

keynote article

Are most positive findings in health psychology false.... or at least somewhat exaggerated?**James C. Coyne^{*1, 2}**¹ University of Pennsylvania School of Medicine, USA² University of Groningen, the Netherlands

Health psychology has expanding opportunities to influence public health policy and promote patient-oriented health care. Health psychologists are in a position to document the role of potentially modifiable behavior in determining biomedical outcomes, and, beyond this, to advocate that improvements in quality of life and the avoidance of unnecessary pain and distress, not just biomedical outcomes, are important factors in evaluating health care. Health psychology research commands increasingly more resources, even if not all that we desire. Yet, this increased influence and these new resources bring new responsibility to be evidence based in the claims made to the public, including clinicians, health policy makers, and health care consumers. Can we trust the health psychology literature to provide this evidence?

Ioannidis (2005a) provocatively declared that “It can be proven that most claimed research findings are false” (e124). Although the title of a later paper, “Why most discovered true associations are inflated” (Ioannidis, 2008) might seem to indicate a retreat from ‘false’ to merely ‘inflated’, the later paper actually presents a cogent defense of the claims in the earlier paper.

Ioannidis (2008) starts by pointing to the accepted criterion for a “discovery” is a statistical test reaching a $p < .05$ level of significance. This fixation on statistical significance, of course, ignores the many nonsignificant results that will not get published, particularly when studies claiming discoveries were underpowered to begin with. Consistent with this, my colleagues and I were surprised to discover when that we scrutinized published meta analyses of health psychology interventions, a considerable number of the studies entered into these analyses were so underpowered that they had only a probability of .20 or less of detecting a moderate effect (e.g., $\sigma = .50$ based on Cohen, 1992) when it was indeed present (Coyne, Thombs, & Hagedoorn, under review). Yet, we found that a substantially higher percentage of the studies included in the meta analyses reported significant effects, even though the overall estimated effects were well below $\sigma = .50$. Based on this, it is reasonable to assume that there must be a pervasive confirmatory bias in the

**James C. Coyne**Professor of Psychology
Department of Psychiatry
University of Pennsylvania, USADirector, Behavioral Oncology Program
Abramson Cancer Center, USAAdjunct Professor of Health Psychology
Department of Health Sciences
University Medical Center Groningen
University of Groningen, the Netherlands

available published studies. Indeed, Ioannidis (2005b) found that across fields, 25% of the most cited clinical trials and 5/6 of the most cited epidemiological findings were exaggerated or simply false.

How do health psychology intervention studies fare? Preparing for the Society of Behavioral Medicine’s “Great Debate” concerning the efficacy of psychosocial interventions for cancer patients, my colleagues and I discovered that many of the claims by authors that their studies demonstrated that psychosocial interventions reduced distress among cancer patients were also exaggerated or simply false (Coyne, Lepore, & Palmer, 2006). We found that most studies that we reviewed did not restrict recruitment to cancer patients who were sufficiently distressed to register a clinically significant reduction in distress. Given that, claims of benefits had to depend on selective reporting of positive results picked from multiple outcomes, emphasis on unplanned subgroup analysis where main analyses did not reveal an effect, and inappropriate use of ►

*Corresponding Author: James Coyne; email: jcoynester@gmail.com



Are most positive findings in health psychology false.... or at least somewhat exaggerated?

multivariate analysis with too many control variables when main analysis revealed no effect, a practice known to lead to a high rate of spurious results that are not likely to generalize to other samples (Babyak, 2004). Of particular concern, we also found that we could not depend on what was reported in the main outcome papers for particular trials to determine how many or what outcome measures had actually been assessed, what subgroup analyses had been conducted, and what control variables had been considered for multivariate analysis. Comparisons among the papers from the same trial often revealed evidence of phantom outcome measures that appeared to have been banished, perhaps because they contradicted claims that interventions were efficacious, of subgroup analyses had been conducted but suppressed for similar reasons, and of phantom degrees of freedom had been consumed in analyses to pick the control variables that portrayed the efficacy of interventions in the most favorable light. In some cases, it was difficult to determine whether papers came from the same trial!

