Contemporary Clinical Trials Communications 2 (2016) 12–15

Contents lists available at ScienceDirect

Contemporary Clinical Trials Communications

journal homepage: www.elsevier.com/locate/conctc

Trial participant representativeness compared to ordinary service users in a work rehabilitation setting

Simon Øverland [a, b, *], Astrid Grasdal [c], Silje Endresen Reme [d]

[a] Department of Public Mental Health, Norwegian Institute of Public Health, Norway
[b] Department of Psychosocial Science, University of Bergen, Norway
[c] Department of Economics, University of Bergen, Norway
[d] Uni Health, Uni Research, Bergen, Norway

ARTICLE INFO

Article history:
Received 28 July 2015
Received in revised form
20 November 2015
Accepted 2 December 2015
Available online 17 December 2015

Key MeSH terms:
Pragmatic clinical trial
Depression
Anxiety
Disability
Sick leave

ABSTRACT

Background: Study representativeness is a major concern for generalizations from trials. The extent of the problem varies with study design and context. There is a strong emphasis on developing interventions to help people remain in the work force despite mental illness. We need to know if results from upcoming trials in this area are valid for those that later might receive the services.
Method: The AWaC trial was a multicenter RCT conducted at six different treatment centers (n = 1193). After the trial was over, the centers were upheld and run as ordinary services. At that time, we surveyed 80 ordinary service users with the same baseline questionnaire as used in the trial, and compared them with those who participated in the trial.
Results: There were a higher proportion of people with the highest level of education (4 years or more at university/college) in the post-trial comparison sample. This sample also reported to be "dissatisfied" with their job more often, but rated their chances for return to work as "bad" less often than the ordinary trial participants. No further significant differences between the two samples in any of the other education categories, or for any of the other demographic, health or work related comparisons were found.
Discussion: Participation bias is likely to depend on study context, but in the setting of a trial to help improve work participation among people who struggle with common mental disorders, the trial participants were overall very similar to those who sought the same services as ordinary practice.

1. Introduction

Running a sound clinical trial is demanding, but implementation of results in the wider population present numerous layers of challenges. Trials commonly employ strict inclusion and exclusion criteria to maintain experimental control and secure internal validity [1]. Furthermore, participation in trials is based on informed consent, and sub-sections of the population might be more or less inclined to participate [2]. Participant selection and bias may arise from criteria definitions and selective participation, and determine generalizability of trial results outside the study sample [3,4].

For clinical trials in mental health, researchers have tried to quantify this problem by assessing ordinary help-seekers with common study inclusion criteria. Of 346 patients with depression, Zimmermann et al. found that only 1/6 would qualify for inclusion in antidepressant efficacy trials employing common criteria [5]. The eligible population also differed from the non-eligible on demographic, clinical and psychological profiles [6]. In a study of actual trial recruitment, 7 out of 8 who volunteered, were turned away due to pre-defined criteria for study inclusion [7]. In a Dutch study comparing participants in RCT's for major depressive disorder to patients attending ordinary care, the trial participants were much more likely to be employed [8].

The core question of whether results from trials can be generalized to a larger population does not seem to converge to one general conclusion, but vary with the context of the trial. Some studies argue that their trial results can represent real-life effects outside the trial settings [1,9], whereas others prompt careful consideration of generalizability [5,6,10,11].

Sickness absence and long-term work disability is a major issue in occupational medicine, and we find it is likely that the number of

* Corresponding author. Department of Public Mental Health, Norwegian Institute of Public Health, Kalfarveien 31, 5018 Bergen, Norway
E-mail address: simon.overland@fhi.no (S. Øverland).

http://dx.doi.org/10.1016/j.conctc.2015.12.005

trials to prevent long-term sickness absence and disability pension from mental illness will increase. In a recent pragmatic trial in this context [12], we defined exclusion/inclusion to allow recruitment of participants who also would be likely to seek the intervention outside the trial. In this paper, we examined if we were successful, and hypothesised that the participants in our large multi-center RCT did not differ substantially from those who attended the same services outside the setting of a trial.

2. Methods

2.1. Context

The context of the AWaC-trial, from which these data were generated (pre-trial registration details at ClinicalTrials.gov, registration number: NCT01146730) can be found in the main effect study and its protocol [12]. In short, the trial tested the effectiveness of the At Work and Coping intervention (AWaC) vs. usual care in helping people who struggle with work participation due to mild to moderate common mental illnesses participate in working life.

2.2. Study samples

This analysis is a comparison of a trial sample and a post-trial comparison sample.

2.3. Trial sample

Trial participants were mainly recruited through referrals from the Norwegian Welfare and Labor Administration, GPs and self-referrals. The main inclusion criterion was that common mental disorder was seen as the main reason why the person struggled with work participation. The accepted age span was 18–60 years, and potential participants had to express a motivation to return to/stay at work. People were excluded if they reported other reasons as the primary cause of work participation problems (e.g. somatic, social, economic and work-related issues), no motivation/desire to work, suffered from severe psychiatric disorders, had high suicide risk or a current substance abuse problem, or was engaged in psychotherapy elsewhere already. Pregnant women were excluded, as were people unable to read or write in Norwegian.

