

Figuring out what happened in a psychotherapy intervention trial

January 14 2016, by James Coyne Phd.

John Ioannidis, the "scourge of sloppy science" has documented again and again that the safeguards being introduced into the biomedical literature against untrustworthy findings are usually inconsistent and ineffective. In Ioannidis' most recent report , his group:

...Assessed the current status of reproducibility and transparency addressing these indicators in a random sample of 441 biomedical journal articles published in 2000–2014. Only one study provided a full protocol and none made all raw data directly available.

As reported in a recent post in "Retraction Watch, Did a clinical trial proceed as planned? New project finds out," psychiatrist Ben Goldacre has a new project with

...The relatively straightforward task of comparing reported outcomes from clinical trials to what the researchers said they planned to measure before the trial began. And what they've found is a bit sad, albeit not entirely surprising.

Ben Goldacre specifically excludes trials of psychotherapy from this project. But there are reasons to believe that the psychotherapy literature is less trustworthy than the [biomedical literature](#) because psychotherapy trials are less frequently registered, adherence to CONSORT reporting standards is less strict, and investigators more routinely refuse to share data when requested.

Untrustworthiness in the psychotherapy literature can have important consequences for patients, clinical practice, and public health and social policy.

The study that I review here twice switched outcomes in its reports, had a poorly chosen comparison control group and flawed analyses, and its protocol was registered after the study started. Yet, the study will very likely provide data for decision-making about what to do with primary care patients with a few unexplained medical symptoms. The message of the investigators is to deny these patients medical tests and workups and instead apply to them an unvalidated psychiatric diagnosis and a treatment that encourages them to believe that their concerns are irrational.

In this post I will attempt to track what should have been an orderly progression from (a) registration of a psychotherapy trial to (b) publishing of its protocol to (c) reporting of the trial's results in the peer-reviewed literature. This exercise will show just how difficult it is to make sense of studies in a poorly documented psychological intervention literature.

I find lots of surprises, including outcome switching in both reports of the trial.

The second article reporting results of the trial that does not acknowledge registration, minimally cites the first reports of outcomes, and hides important shortcomings of the trial. But the authors inadvertently expose new crucial shortcomings without comment.

Detecting important inconsistencies between registration and protocols and reports in the journals requires an almost forensic attention to detail to assess the trustworthiness of what is reported. Some problems hide in plain sight if one takes the time to look, but others require a certain

clinical connoisseurship, a well-developed appreciation of the subtle means by which investigators spin outcomes to detect.

Outcome switching and inconsistent cross-referencing of reports of a clinical trial will bedevil any effort to integrate the results of the trial into the larger literature in a systematic review or meta-analysis.

Two journals – *Psychosomatic Medicine* and particularly *Journal of Psychosomatic Research* failed adequate review of articles based on this trial, both in terms of this regard to trial registration, outcome switching, and allowing multiple reports of what could be construed as primary outcomes the same trial into the literature.

Despite serious problems in their interpretability, results of this study are likely to be cited and establishing far-reaching public policies.

The generalizability of results of this exercise is unclear, but my findings encourage skepticism about reports of psychotherapy interventions. It is distressing that more alarm bells have not been sounded about reports of the study.

The publicly accessible registration of the trial is:

Cognitive Behaviour Therapy for Abridged Somatization Disorder (Somatic Symptom Index [SSI] 4,6) patients in primary care. Current controlled trials ISRCTN69944771

The publicly accessible full protocol is:

Magallón R, Gili M, Moreno S, Bauzá N, García-Campayo J, Roca M, Ruiz Y, Andrés E. Cognitive-behaviour therapy for patients with Abridged Somatization Disorder (SSI 4, 6) in primary care: a randomized, controlled study. *BMC Psychiatry*. 2008 Jun 22;8(1):47.

The second report of treatment outcomes in *Journal of Psychosomatic Research*

Readers can more fully appreciate the problems that I uncovered if I work backwards from the second published report of outcomes from the trial. Published in *Journal of Psychosomatic Research*, the article is behind a pay wall, but readers can write to the corresponding author for a PDF: mgili@uib.es. Because this person is also the corresponding author for the second paper in *Psychosomatic Medicine*, readers might want to request both papers.

