Published in final edited form as:

Psychooncology. 2010 March ; 19(3): 313–317. doi:10.1002/pon.1561.

Published versus Unpublished Dissertations in Psycho-Oncology Intervention Research

Anne Moyer, Stefan Schneider, Sarah K. Knapp-Oliver, and Stephanie J. Sohl
Stony Brook University

Abstract

Objective—There are conflicting views regarding whether gray literature, including unpublished doctoral dissertations, should be included in systematic reviews and meta-analyses. Although publication status frequently is used as a proxy for study quality, some research suggests that dissertations are often of superior quality to published studies.

Methods—We examined 107 projects involving doctoral dissertations (42 published, 65 unpublished) that studied psychosocial interventions for cancer patients.

Results—Published dissertations were more likely to be supported by research funding but were not more likely than unpublished dissertations to examine specific types of interventions. Across several indices of methodological quality there were minimal differences. Dissertations with significant findings tended to be more likely to be published than those without significant findings.

Conclusions—Unpublished dissertations focusing on psychosocial interventions for cancer patients are not necessarily of vastly inferior quality to those that eventually are published. Because doctoral dissertations are easy to access relative to other forms of gray literature, are free from some types of bias, and are reported thoroughly, they merit inclusion in comprehensive literature reviews.

Key terms

dissertations; cancer; oncology; psychosocial; intervention; meta-analysis

There are conflicting opinions about whether gray literature, including unpublished doctoral dissertations, should be included in systematic reviews and meta-analyses [1]. On the one hand, doctoral dissertations are not peer reviewed in the same fashion as published journal articles; on the other hand, they are subjected to intense scrutiny by dissertation committees, and their reporting is typically quite thorough. Often publication status is used as a convenient proxy for methodological quality [2] and this practice may have some merit. However, even peer-reviewed journals differ in their methodological rigor, and published articles vary in their reporting quality due to differences in journal space limitations. Finally, some argue for the high quality of doctoral dissertations, asserting, “often the culmination of months or years of work, the dissertation is the concrete manifestation of a doctoral student’s best thinking, guided and refined by the sage suggestions of the student’s committee members” [3, p. 537].

The small number of empirical studies evaluating the quality of gray literature has generally attested that dissertations actually have stronger methodology than published studies [4] or

that there is limited evidence that trials in the gray literature are of lower quality than published trials [5]. However, it is unclear whether these conclusions apply to studies in psycho-oncology research. Because searching for and retrieving gray literature, including dissertations, is time-consuming [6], decisions to include them in reviews are complex and should be based on empirical evidence [7].

In a prior, larger review of 488 published and unpublished projects investigating psychosocial interventions for cancer patients we characterized the methodological quality of this literature as a whole and focused on trends over time (Moyer, Sohl, Knapp-Oliver, Schneider, in press). We found that particular strengths of this research, present in a majority of projects, included using randomized designs, testing for baseline group equivalence, and monitoring treatment, the use of which rose significantly over time. In that review we identified a subset of 107 projects that included at least one doctoral dissertation, suggesting that this is a fruitful area for budding investigators. Sixty-five of these projects, however, included dissertations that remained unpublished, prompting the question of whether they were of lower quality than projects with published results.

In this review we focus on the characteristics and methodological quality of just these 107 published and unpublished dissertations. Because prior reviews have contrasted unpublished dissertations with published studies in general, they have not explored the potential for publication bias resulting from only some dissertations making their way into the published literature. Thus, this review, in comparing only published versus unpublished research originating from doctoral dissertations, may shed some light on reasons for differential publication.

It may also be that publication was more likely when the dissertation reported significant findings. This “publication bias” is a well-documented problem [8] and can result from authors selectively submitting, or from journals selectively accepting, those manuscripts that show significant results [9]. Thus, an additional goal of this review was to examine whether dissertations that reported significant findings were more likely to be published.

Method

Study Identification

Studies included in the larger review (Moyer, Sohl, Knapp-Oliver, Schneider, in press) examined psychosocial interventions for adult cancer patients that: (1) reported outcomes on psychological, emotional, behavioral, physiological, functional, or medical status; (2) were first reported as a published article or dissertation between January 1980 and December 2005; and (3) included 10 or more individuals per group. Electronic databases (PsycINFO, which indexes dissertations; PubMed; and Dissertation Abstracts Online, which indexes doctoral theses from 1861 to the present from U.S., Canadian, British, and European universities [10] were searched using key terms (e.g., *cancer*, *neoplasms*, *tumor*, and *psychosocial intervention*, *psychotherapy*, *psychological treatment*, *education*, *cognitive behavioral*, *relaxation*, *stress management*, *support group*, *self-help group*, *nursing intervention*, *biofeedback*). The reference lists of included reports and of 94 prior reviews and meta-analyses also were examined. Searches for articles citing prior reviews and the tables of contents of several journals (*Psycho-Oncology*, *Journal of Clinical Oncology*, *Cancer*, *Journal of Psychosocial Oncology*, *European Journal of Cancer*, and *Cancer Nursing*) were conducted.

