Prevalence of depression in adults with cancer: a systematic review

J. Walker1*, C. Holm Hansen2, P. Martin2, A. Sawhney2, P. Thekkumpurath2, C. Beale2, S. Symeonides3, L. Wall3, G. Murray4 & M. Sharpe1

1Psychological Medicine Research Department of Psychiatry, University of Oxford, Oxford; 2Psychological Medicine Research, Edinburgh Cancer Research Centre, University of Edinburgh; 3Edinburgh Cancer Centre, Western General Hospital; 4Centre for Population Health Sciences, University of Edinburgh, Edinburgh, UK

Received 31 October 2011; revised 27 April 2012 & 11 September 2012; accepted 28 September 2012

Background: Depression has substantial effects on cancer patients’ quality of life. Estimates of its prevalence vary widely. We aimed to systematically review published studies to obtain the best estimate of the prevalence of depression in clinically meaningful subgroups of cancer patients.

Design: Systematic review that addressed the limitations of previous reviews by (i) including only studies that used diagnostic interviews; (ii) including only studies that met basic quality criteria (random or consecutive sampling, ≥ 70% response rate, clear definition of depression caseness, sample size ≥ 100); (iii) grouping studies into clinically meaningful subgroups; (iv) describing the effect on prevalence estimates of different methods of diagnosing depression.

Results: Of 66 relevant studies, only 15 (23%) met quality criteria. The estimated prevalence of depression in the defined subgroups was as follows: 5% to 16% in outpatients, 4% to 14% in inpatients, 4% to 11% in mixed outpatient and inpatient samples and 7% to 49% in palliative care. Studies which used expert interviewers (psychiatrists or clinical psychologists) reported lower prevalence estimates.

Conclusions: Of the large number of relevant studies, few met our inclusion criteria, and prevalence estimates are consequently imprecise. We propose that future studies should be designed to meet basic quality criteria and employ expert interviewers.

Key words: cancer, depression, prevalence, review, systematic

At first glance, this information appears easy to come by. The briefest of electronic searches reveals that there are hundreds of articles that might be relevant, reflecting the importance of and interest in the topic. A number of reviews, including systematic reviews and meta-analyses have also been published [10–22]. However, a closer inspection of these reviews reveals that, despite the wealth of research publications summarised, the prevalence of depression in clinically meaningful subgroups of people with cancer remains unclear with widely varying estimates that are difficult to apply clinically. This is because the published reviews have been limited by one or more of the following problems. The first problem is the inclusion of studies that have not used diagnostic interviews to assess whether participants were depressed. The most commonly used diagnostic criteria in psychiatry are those of the Diagnostic and Statistical Manual of Mental Disorders (DSM) and the International Classification of Diseases (ICD) [23, 24]. These diagnostic criteria describe the nature, severity and duration of symptoms required to make an interview-based diagnosis of depression. Although rating scales can be used to identify patients who require...
further assessment (e.g. as a first stage in screening), or to monitor the course of diagnosed depression, they cannot be used to diagnose depression. The second problem is the inclusion of studies of varying methodological quality. Lack of attention to study quality is a common problem in reviews of observational studies and is important because the results of low-quality studies are likely to be biased and therefore to provide misleading estimates [25]. The third problem concerns the pooling of data from heterogeneous samples into one overall estimate of depression prevalence for all cancer patients. This strategy makes the questionable assumptions that the prevalence is the same in different patient subgroups and that a pooled estimate is clinically meaningful. The final problem is a failure to consider the effect of different methods of diagnosing depression on prevalence estimates. Diagnostic criteria can be applied using different interview schedules administered by people with a range of expertise. In addition, the most commonly used diagnostic criteria include a number of physical symptoms which may also arise from having cancer or cancer treatments. Researchers may decide to apply these criteria without any assessment of the cause of patients' physical symptoms (the ‘inclusive approach’), to exclude symptoms they judge to be cancer-related or to use alternative criteria without physical symptoms.

Although previous reviews have tried to address one or more of these problems, none has addressed all of them. We, therefore, aimed to answer the question, ‘How common is depression in people with cancer?’ by conducting a systematic review of relevant published studies in a way that addressed all the aforementioned problems by (i) including only studies that used diagnostic interviews to determine depression caseness; (ii) including only studies that met basic quality criteria; (iii) grouping studies into clinically meaningful subgroups of people with cancer; (iv) describing the effect on prevalence estimates of different methods of diagnosing depression.