Gili M, Magallón R, López-Navarro E, Roca M, Moreno S, Bauzá N, García-Cammpayo J. Health related quality of life changes in somatising patients after individual versus group cognitive behavioural therapy: A [randomized clinical trial](#). *Journal of Psychosomatic Research*. 2014 Feb 28;76(2):89-93.

The title is misleading in its ambiguity because "somatising" does not refer to an established diagnostic category, but an unvalidated category that encompasses a considerable proportion of primary care patients, usually those with comorbid anxiety or depression. More about that later.

PubMed, which usually reliably attaches a trial registration number to abstracts, doesn't do so for this article

The article does not list the registration, and does not provide the citation when indicating that the trial protocol is available. The only subsequent citations of the trial protocol are ambiguous:

More detailed design settings and study sample of this trial have been described elsewhere [14,16], which explain the effectiveness of CBT reducing number and severity of somatic symptoms.

The above quote is also the sole citation of a key previous paper, which presents outcomes for the trial. Only an alert and motivated reader would catch this. No opportunity within the article is provided for comparing and contrasting results of the two papers.

The brief introduction displays a decided puffer fish phenomenon, exaggerating the prevalence and clinical significance of the unvalidated "abridged somatization disorder." Essentially, the authors invoke problematic, but accepted psychiatric diagnostic categories somatoform or somatization disorders. Oddly, the category has different criteria when applied to men and women: men require four unexplained medical symptoms, whereas women require six.

I haven't previously counted the term "abridged" in psychiatric diagnosis. Maybe the authors mean "subsyndromal," as in "subsyndromal depression." This is a dubious labeling because it suggested all characteristics needed for diagnosis are not present, some of which may be crucial. Think of it: is a persistent cough subsyndromal lung cancer or maybe emphysema? References to symptoms being "subsyndromal" often occur in context where exaggerated claims about prevalence are made and inappropriate inferences made about treatment of milder cases from the more severe.

A casual reader might infer that the authors are evaluating a psychiatric treatment with wide applicability to as many as 20% of primary care patients. As we will see, the treatment focuses on discouraging any diagnostic medical tests and trying to convince the patient that their concerns are irrational.

The introduction identifies the primary outcome of the trial:

The aim of our study is to assess the efficacy of a cognitive behavioural intervention program on HRQoL [health-related quality of life] of

patients with abridged somatization disorder in primary care.

This primary outcome is inconsistent with what was reported in the registration, the published protocol, and the first article reporting outcomes. The earlier report does not even mention the inclusion of a measure of HRQoL, measured by the SF-36. It is listed in the study protocol as a "secondary variable."

The opening of the methods section declares that the trial is reported in this paper consistent with the Consolidated Standards of Reporting Clinical Trials (CONSORT). This is not true because the flowchart describing patients from recruitment to follow-up is missing. We will see that when it is reported in another paper, some important information is contained in that flowchart.

The methods section reports only three measures were administered: a Standardized Polyvalent Psychiatric Interview (SPPI), a semistructured interview developed by the authors with minimal validation; a screening measure for somatization administered by primary care physicians to patients whom they deemed appropriate for the trial, and the SF-36.

Crucial details are withheld about the screening and diagnosis of "abridged somatization disorder" that, if presented, might lead a reader to further doubt the validity of this unvalidated and idiosyncratic diagnosis.

Few readers, even primary care physicians or psychiatrists, will know what to make of the Smith's guidelines (Googling it won't yield much), which is essentially a matter of simply sending a letter to the referring GP. Sending such a letter is a notoriously ineffective intervention in primary care. It mainly indicates that patients referred to a trial did not assigned to an active treatment. As I will document later, the authors were well aware that this would be an ineffectual control/comparison

intervention, but it would ensure that their preferred intervention would look quite good in terms of effect size.

The two active interventions are individual- and group-administered CBT which is described as:

Experimental or intervention group: implementation of the protocol developed by Escobar [21,22] that includes ten weekly 90-min sessions. Patients were assessed at 4 time points: baseline, post-treatment, 6 and 12 months after finishing the treatment. The CBT intervention mainly consists of two major components: cognitive restructuring, which focuses on reducing pain-specific dysfunctional cognitions, and coping, which focuses on teaching cognitive and behavioural coping strategies. The program is structured as follows. Session 1: the connection between stress and pain. Session 2: identification of automated thoughts. Session 3: evaluation of automated thoughts. Session 4: questioning the automatic thoughts and constructing alternatives. Session 5: nuclear beliefs. Session 6: nuclear beliefs on pain. Session 7: changing coping mechanisms. Session 8: coping with ruminations, obsessions and worrying. Session 9: expressive writing. Session 10: assertive communication.