Potential participants were informed about the project, and screened for inclusion and exclusion criteria at the centers. Eligible and willing participants signed the informed consent and completed the baseline questionnaire. After random allocation, participants were written about the outcome and the intervention group were given a date for their first session.

1416 potential participants were referred and considered for inclusion. Of these, 197 did not fulfill the inclusion criteria, 17 did not consent to participate, and 9 withdrew their consent and required data deletion (2 from the intervention group and 7 from the control group). In total, 1193 participants entered the trial and were randomized.

2.4. Post-trial comparison sample

The sub sample of 80 persons was recruited after the completion of the trial, and we obtained specific ethical approval for this data collection. The data were collected at the 6 centers that were part of the multicenter trial in June 2012, when the centers no longer recruited or evaluated potential patients for trial inclusion, or had trial participants in treatment. New cases enrolled at the centers over a period of one month were invited to participate. Like the trial participants, they came to the centers after referrals from the Norwegian Welfare and Labor Administration, their GP, self-referrals or through other channels. We do not have the exact figures on how many attended the centers during June 2013, but the final n of 80 exceeded the average number included in the trial per month (total n = 1193 included over a period of 18 months equals an average of 67 participants per month). Those who were willing to participate in the post-trial sample, were asked to complete a shortened version of the baseline questionnaire used in the trial.

2.5. Statistical comparisons

The aim of this study was to compare the trial participants with those who found their way to the same treatment centers after the trial was over. We did this by comparing the two samples in terms of self-reported demographic characteristics and scores on key health variables using chi-square tests for categorical variables and t-tests for continuous variables.

2.6. Ethics

The trial and the post-trial comparison survey were both approved by the regional committee for medical research and all participants provided informed consent.

3. Results

There was no overall difference between the samples on any of the variables. When examining single response levels separately there was a higher proportion of people with the highest level of education (5 years or more at university/college) in the post-trial comparison sample. They also reported to be “dissatisfied” with their job more often, but rated their chances for return to work as “bad” less often than the ordinary trial participants. Beyond that, there was no significant difference between the two samples in any of the other education categories, or for any of the other demographic, health or work related comparisons (Table 1).

4. Discussion

The data supported the hypothesis that those who participated in a pragmatic multicenter randomized controlled trial were comparable to those who attended the same services outside the context of a trial in terms of self-reported health and demographic characteristics.

There are limitations to our study that should be considered. There could of course be a common self-selection process in both these samples, where those unwilling to participate in research at all – both concerning the trial and the post-trial comparison – share the same characteristics and health status. Anecdotal evidence however suggested that disapproval against randomization was an important hindrance to participation for individuals, but also their referring doctors or case-managers.

By logic of multiple testing, one in twenty comparisons should appear significant despite no true underlying difference. Here, we did a total of 35 comparisons, and found three statistical differences between the samples when directly comparing individual scores on a scale, despite no overall difference for those scales. Thus, even in the presence of these statistically significant differences between the samples, we argue these are insufficient to reject our initial hypothesis of similar samples.

Anecdotally, the treatment centers reported increased referral rates after the trial inclusion period, which resonates with the higher number of participants in the post-trial comparison sample compared to the average inclusion per month during the trial. Individuals, case-managers and general practitioners could have been skeptical to participation as trials are more uncommon in this

Table 1
Background characteristics, trial-participants and regular service users (post-trial comparison sample) compared in t-tests and chi-square tests for continuous and categorical variables respectfully.

Variable	Trial sample (n = 1193)	Post-trial comparison sample (n = 80)	p-value	F-statistic (p-value)
Female (%)	67.25	63.75	0.519	
Married (%)	31.23	33.75	0.640	
Age (%)				0.28 (0.596)
<30	12.81	13.75	0.809	
30–39	30.31	31.25	0.861	
40–49	32.41	33.75	0.805	
50+	24.45	21.25	0.517	
Education (%)				2.66 (0.103)
Primary	7.11	5.00	0.427	
Senior high	32.08	30.00	0.700	
University/College 1–4 yrs	36.02	26.25	0.080	
University/College 5 + yrs	19.23	35.00	**0.001**	
Other	5.44	3.75	0.514	
Self assessed health (%)				0.22 (0.643)
Good	37.27	40.00	0.625	
Medium	45.98	43.75	0.698	
Poor	15.91	15.00	0.829	
Employed (%)	66.67	65.00	0.760	
Job satisfaction[a]				1.70 (0.193)
Very dissatisfied	5.53	6.25	0.785	
Dissatisfied	13.07	21.25	**0.039**	
Neither satisfied nor dissatisfied	17.42	13.75	0.400	
Satisfied	29.06	26.25	0.591	
Very satisfied	10.05	10.00	0.988	
Have no job	21.52	21.25	0.954	
Return-to-work prospects[b]				0.21 (0.650)
Very good	12.40	12.50	0.978	
Good	15.00	18.75	0.365	
Neither good nor bad	26.13	27.50	0.787	
Bad	17.34	8.75	**0.047**	
Very bad	13.65	18.75	0.203	
Missing data	15.49	13.75	0.676	
HADS[c] Anxiety (mean)	10.67	10.91	0.615	
HADS[c] Depression (mean)	8.07	8.19	0.801	
Subjective health complaints[d] (mean)	11.78	11.36	0.467	