methods

search strategy

We identified relevant published research articles by a systematic search of the following electronic databases conducted in January 2012: Medline (1950 to 2012), PsycINFO (1806 to 2012), EMBASE Classic+EMBASE (1947 to 2012), Web of Science and BIOSIS. Searches were run for the combination of ‘prevalence’, ‘cancer’ and ‘depression’, using both standardised subject terms and free text terms, including synonyms and alternative spellings. Full details of the searches used are given in the supplementary appendix, available at Annals of Oncology online. We also manually searched the reference lists of all the study reports selected for inclusion in the review and of review articles obtained through the electronic searches.

selection criteria

We judged studies to be relevant if they met all the following criteria: (i) the study clearly aimed to estimate the prevalence of depression (i.e. studies that were designed to address a different research question but happened to include a prevalence estimate, such as clinical trials or questionnaire validation studies, were not included); (ii) all study participants were adults (aged 18 or older); (iii) all study participants (or a clearly defined subgroup for which there was an estimate of depression prevalence) had a definite cancer diagnosis (e.g. histological diagnosis or attending for cancer treatment); (iv) depression caseness was determined using diagnostic interviews.

We included only primary studies (i.e. not reviews) for which we could obtain the full paper for data extraction. We also applied quality criteria to the study methods. To ensure a consistent and transparent approach to quality assessment, we developed and used a checklist of specific inclusion criteria, informed by the work of Loney et al., rather than using a continuous quality score [26–29]. We included only studies that met all of the following criteria which we considered to be basic and relatively undemanding markers of quality: (i) the study sample was obtained using a random or consecutive sampling method; (ii) data were available for analysis on at least 70% of the eligible patients (either as reported by the authors or derived from presented data); (iii) depression caseness was defined using standard diagnostic criteria, for example major depression from the DSM or depressive episode from the ICD [23, 24]; (iv) at least 100 study participants were assessed for depression. The first two of these criteria relate to the minimisation of selection bias, where participants are not representative of the target population. The third criterion was included to ensure that estimates could be compared across studies. The final (sample size) criterion was included because a small sample, although not in itself an indicator of bias, is prone to result in chance and inaccurate findings and, when reported individually, can be misleading. An error rate of ±5% has, therefore, been recommended for prevalence studies [29]. We included studies of 100 participants or more, thereby striking a balance between having a reasonable number of studies in our own review and including studies of reasonable accuracy (100 participants are required to estimate an expected prevalence of 10% with a <6% error rate).

data collection and analysis

We imported all articles identified by the searches into a database and screened their titles and abstracts to determine whether each might meet the selection criteria. We reviewed the full text of the article, with the help of a translator where necessary, if there was any possibility that it might be relevant. This process was conducted independently by two researchers and a decision whether to include each study was made by consensus. Two researchers independently extracted data from all the articles judged relevant, using a specially designed, standardised data extraction form, and with the help of a translator for non-English papers. We extracted data on study setting and design, number and characteristics of participants included, method of depression diagnosis (including interviewer expertise, diagnostic criteria used and how these were applied) and reported the prevalence of depression in the sample. For cohort studies, we reported depression prevalence at the first time point only. If studies assessed both current and previous depression, we reported current depression prevalence. If studies compared the prevalence of depression in patients with cancer with that in a non-cancer population, we used only the prevalence in the cancer sample. We assessed the methodological quality of each study at the same time as data extraction. In order to provide subgroups that might be meaningful to clinicians, we aimed to group studies by clinical setting (outpatient, inpatient, palliative care) and further subdivide these by primary cancer site and disease stage. For each subgroup, we reported the prevalence of depression found in the primary studies and considered whether meta-analysis was appropriate for statistical synthesis of their findings. We also considered whether the data were adequate to test the hypothesis that depression prevalence is the same across cancer subgroups. We used a forest plot to display the results of the primary studies graphically. In order to describe the effect of different methods of diagnosing depression, we described the diagnostic criteria that
were used in each study and how these were applied as well as the interview schedules employed and the interviewers’ expertise.

results

literature overview

Searches and initial screening of titles and abstracts yielded 499 potentially eligible studies, of which 42 required translation. After reviewing the full articles, 433 were considered not relevant (the most common reason was that the study was not designed to estimate depression prevalence or depression was not diagnosed using an interview), leaving 66 relevant studies. Of these, 15 (23%) met our quality criteria and were included [4, 30–43]; 12 of the included studies had a cross-sectional design and 3 were prospective cohort studies. Sample sizes ranged from 100 to 3938 (median 129, mean 399). Full details are shown in the supplementary flowchart, available at *Annals of Oncology* online.

quality of studies

During data extraction, we noted that reporting was often unclear, both for the methods used and results obtained; few of the publications had adhered to current reporting guidelines [44]. Of the 15 studies we included, only 2 reported confidence intervals for their estimate of depression prevalence.