There is sparse presentation of data from the trial in the results section, but some fascinating details await a skeptical, motivated reader.

Table 1 displays social demographic and clinical variables. Psychiatric comorbidity is highly prevalent. Readers can't tell exactly what is going on, because the authors' own interview schedule is used to assess comorbidity. But it appears that all but a small minority of patients diagnosed with "abridged somatization disorder" have substantial anxiety and depression. Whether these symptoms meet formal criteria cannot be determined. There is no mention of physical comorbidities.

But there is something startling awaiting an alert reader in Table 2.

Scales Mean (SD)	Physical functioning
TAU	
<i>Pre</i>	44.17 (26.97)
<i>Post</i>	52.08 (29.45)
<i>6 months</i>	51.04 (30.07)
<i>12 months</i>	50.63 (22.71)
Individual CBT	
<i>Pre</i>	66.86 (27.65)
<i>Post</i>	77.33 (20.77)
<i>6 months</i>	80.81 (22.23)
<i>12 months</i>	79.42 (24.52)
Group CBT	
<i>Pre</i>	51.28 (29.82)
<i>Post</i>	56.86 (27.21)
<i>6 months</i>	59.19 (25.35)
<i>12 months</i>	61.28 (25.45)

There is something very odd going on here, and very likely a breakdown of randomization. Baseline differences in the key outcome measure, SF-36 are substantially greater between groups than any within-group change. The treatment as usual condition (TAU) has much lower functioning [lower scores mean lower functioning] than the group CBT condition, which is substantially below the individual CBT difference.

If we compare the scores to adult norms, all three groups of patients are poorly functioning, but those "randomized" to TAU are unusually impaired, strikingly more so than the other two groups.

Keep in mind that evaluations of active interventions, in this case CBT, in randomized trials always involve a between difference between

groups, not just difference observed within a particular group. That's because a comparison/control group is supposed to be equivalent for nonspecific factors, including natural recovery. This trial is going to be very biased in its evaluation of individual CBT, a group within which patients started much higher in physical functioning and ended up much higher. Statistical controls fail to correct for such baseline differences. We simply do not have an interpretable clinical trial here.

The first report of treatment outcomes in Psychosomatic Medicine

Moreno S, Gili M, Magallón R, Bauzá N, Roca M, del Hoyo YL, Garcia-Campayo J. Effectiveness of group versus individual cognitive-behavioral therapy in patients with abridged somatization disorder: a randomized controlled trial. *Psychosomatic medicine*. 2013 Jul 1;75(6):600-8.

The title indicates that the patients are selected on the basis of "abridged somatization disorder."

The abstract prominently indicates the trial registration number (ISRCTN69944771), which can be plugged into Google to reach the publicly accessible registration.

If a reader is unaware of the lack of validation for "abridged somatization disorder," they probably won't infer that from the introduction. The rationale given for the study is that

A recently published meta-analysis (18) has shown that there has been ongoing research on the effectiveness of therapies for abridged somatization disorder in the last decade.

Checking that meta-analysis, it did include a single null trial for treatment of abridged somatization disorder. This seems like a

gratuitous, ambiguous citation.

I was surprised to learn that in three of the five provinces in which the study was conducted, patients

... Were not randomized on a one-to-one basis but in blocks of four patients to avoid a long delay between allocation and the onset of treatment in the group CBT arm (where the minimal group size required was eight patients). This has produced, by chance, relatively big differences in the sizes of the three arms.