Separate reports based on the same sample (e.g., separate articles reporting outcomes at 3-month and 12-month follow-ups) were consolidated as being from the same project. The larger sample included 673 reports comprising 488 projects, 107 of which included

dissertations. As some projects yielded multiple dissertations (and some dissertations reported on more than one project), there were 112 dissertations conducted across the 107 projects. However, our analyses were at the project level.

A follow-up search of the PubMed and PsycINFO bibliographic databases determined if unpublished dissertations had been published by November, 2008. For projects that yielded more than one dissertation, if at least one was published, we coded the project as published. In all, 42 projects comprised (at least one) dissertation that had been published and 65 projects comprised (at least one) dissertation had not been published; for simplicity, we refer to these as *published dissertations* and *unpublished dissertations*, respectively.

Study Coding

Coding of descriptive project characteristics involved the types of interventions that had been investigated and whether the project had received research funding. Because some authors have noted difficulties in tracking dissertations when female investigators have married and changed their names [6], we also examined whether our coding of publication status was linked to the gender of the principal investigator. Coding items assessing the quality of study methodology and reporting were adapted from prior work [11]. Although consensus on essential areas of methodological quality has yet to be reached, and no one scale is considered appropriate for all research topic areas [12] we included aspects of quality conventionally considered important. These involved aspects of the sample description; the research design, including the quality of randomization, where applicable; intervention specification and provision; and data analyses, such as whether intent-to-treat analyses were conducted. Due to the low feasibility of keeping participants and interventionists blind to treatment groups for psychosocial interventions, items assessing this were not included. Similarly, because outcomes in this area are predominantly based on self-report, assessments of blinding of outcome assessors also were not included. Aspects of reporting from the CONSORT [13], such as noting the number dropping out of treatment, also were evaluated. Combining elements related to different dimensions of methodological quality is not advised because they are theoretically independent and may be negatively related [2]. Therefore, we report these elements of methodological quality separately. Because prior research has shown that studies in the published literature have more participants than studies in the gray literature [5], we also examined sample size. Finally, to address the question of publication bias, we coded whether or not a project reported at least one significant intervention effect.

Coding according to a detailed manual was conducted by the PI and two teams of thoroughly-trained graduate-level coders. Coders met regularly to prevent coding drift, discuss coding dilemmas, and to reach consensus on independently-coded projects (representing 9.2% of the total sample) used for reliability estimation. Ten key continuous a priori coding items were examined for reliability. The agreement for the ratings of the PI, Coder 1, and Coder 2 was .83 and for the ratings of the PI, Coder 3, Coder 4, and Coder 5 was .90 (average two-way mixed effect intraclass correlation [14]). Ten key categorical a priori coding items were also examined. The agreement for the ratings of the PI, Coder 1, and Coder 2 was .72 and for the ratings of the PI, Coders 3, Coder 4, and Coder 5 was .61 (average generalized kappa [15]).

Statistical Analyses

For continuous outcome variables, we used *t*-tests to compare published and unpublished dissertations. For categorical outcome variables, we used chi-square tests and Fisher's exact test when cells had an expected count of less than 5. A Bonferroni-corrected alpha level of 0.002 (0.05/26) was adopted to account for the multiple tests of significance conducted.

Results

Investigator, Report, and Intervention Characteristics

Published dissertations were significantly more likely to have had some form of research funding (73.7%) than unpublished dissertations (26.3%, $\chi^2 [1, N = 107] = 29.30, p = 0.000$). The proportion of female principal investigators was somewhat greater (85.7%) for published dissertations than for unpublished dissertations (67.2%, $\chi^2 [1, N = 106] = 4.59, p = 0.032$), but this difference was not significant at the corrected alpha level. There were no significant differences in the likelihood of published versus unpublished dissertations' examining cognitive, behavioral, or cognitive-behavioral interventions (47.6% vs. 47.7%, $\chi^2 [1, N = 107] = 0.00, p = 0.994$); non-behavioral counseling or psychotherapy interventions, (4.8% vs. 13.8%, Fisher's Exact $[N = 107] p = 0.195$); educational or informational interventions, (19.0% vs. 10.8%, $\chi^2 [1, N = 107] = 1.45, p = 0.228$); social support interventions, (4.8% vs. 4.6%, Fisher's Exact $[N = 107] p = 1.00$); multimodal interventions, (19.0% vs. 26.2%, $\chi^2 [1, N = 107] = 0.72, p = 0.396$); or interventions using complementary or alternative medicine approaches, (21.4% vs. 9.2%, $\chi^2 [1, N = 107] = 3.15, p = 0.076$).