characteristics of study participants

The studies had been carried out in seven different countries and included samples from a range of settings; the majority were in specialist cancer or palliative care services. Five studies focused on a specific primary cancer (e.g. only two included patients with breast cancer) and the remaining 10 included patients with various cancer types. Further details are given in the supplementary tables, available at *Annals of Oncology* online.

depression diagnosis

Thirteen studies used the diagnostic criteria for major depression from DSM to define caseness, and the other two used the ICD criteria. Four studies reported that they had taken an inclusive approach to diagnosis (including all potential symptoms of depression without attempting to judge what might have caused them); one noted that the interviewer had not judged whether patients were impaired by their symptoms; two had excluded all symptoms that the interviewer judged to be related to the patient’s cancer or its treatments and one noted that ‘attention had been paid’ to physical symptoms. The other seven studies did not specify how they had applied the diagnostic criteria. One study also reported the number of participants who met the Endicott diagnostic criteria for depression in which physical symptoms are substituted by other criteria [45].

The most commonly used interview schedule was the Structured Clinical Interview for DSM-IV (SCID), some form of which was used by 10 studies to diagnose major depression, including one study that used a two-stage procedure to identify depressed participants [46]; 14 of the 15 studies reported the interviewers’ professions; 10 studies had employed a psychiatrist or clinical psychologist to conduct the diagnostic interviews and the other 4 had used research assistants, research nurses, oncologists and students. References were made to interviewer training in a number of the publications, but there was little description of what the training had involved or how the interviewers (especially those who were not mental health professionals) had been deemed competent.

prevalence of depression

Figure 1 shows the prevalence of depression in each study, together with our calculated confidence intervals, grouped by clinical setting. We were unable to further subdivide the studies by primary cancer site and disease stage due to the small number of studies available. We had considered using meta-analysis to combine the prevalence estimates within subgroups, but in the light of the degree of clinical heterogeneity (e.g. cancer sites and treatment being received) and variety of methods of diagnosing depression, we considered meta-analysis to be inappropriate. There were also insufficient data to test the hypothesis that depression prevalence is equal across cancer subgroups.

outpatients

Six studies were of cancer outpatients and reported a current depression prevalence ranging from 5% to 16%. Two of these assessed women with breast cancer for major depression and reported prevalence of 9% and 16%. In the first study, a clinical psychologist administered the SCID to women with all stages of disease, whereas in the latter a research nurse used the Mini Neuropsychiatric Interview to assess women whose cancer was in remission. One study was of patients 1 month after surgery for lung cancer and found a depression prevalence of 5% using the SCID administered by psychiatrists. The other three studies included patients with various primary cancer sites attending for outpatient treatment or follow-up: one used trained researchers to administer the interview, the second used a psychiatrist and the third used untrained oncologists; they reported prevalence of 8%, 8% and 12%, respectively.

inpatients

Three studies used psychiatrists to assess cancer inpatients for major depression, using the SCID: one study focused on patients with head and neck cancer and found that 4% of newly diagnosed patients admitted for initial treatment had major depression. The other two were of patients with various cancer types; one reported a current depression prevalence of 14%, while the second reported that 30% of participants had had major depression at some time in the previous 12 months.

outpatients and inpatients

Three studies included both inpatients and outpatients. One used a psychiatrist-administered SCID to assess patients with unresectable lung cancer and reported a prevalence of 5%. The other two assessed patients with various cancer types; one of these reported prevalence of 4% for current depression as assessed by a clinical psychologist or psychiatrist and the other reported that 11% of participants had major depression as assessed by a psychiatrist.
Three studies were of patients with various cancer types who had been referred for palliative care: one reported that 49% of patients had major depression in interviews administered by trained psychology students, a second found a prevalence of 7% using psychiatrist-delivered interviews and a third reported that 22% of patients had depression using 'trained interviewers'.

**effect on prevalence estimates of different methods of diagnosing depression**

**diagnostic criteria**

Cases of depression were defined using DSM and ICD criteria in all the studies we included. Some attempted to address the concern that cancer-related physical symptoms might result in an overestimate of depression prevalence by judging the cause of these symptoms in individual patients. We were not able to discern a clear pattern in prevalence estimates based on these approaches (i.e. it was not clear that using an inclusive approach led to higher prevalence estimates within the subgroups). One study reported two prevalence estimates, one using DSM criteria for major depression and a much lower prevalence using the Endicott criteria when these were applied by the same interviewers.

**interview schedules**

In the 'outpatients' subgroup, two studies used the SCID alone, one used a two-stage process that included the SCID and the other three studies each used a different interview schedule. While the small number of studies makes comparisons difficult, those using the SCID tended to report lower depression prevalence (range 5% to 9%) compared with those that used other interviews (range 8% to 16%). However, it should be noted that two of the SCID-based studies used expert interviewers.