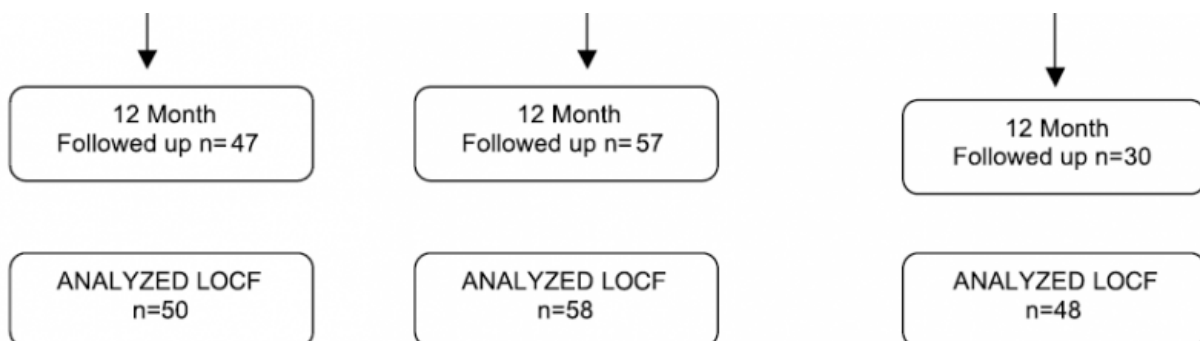
This departure from one-to-one randomization was not mentioned in the second article reporting results of the study, and seems an outright contradiction of what is presented there. Neither is it mentioned nor in the study protocol. This patient selection strategy may have been the source of lack of baseline equivalence of the TAU and to intervention groups.

For the vigilant skeptic, the calculation of sample size is an eye-opener. Sample size estimation was based on the effectiveness of TAU in primary care visits, which has been assumed to be very low (approximately 10%).

Essentially, the authors are justifying a modest sample size because they don't expect the TAU intervention has the efficacy. How could authors believe there is equipoise, that the comparison control treatments could be expected to be equally effective? They seem to be saying they don't believe this, but equipoise is an ethical and practical requirement for a clinical trial to which human subjects are being recruited. In terms of trial design, do the authors really think this poor treatment provides an adequate comparison/control?

In the methods section, the authors also provide a study flowchart, which

was required for the other paper to adhere to CONSORT standards but was missing. Note the flow at the end of the study for the TAU comparison/control condition at the far right. There was substantially more dropout in this group. The authors chose to estimate the scores with the Last Observation Carried Forward (LOCF) method which assumes the last available observation can be substituted for every subsequent one. This is a discredited technique and particularly inappropriate in this context. Think about it: the TAU condition was expected by the authors to be quite poor care and so more patients dropped out. But they might've been dropped out deteriorating. Certainly it cannot be assumed that the smaller number of dropouts from the other conditions were from the same reason. We have a methodological and statistical mess on our hands, but it was hidden from us in our discussion of the second report.



Six measures are mentioned: (1) the Othmer-DeSouza screening instrument used by clinicians to select patients; (2) the Screening for Somatoform Disorders (SOMS, a 39 item questionnaire that includes all

bodily symptoms and criteria relevant to somatoform disorders according to either DSM-IV or ICD-10; (3) a Visual Analog Scale of somatic symptoms (Severity of Somatic Symptoms scale) that patients use to assess changes in severity in each of 40 symptoms; (4) the authors own SPPI semistructured psychiatric interview for diagnosis of psychiatric morbidity in primary care settings; (5) the clinician administered Hamilton Anxiety Rating Scale; and the (6) Hamilton Depression Rating Scale.

We are never actually told what the primary outcome is for the study, but it can be inferred from the opening of the discussion:

The main finding of the trial is a significant improvement regardless of CBT type compared with no intervention at all. CBT was effective for the relief of somatization, reducing both the number of somatic symptoms (Fig. 2) and their intensity (Fig. 3). CBT was also shown to be effective in reducing symptoms related to anxiety and depression.

But I noticed something else here, after a couple of readings. The items used to select patients and identify them with "abridged somatization disorder" has 39 or 40 symptoms, that men only need four and women only need six symptoms for a diagnosis. That means that most pairs of patients receiving a diagnosis will not have a symptom in common. Whatever "abridged somatization disorder" means, patients who received this diagnosis are likely to be different from each other in terms of somatic symptoms, but probably have other characteristics in common.

Comparison of this report to the outcomes paper we were reviewed earlier, shows none of these outcomes are mentioned as being assessed and certainly not has outcomes.

Comparison of this report to the published protocol reveals that number

and intensity of somatic symptoms are two of the three main outcomes, but this article makes no mention of the third, utilization of healthcare.