Quality of Study Methodology and Reporting

Table 1 compares the methodological and reporting quality of published versus unpublished dissertations. Of 16 statistical comparisons, there were no significant differences between published and unpublished dissertations. Although published dissertations were somewhat more likely to use randomized as opposed to other types of designs (83.3% versus 61.5%, $\chi^2 [1, N = 107] = 5.78, p = 0.016$), this trend was not significant. In addition, published dissertations did not have significantly larger numbers of initial participants per group ($M = 37.70, SD = 42.46$) than unpublished dissertations ($M = 29.39, SD = 23.94, t[105] = 1.29, p = 0.199$). Finally, there was a trend for a larger proportion of published dissertations to report significant findings than unpublished dissertations, (100.0% versus 84.6%, Fisher's Exact $[N = 107] = 0.006$), although it did not reach significance at the corrected alpha level.

Discussion

We examined 107 projects comprising 42 published and 65 unpublished dissertations examining psychosocial interventions for cancer patients, comparing their characteristics and methodological quality. Published versus unpublished dissertations were not more likely to examine different types of interventions, but were more likely to be supported by research funding. On several indices of methodological quality, differences were minimal. However, there was a trend for dissertations in this area to exhibit publication bias; whereas every published dissertation reported a significant finding, only a proportion of unpublished dissertations reported one. Thus, excluding unpublished dissertations from systematic reviews may ignore null-results from methodologically sound research.

It was not clear whether having research support for the study was linked to the likelihood of using a rigorous randomized design, which may lead to a more publishable study, or if having funding is linked to publication status independently, perhaps by producing an imperative to persevere in publishing a study. We found no association between having funding and using randomized designs ($\chi^2 [1, N = 107] = 0.36, p = 0.55$), suggesting that being awarded funding may encourage authors to pursue publication, perhaps through pressure from funding agencies that expect timely publication of projects and use this in decisions for future funding.

Limitations of this analysis include the fact that our search strategy involved only three bibliographic databases, although additional channels of study identification were pursued. Although Dissertation Abstract Online indexes dissertations conducted at most North

American, UK, and many European universities, coverage of other countries is limited. Consequently, some existing dissertations likely remained unidentified.

Findings support the notion that unpublished doctoral dissertations, at least in this area of research, are not necessarily of vastly inferior quality. Because doctoral dissertations are easy to identify in bibliographic databases, are not subject to publication-review bias and the file-drawer effect due to automatic entry into the Dissertation Abstracts International database, can often be obtained via institutions' interlibrary loan services (or purchased commercially), and have thorough reporting, we concur with other authors [4,6] that they merit inclusion in comprehensive literature reviews.

Acknowledgments

This work was supported by a grant from the National Cancer Institute (R01 CA100810) to Anne Moyer. We are grateful to John Finney for helpful feedback on a prior version of this manuscript.