**interviewers’ expertise**

The 'outpatients' and 'palliative care' clinical subgroups included studies that used expert interviewers (psychiatrists or clinical psychologists) as well as those that used less expert interviewers. In outpatients, the range of depression prevalence was 5% to 9% when assessed by experts and 8% to 16% when assessed by less expert interviewers. In palliative care settings, there were fewer studies but the one study that used expert assessments reported a prevalence of 7% compared with 22% and 49% in the other two studies.

**discussion**

**main findings**

Of 66 relevant studies, it was notable that only 15 (23%) met our basic quality criteria for inclusion in this review. The reporting of both study methods and results was often unclear, even in the studies we included, and few publications had adhered to current reporting guidelines. There were too few comparable studies in the subgroups for us to conduct meta-analyses that would be likely to yield meaningful results. The studies we reviewed reported current depression prevalence estimates ranging from 5% to 16% in outpatients, 4% to 14% in inpatients, 4% to 11% in mixed outpatient and inpatient samples and 7% to 49% in palliative care. Despite the fact that all studies used standard diagnostic criteria to define depression caseness, those in which an expert (psychiatrist or clinical psychologist) administered interviews reported a lower estimate of current depression prevalence than studies that employed less expert interviewers.
previous reviews
Our finding that there are a large number of publications reporting the prevalence of depression in patients with cancer, but that few of these reported studies met basic quality criteria, echoes observations made in previous systematic reviews and their accompanying editorials [47]. To the best of our knowledge, our review is the first to use a systematic checklist of transparent quality criteria to determine which studies should be included in the review. We were unable to find any consistent pattern in prevalence estimates based on the approaches that researchers had taken to physical symptoms. This is consistent with other work that suggests that the effect of such modifications to diagnostic criteria may be smaller than previously thought [48, 49]. It is noteworthy that a recent meta-analysis of all interview-based studies found a pooled prevalence of ∼16%, which is substantially higher than the estimates we obtained from the subset of studies in which an expert had administered the interviews [10]. Wasteson et al. [50] have previously commented that lack of consistent measurement and definition of depression in people with cancer makes comparisons of prevalence estimates problematic and that the range of experience and training of interviewers employed in prevalence studies adds to this problem.

strengths and limitations
This review has a number of strengths. We searched for articles systematically and included studies using clearly defined criteria to minimise selection bias. We also judged studies’ methodological quality and excluded those that had specific design flaws rather than merely assigning quality scores, thereby maximising the transparency and reproducibility of the review [51, 52]. Our review also has limitations. First, while our quality criteria were based on relevant literature, guidelines for assessing the quality of observational studies are less well-defined than those for clinical trials, making some of the criteria we used, such as sample size, relatively arbitrary. However, we believe that our quality criteria may be considered a reasonable minimum and were not over stringent (e.g. a sample of 100 participants is lower than that recommended). Second, our assessments of studies, both for relevance and quality, were based on the information available in the published reports. This may mean that we excluded studies that were well conducted but simply poorly reported. However, even if this were the case, it would be unlikely to explain much of the methodological shortcomings apparent in the reports. Third, we grouped studies using clinically relevant settings, as patients attending outpatient clinics are likely to be less unwell than inpatients and those receiving palliative care. However, these groups are not homogenous, and the wide range of depression prevalence estimates in the palliative care setting, for example, may reflect the contributions of cultural and treatment factors to depression causation that we were unable to study. Finally, we did not attempt to include grey literature by contacting relevant experts for unpublished manuscripts. However, we did make substantial efforts to find all relevant published studies through inclusive search strategies and the use of translators.

implications
Depression is an important and potentially fatal but treatable complication of cancer. As well as having substantial effects on quality of life, depression contributes to non-adherence to medical treatments [53]. In order to plan effective services, we need accurate estimates of its prevalence in clinically meaningful subgroups of cancer patients.

Despite a large number of relevant publications, it was striking that few studies met our basic quality criteria and we, therefore, currently lack adequately precise and useful data on the prevalence of depression in clinically relevant subgroups of cancer patients. Studies that used expert interviewers to diagnose depression were more consistent in their findings and reported a lower prevalence than studies that used less expert interviewers. Finally, we wish to make a plea for an improvement in the quality of the research published in this area and suggest that the quality criteria used in this review, along with the use of expert interviewers, are a prerequisite for the funding and publication of future studies on this topic.

acknowledgements
The authors would like to acknowledge the support of Anne Byrne and Marshall Dozier.

funding
This work was supported by Cancer Research UK (grant number C5547/A7375).

disclosure
The authors have declared no conflicts of interest.

references