linden & Girgis, 2012).

NICE guidance state the need for consideration of psychological effects of head and neck cancer and the need for psychological interventions; but also identify that research into this area is lacking (NICE, 2004). Future studies would benefit from improved reporting; on occasion, information regarding intervention content, training of the delivering professional, and adherence and integrity of the intervention was very limited. Extractable data was also insufficient in some studies, thus precluding the pooling of data and meta-analysis. Consistent outcome measurement times, using well-validated outcome measures is necessary for future research to be reviewed and synthesised to develop this evidence base. Furthermore, follow-up time points are of particular importance as some studies have suggested that this is where the main differences may be observed. Studies now need to develop past pilot and feasibility studies to utilise larger sample sizes based on power calculations and to recruit from multiple centres to achieve these numbers. Studies including clearly defined control groups, specifying allocation and blinding are also imperative in improving the quality of evidence within this area.

Further evidence would also be beneficial in investigating differential effects between psychological interventions and psychoeducation/supportive care. As demonstrated by Kangas et al. (2013), interventions which rely upon psychoeducation and emotional support from professionals (“non-directive supportive counselling” – control group) can offer positive effects in reducing psychological distress and improving quality of life. However, effects from psychological interventions appeared more clinically meaningful and enduring. This research needs to be developed further. Despite the currently limited evidence-base for psychological interventions for this population, the broader evidence-base for the efficacy of CBT (Jassim et al., 2015; Moghaddam & Dawson, 2016) suggests that CBT-based interventions warrant continued investigation, with suitably rigorous study designs. As applied to head and neck cancer, studies to date suggest that there may be difficulties recruiting/retaining participants, which may reflect on the burdensomeness of study procedures, but
could also have implications for optimising the acceptability of CBT content/delivery for this population – ensuring that suitable feasibility or internal pilot work is undertaken prior to trialling effectiveness (there may be important population-specific factors that are limiting intervention engagement and effectiveness in studies to date). Alongside continuing investigation of CBT, there are promising new developments within this field, based on alternative theoretical models in the context of cancer (Hulbert-Williams, Storey, & Wilson, 2015), which could also be explored.

References