Readers can find something strange in Table 2 presenting what seems to be one of the primary outcomes, severity of symptoms. In this table the order is TAU, group CBT, and individual CBT. Note the large difference in baseline symptoms with group CBT being much more severe. It's difficult to make sense of the 12 month follow-up because there was differential drop out and reliance on an inappropriate LOCR imputation of missing data. But if we accept the imputation as the authors did, it appears that there were no differences between TAU and group CBT. That is what the authors reported with inappropriate analyses of covariance.

Severity of Somatic Symptoms scale total score			
Baseline	71.48 (57.78–85.19)	92.06 (79.02–105.1)	63.26 (50.87–75.65)
Posttreatment	59.98 (50.44–69.51)	41.52 (32.69–50.35)	34.9 (26.42–43.38)
6-Month follow-up	64.02 (56.24–71.8)	54.27 (47.06–61.47)	45.31 (38.39–52.23)
12-Month follow-up	67.7 (60.42–74.98)	67.84 (61.03–74.65)	55.53 (49–62.06)

The authors' cheerful take away message?

This trial, based on a previous successful intervention proposed by Sumathipala et al. (39), presents the effectiveness of CBT applied at individual and group levels for patients with abridged somatization (somatic symptom indexes 4 and 6).

But hold on! In the introduction, the authors' justification for their trial was

Evidence for the group versus individual effectiveness of cognitive-behavioral treatment of medically unexplained physical symptoms in the primary care setting is not yet available.

And let's take a look at Sumathipala et al.

Sumathipala A, Siribaddana S, Hewege S, Sumathipala K, Prince M, Mann A. Understanding the explanatory model of the patient on their medically unexplained symptoms and its implication on treatment development research: a Sri Lanka Study. *BMC Psychiatry*. 2008 Jul 8;8(1):54.

The article presents speculations based on an observational study, not an intervention study so there is no success being reported.

The formal registration

The registration of psychotherapy trials typically provides sparse details. The curious must consult the more elaborate published protocol. Nonetheless, the registration can often provide grounds for skepticism, particularly when it is compared to any discrepant details in the published protocol, as well as subsequent publications.

This protocol declares

Study hypothesis

Patients randomized to cognitive behavioural therapy significantly improve in measures related to quality of life, somatic symptoms, psychopathology and health services use.

Primary outcome measures

Severity of Clinical Global Impression scale at baseline, 3 and 6 months and 1-year follow-up

Secondary outcome measures

The following will be assessed at baseline, 3 and 6 months and 1-year follow-up:

- 1. Quality of life: 36-item Short Form health survey (SF-36)*
- 2. Hamilton Depression Scale*
- 3. Hamilton Anxiety Scale*
- 4. Screening for Somatoform Symptoms [SOMS]*

Overall trial start date

15/01/2008

Overall trial end date

01/07/2009

Primary outcome

Main outcome variables:

– SSS (Severity of somatic symptoms scale) [22]: a scale of 40 somatic symptoms assessed by a 7-point visual analogue scale.

– SSQ (Somatic symptoms questionnaire) [22]: a scale made up of 40 items on somatic symptoms and patients' illness behaviour.

When I searched for, Severity of Clinical Global Impression, the

primary outcome declared in the registration , and I could find no reference to it.

The protocol was submitted on May 14, 2008 and published on June 22, 2008. This suggests that the protocol was submitted after the start of the trial.

To calculate the sample size we consider that the effectiveness of usual treatment (Smith's norms) is rather low, estimated at about 20% in most of the variables [10,11]. We aim to assess whether the new intervention is at least 20% more effective than usual treatment.

Comparison group

Control group or standardized recommended treatment for somatization disorder in [primary care](#) (Smith's norms) [10,11]: standardized letter to the family doctor with Smith's norms that includes: 1. Provide brief, regularly scheduled visits. 2. Establish a strong patient-physician relationship. 3. Perform a physical examination of the area of the body where the symptom arises. 4. Search for signs of disease instead of relying of symptoms. 5. Avoid diagnostic tests and laboratory or surgical procedures. 6. Gradually move the patient to being "referral ready".