References

1. McAuley L, Pham B, Tugwell P, Moher D. Does the inclusion of grey literature influence estimates of intervention effectiveness reported in meta-analyses? *Lancet*. 2000; 356:1228–1231. [PubMed: 11072941]
2. Conn VS, Valentine JC, Cooper HM, Rantz MJ. Grey literature in meta-analyses. *Nurs Res*. 2003; 52:256–261. [PubMed: 12867783]
3. Conn VS. The light under the bushel basket: unpublished dissertations. *West J Nurs Res*. 2008; 30:537–538. [PubMed: 18612094]
4. McLeod BD, Weisz JR. Using dissertations to examine potential bias in child and adolescent clinical trials. *J Cons Clin Psychol*. 2004; 72:235–251.
5. Hopewell S, McDonald S, Clarke M, Egger M. Grey literature in meta-analyses of randomized trials of health care interventions. *Cochrane Database Syst Rev*. 2007:MR000010. [PubMed: 17443631]
6. Vickers AJ, Smith C. Incorporating data from dissertations in systematic reviews. *Int J Tech Asses Health Care*. 2000; 16:711–713.
7. Benzie KM, Premji S, Hayden KA, Serrett K. State-of-the-evidence reviews: advantages and challenges of including grey literature. *Worldviews Evid Based Nurs*. 2006; 3:55–61. [PubMed: 17040510]
8. Dubben HH, Beck-Bornholdt HP. Systematic review of publication bias in studies on publication bias. *BMJ*. 2005; 331:433–434. [PubMed: 15937056]
9. Kraemer HC, Gardner C, Brooks JO, Yesavage JA. Advantages of excluding underpowered studies in meta-analysis: Inclusionist versus exclusionist viewpoints. *Psychol Methods*. 1998; 3:23–31.
10. ProQuest Information and Learning. Dissertation Abstracts Online (Dissertations). 2009. Retrieved February 1, 2009 from <http://newfirstsearch.oclc.org/WebZ/FSHelp?show=Dissertations:entityhdbname=Dissertations:sessionid=fsapp1-57468-fqo4uq16-1e4aws:entitypagenum=2:0:code=&badcode>
11. Moyer A, Finney JW, Swearingen CE. Methodological characteristics and quality of alcohol treatment outcome studies, 1970–98: An expanded evaluation. *Addiction*. 2002; 97:253–263. [PubMed: 11964099]
12. Moyer A, Finney JW. Rating methodological quality: toward improved assessment and investigation. *Account Res*. 2005; 12:299–313. [PubMed: 16578923]
13. Moher D, Schulz KF, Altman D. The CONSORT statement: revised recommendations for improving the quality of reports of parallel-group randomized trials. *JAMA*. 2001; 285:1987–1991. [PubMed: 11308435]
14. Shrout PE, Fleiss JL. Intraclass correlation: Uses in assessing rater reliability. *Psychol Bulletin*. 1979; 86:420–428.

15. Siegel, S.; Castellan, NJ. Nonparametric statistics for the behavioral sciences. 2. New York: McGraw Hill; 1988.

Table 1

Quality of Study Methodology and Reporting

	Published dissertations (<i>n</i> = 42)	Unpublished dissertations (<i>n</i> = 65)	χ^2	<i>p</i>
	%	%		
Sample description				
Reported number approached to participate	45.2	44.6	0.004	0.950
Reported number initially participating	90.5	90.8	F.E.	1.000
Compared characteristics of participants to eligible non- participants	6.2	14.3	F.E.	0.186
Reported number dropping out of treatments	50.0	44.6	0.297	0.586
Research Design ^{ab}			5.782	0.016
One group pre and post test	11.9	13.8		
Nonequivalent control group without matching or statistical control	0.0	21.5		
Nonequivalent control group with matching or statistical control	2.4	0.0		
Randomized experiment	83.3	61.5		
Other design/design not indicated	2.5	3.0		
Quality of randomization, for randomized experiments (<i>N</i> = 75) ^c			F.E.	0.528
Method only stated to be randomized	48.6	50.0		
Randomization procedure described but no measures to prevent subterfuge included	40.0	32.5		
Randomization procedure described and measures to prevent subterfuge included	11.4	17.5		
Intervention provision and specification				
Manuals used to guide treatment (where applicable, <i>N</i> = 140 applicable intervention conditions)	46.4	44.0	0.077	0.782
Intervention implementation assessed ^d				
Intervention monitored	66.7	49.2	3.149	0.076
Assessed immediate effects of intervention	9.5	13.8	F.E.	0.561
Additional intervention monitored	26.2	24.6	0.034	0.855
Contamination monitored (where applicable, <i>N</i> = 90)	8.1	7.5	F.E.	1.000
Process analyses conducted ^d				
Linked intervention elements or duration to outcome	31.0	16.9	2.886	0.089
Examined mediating factors	19.0	13.8	0.517	0.472
Data Analyses				
Groups compared for equivalence at baseline(where applicable, <i>N</i> = 93)	81.1	85.7	0.354	0.552
Interaction between treatment condition and participant baseline characteristics in relation to dropout status (where applicable <i>N</i> = 87)	2.8	0.0	F.E.	0.414
Intention-to-treat analyses conducted	11.9	6.2	F.E.	0.311

Note. All analyses χ^2 (1, *N* = 107), except where *N* is indicated otherwise. F.E. = Fisher's Exact Test.

^aCategories do not add to 100.0% due to rounding.

^b Comparison between randomized versus all other types of design.

^c Comparison between randomization procedures described and methods to prevent subterfuge included versus all other levels of quality of randomization.

^d Subcategories do not add to 100% because they are not mutually exclusive.