Yet, it is well known that confirmatory bias is pervasive in the reporting of clinical trials. Ioannidis (2008) notes how selective reporting of analyses and outcomes contributes to inflated estimates of the efficacy of interventions across fields. Is health psychology really any worse than biomedicine? A number of factors suggest that this might be the case. First, there has been much less discussion and documentation of the problem of confirmatory bias in health psychology and much less embracing of what are at least partial solutions. Psychology was slow to adopt empirically based reporting standards, particularly those delineated in CONSORT (Altman, et al., 2001), and in the meantime studies with incomplete reporting of flawed methods and analyses have accumulated (Cook, Palmer, Hoffman, & Coyne, 2007) and continue to dominate the literature (Coyne et al., submitted). Second, health psychology remains a lot more insecure than biomedicine in its sense of legitimacy. A large scale trial that produces null results is disappointing in biomedicine, but it does not threaten the legitimacy of biomedicine in quite the same way that it would health psychology. Indeed, I suspect that the reason why many health psychologists have clung to the claim that psychotherapy extends the life of cancer patients in the utter absence of credible evidence (Coyne, Stefanek, & Palmer, 2007) was that they believed that the credibility of the field would more be generally undermined if the claim about this hard biomedical outcome were shown to be false.

However, we need to be cognizant of the risks of holding onto claims when it becomes apparent that they are unrealistic. Science is self-correcting, even if frustratingly slow, and unrealistic claims will ultimately

be discovered to be such. The first risk in clinging to unsustainable claims is that health psychology will be seen as having weak standards of evidence and a field prone to exaggerated claims. A lost credibility will be difficult to restore. The second risk is that when we are slow to abandon unrealistic claims, other claims that were valid all along come to be seen as an undignified retreat. Thus, psychological interventions can reduce distress and improve quality of life in cancer patients, particularly when the interventions are provided to patients with significant clinical distress (as opposed to all patients, even those who are not distressed). This outcome is important in its own right, but will likely be seen less so if we cling much longer to untenable claims that these interventions can or should extend life. Oncologists already skeptical about the value of psychosocial interventions will likely feel justified in observing that health psychologists made unwarranted claims that they could extend life and now they concede that psychosocial interventions *merely* reduce distress and improve quality of life.

Pressures to claim positive results occur all along the continuum -- from investigators' analysis, reporting, and interpretation of their data to reviewers' evaluation to editors' decisions whether to accept a manuscript to immediate post-publication publicity and, finally, to citations of results in secondary sources such as reviews and practice guidelines. Recently, reports of two clinical trials evaluating the effects of psychological intervention on the survival of cancer patients appeared in print. The first study (Boesen et al., 2007) reported a lack of an effect, despite being well designed and adequately powered. The second (Andersen et al., 2008) claimed that cancer patients receiving psychosocial intervention had fewer recurrences and overall greater survival. The first study received almost no attention in the media, the second received extensive world wide coverage of its claims despite depending on an inappropriately conducted multivariate analysis. As in the first trial, simple bivariate analyses did not reveal that patients receiving the intervention lived longer. We would like to believe that psychosocial intervention can improve both the quantity and the quality of life, but based on the uniform consistency of the results of past studies (Coyne, Thombs, Stefanek, & Palmer, 2009), we should be skeptical on renewed claims of positive effects of psychosocial intervention on survival of cancer patients. As Ioannidis (2005a) clearly demonstrates, when pre-study probabilities of a positive result are low, then we should expect that reports of positive findings have a high probability of being false or exaggerated. ►

Are most positive findings in health psychology false.... or at least somewhat exaggerated?

In summary, we need a fundamental shift in the culture of reporting and publishing of studies, towards more frank and full disclosure of the results. Carrying out well designed trials and honestly reporting their results should become more valued than tortured analyses of data and selective reporting to guarantee positive results. Adherence to CONSORT as a condition of publishing reports of trials is an important first step, but CONSORT only increases the transparency in the reporting of the details of how studies were done, but does not necessarily the quality of the studies or the accuracy with which results are represented. Requiring that clinical trials be registered and that the designs be available on the Web prior to the collection of any data are collected is a practice worth borrowing from the biomedical journals. Being able to cite a prior web-based registered design, including designation of primary outcomes should become a condition for publishing in health psychology journals, just as it has become for the best of biomedical journals since 2004 (DeAngelis et al., 2004). Ioannidis (2008) points out that the success of such measures in eliminating bias in the biomedical journals has been far from complete, but that should not discourage health psychology from adopting them to reduce bias in its literature.