Basically, TAU, the comparison/control group involves simply sending a letter to referring physicians encouraging them simply to meet regularly with the patients but discouraged diagnostic test or medical procedures. Keep in mind that patients for this study were selected by the physicians because they found them particularly frustrating to treat. Despite the authors' repeated claims about the high prevalence of "abridged somatization disorder," they relied on a large number of general practice settings that would each contribute relatively few patients for setting. These patients a very heterogeneous in terms of somatic symptoms, but most share anxiety or depressive symptoms.

There is an uncontrolled selection bias here that makes generalization from results of the study problematic. Just who are these patients? I wonder if these patients have some similarity to the frustrating GOMERS (Get Out Of My Emergency Room) in the classic House of God, described by Amazon as "an unvarnished, unglorified, and amazingly forthright portrait revealing the depth of caring, pain, pathos, and tragedy felt by all who spend their lives treating patients and stand at the crossroads between science and humanity."

Imagine the disappointment about the physicians and the patients when randomization simply left them in routine care. It's no wonder that the patients deteriorated and that patients assigned to this treatment dropped out more.

Whatever active ingredients the individual and group CBT have, they also include some nonspecific factors missing from the TAU comparison group: frequency and intensity of contact, reassurance and support, attentive listening, and positive expectations. These nonspecific factors can readily be confused with active ingredients and may account for any differences between the active treatments and the TAU comparison. What terrible study.

The two journals providing reports of the studies failed to responsibility to the readership and the larger audience seeking clinical and public policy relevance. Authors have ample incentive to engage in questionable publication practices, including ignoring and even suppressing registration, switching outcomes, and exaggerating the significance of their results. Journals of necessity must protect authors from their own inclinations, as well as the readers and the larger medical community from on trustworthy reports. *Psychosomatic Medicine* and *Journal of Psychosomatic Research* failed miserably in their peer review of these articles. Neither journal is likely to be the first choice for authors seeking to publish findings from well-designed and well reported

trials. Who knows, maybe the journals' standards for compromised by the need to attract randomized [trials](#) for what is construed as a psychosomatic condition, at least by the psychiatric community.

Postscript: Imagine what will be done with the results of this study

[In a recent blog post](#), I examined a registration for a protocol for a systematic review and meta-analysis of interventions to address medically unexplained symptoms. The protocol was inadequately described, had undisclosed conflicts of interest, and one of the senior investigators had a history of switching outcomes and refusing to share data for independent analysis. Undoubtedly, the study we have been discussing meets the vague criteria for inclusion in this meta-analysis. But what outcomes will be chosen, particularly when they should only be one outcome per study? And will be recognized that these two reports are actually the same study? Will key problems in the designation of the TAU control group be recognized, with its likely inflation of treatment effects, when used to calculate effect sizes?

As you can see, it took a lot of effort to compare and contrast documents that should have been in alignment. Do you really expect those who conduct subsequent meta-analyses to make those multiple comparisons or will they simply extract multiple effect sizes from the two papers so far reporting results?

More information: Margalida Gili et al. Health related quality of life changes in somatising patients after individual versus group cognitive behavioural therapy: A randomized clinical trial, *Journal of Psychosomatic Research* (2014). [DOI: 10.1016/j.jpsychores.2013.10.018](https://doi.org/10.1016/j.jpsychores.2013.10.018)

Athula Sumathipala et al. Understanding the explanatory model of the patient on their medically unexplained symptoms and its implication on treatment development research: a Sri Lanka Study, *BMC Psychiatry* (2008). DOI: [10.1186/1471-244X-8-54](https://doi.org/10.1186/1471-244X-8-54)

Rosa Magallón et al. Cognitive-behaviour therapy for patients with Abridged Somatization Disorder (SSI 4,6) in primary care: a randomized, controlled study, *BMC Psychiatry* (2008). DOI: [10.1186/1471-244X-8-47](https://doi.org/10.1186/1471-244X-8-47)

This story is republished courtesy of PLOS Blogs: blogs.plos.org.

Provided by Public Library of Science

Citation: Figuring out what happened in a psychotherapy intervention trial (2016, January 14) retrieved 24 April 2024 from <https://medicalxpress.com/news/2016-01-figuring-psychotherapy-intervention-trial.html>

This document is subject to copyright. Apart from any fair dealing for the purpose of private study or research, no part may be reproduced without the written permission. The content is provided for information purposes only.