p-values from chi-square test for dichotomous variables, and t-test for continuous variables.
Significant differences ($p < 0.05$) between the samples in bold.
[a] "Everything considered, how satisfied are you with your current job?"
[b] Consider this statement: "I count on being back at work within a couple of weeks".
[c] Hospital Anxiety and Depression Scale, 14 items (7 on anxiety and 7 on depression) scored on a four point ordinal scale (0–3). A score of 8 or above is regarded as a case-defining symptom level.
[d] SHC-29, list of 29 subjectively reported common health complaints, four response levels on an ordinal scale (0–3).

sector. Some practitioners might have delayed referrals if they believed the intervention would yield positive results and be continued after the trial. By holding back their patients until after the service was defined as a research project, the GP's could avoid their patient ending up in the control group.

An amassing body of observational studies place common mental disorders as a prominent risk factor for adverse occupational outcomes across countries and welfare systems. We will likely see more trials being conducted in this area in the years to come, and evidence to support the representativeness of trials in this context must follow. This problem is often overlooked and more careful implementation of trial results, moving them from the controlled to the pragmatic contexts, is needed. Our analysis contributes as a single result for the AWaC -trial, and supports that the trial results are valid for those who attend the service under normal circumstances. In lack of evidence to the contrary, self-selection out of trial participation might not be a major problem in this context. This paper also demonstrates a low-cost approach to study representativeness that other trials could include in protocols.

Acknowledgments

The AWaC trial was commissioned by the Norwegian Ministry of Health and Ministry of Labour, and financed through the National Strategy on Work and Mental Health (2007–2012), but no specific funding was acquired for this sub-project. The authors declare that they have no competing interests and therefore have nothing to disclose.

References

[1] A. von Wolff, M. Jansen, L.P. Holzel, et al., Generalizability of findings from efficacy trials for chronic depression: an analysis of eligibility criteria, Psychiatr. Serv. 65 (7) (2014) 897–904.
[2] P.N. Pfeiffer, B. Szymanski, N. Dhawan, et al., Is there an "unhealthy volunteer effect" in mental health outpatient research? Psychiatry Res. 176 (2–3) (2010) 224–228.
[3] M. Flather, N. Delahunty, J. Collinson, Generalizing results of randomized trials to clinical practice: reliability and cautions, Clin. Trials 3 (6) (2006) 508–512.
[4] P.M. Rothwell, External validity of randomised controlled trials: "To whom do the results of this trial apply?", Lancet 365 (9453) (2005) 82–93.
[5] M. Zimmerman, J.I. Mattia, M.A. Posternak, Are subjects in pharmacological treatment trials of depression representative of patients in routine clinical practice? Am. J. Psychiatry 159 (3) (2002) 469–473.
[6] M. Zimmerman, I. Chelminski, M.A. Posternak, Generalizability of antidepressant efficacy trials: differences between depressed psychiatric outpatients

who would or would not qualify for an efficacy trial, Am. J. Psychiatry 162 (7) (2005) 1370–1372.
[7] G.I. Keitner, M.A. Posternak, C.E. Ryan, How many subjects with major depressive disorder meet eligibility requirements of an antidepressant efficacy trial? J. Clin. Psychiatry 64 (9) (2003) 1091–1093.
[8] R. van der Lem, P.M. Stamsnieder, N.J. van der Wee, et al., Influence of sociodemographic and socioeconomic features on treatment outcome in RCTs versus daily psychiatric practice, Soc. Psychiatry Psychiatr. Epidemiol. 48 (6) (2013) 975–984.
[9] R. van der Lem, W.W. de Wever, N.J. van der Wee, et al., The generalizability of psychotherapy efficacy trials in major depressive disorder: an analysis of the influence of patient selection in efficacy trials on symptom outcome in daily practice, BMC Psychiatry 12 (2012) 192.
[10] N. Hoertel, Y. Le Strat, C. Blanco, et al., Generlizability of clinical trial results for generalized anxiety disorder to community samples, Depression Anxiety 29 (7) (2012) 614–620.
[11] N. Hoertel, Y. Le Strat, P. De Maricourt, et al., Are subjects in treatment trials of panic disorder representative of patients in routine clinical practice? results from a national sample, J. Affect. Disord. 146 (3) (2013) 383–389.
[12] S. Reme, A. Grasdal, C. Løvvik, et al., Work-focussed behavioural therapy and individual job support to increase work participation in common mental disorder. A randomized controlled multicenter trial, Occup. Environ. Med. 72 (2015) 745–752.