Beyond these recommendations, editors need to encourage the publishing of well designed trials that nonetheless yield null results. Moreover, if negative results are not considered publishable because a study only had 15 or so patients provided with the intervention, then positive results from such an underpowered study should not be considered publishable either. Furthermore, peer review is fallible and editors need to encourage the corrective process through letters to the editors. They should encourage more letters by removing overly strict limitations on the number of words allowed or time frames in which they must be submitted. Editors can also ensure that critical responses are acknowledged where the original studies are downloaded from electronic journal websites and that they are accessible with search engines.

The credibility of the health psychology literature is worth taking such steps to preserve. Patients, clinicians, and policy makers are depending on our delivery of accurate, helpful information that improves health outcomes. ■

References

- Altman, D. G., Schulz, K. F., Moher, D., Egger, M., Davidoff, F., Elbourne, D. et al. (2001). The revised CONSORT statement for reporting randomized trials: Explanation and elaboration. *Annals of Internal Medicine*, 134, 663-694.
- Andersen, B. L., Yang, H. C., Farrar, W. B., Golden-Kreutz, D. M., Emery, C. F., Thornton, L. M., et al. (2008). Psychologic Intervention Improves Survival for Breast Cancer Patients A Randomized Clinical Trial. *Cancer*, 113(12), 3450-3458.
- Boesen, E. H., Boesen, S. H., Frederiksen, K., Ross, L., Dahlstrom, K., Schmidt, G., et al. (2007). Survival after a psychoeducational intervention for patients with cutaneous malignant melanoma: A replication study. *Journal of Clinical Oncology*, 25(36), 5698-5703.
- Cohen, J. (1992). A power primer. *Psychological Bulletin*, 112, 155-159.
- Cook, J.M., Palmer, S., Hoffman, K., & Coyne, J.C. (2007). "Evaluation of clinical trials appearing in Journal of Consulting and Clinical Psychology: CONSORT and beyond." *The Scientific Review of Mental Health Practice* 5, 69-80, 2007.
- Coyne, J.C., Lepore, S. J., & Palmer, S. C. (2006). Efficacy of psychosocial interventions in cancer care: Evidence is weaker than it first looks. *Annals of Behavioral Medicine*, 32, 104-110.
- Coyne, J.C., Stefanek, M., & Palmer, S. C. (2007). Psychotherapy and survival in cancer: The conflict between hope and evidence. *Psychological Bulletin*, 133(3), 367-394.
- Coyne, J.C., Thombs, B., & Hagedoorn, M. (Manuscript Under Review). Ain't Necessarily So: Review and Critique of Recent Meta-Analyses of Behavioral Medicine Interventions in *Health Psychology*.
- Coyne, J.C., Thombs, B. D., Stefanek, M., & Palmer, S. C. (2009). Time to Let Go of the Illusion That Psychotherapy Extends the Survival of Cancer Patients: Reply to Kraemer, Kuchler, and Spiegel (2009). *Psychological Bulletin*, 135(2), 179-182.
- DeAngelis, C. D., Drazen, J. M., Frizelle, F. A., Haug, C., Hoey, J., Horton, R., et al. (2004). Clinical trial registration - A statement from the International Committee of Medical Journal Editors. *JAMA-Journal of the American Medical Association*, 292(11), 1363-1364.
- Ioannidis, J. P. A. (2005a). Why most published research findings are false. *Plos Medicine*, 2(8), 696-701.
- Ioannidis, J. P. A. (2005). Contradicted and initially stronger effects in highly cited clinical research. *JAMA-Journal of the American Medical Association*, 294(2), 218-228.
- Ioannidis, J. P. A. (2008). Why most discovered true associations are inflated. *Epidemiology*, 19(5), 640-648.

Acknowledgements

Special thanks to Brett Thombs, Mariet Hagedoorn, Stephen C. Palmer, Howard Leventhal and Michael Stefanek. Over time, discussions with them contributed to the ideas presented in this article and their support was important in encouraging its writing.

conference announcements

STAR 2010: 31st World Conference on Stress & Anxiety Research. 4-Aug-10 to 06-Aug-10. Galway, Ireland.

UK Society of Behavioural Medicine 5th Scientific Meeting. 14-Dec-09 to 15-Dec-09. Southampton, England.

11th International Congress of Behavioral Medicine. 4-Aug-10 to 4-Aug-10. Washington DC, USA.