Gidlow, C., Johnston, L. H., Crone, D., & James, D. V. (2008). State of the art reviews: methods of evaluation: issues and implications for physical activity referral schemes. *American Journal of Lifestyle Medicine*, *2*(1), 46-50.

Title page

Evaluation of Physical Activity Referral Schemes: issues and implications

Running Head:

Evaluating Physical Activity Referral

Authors:

Christopher Gidlow BSc MSc PhD✉

Centre for Sport & Exercise Research

Staffordshire University,

Mellor Building, College Road

Stoke on Trent, ST4 2DE, UK

Email: c.gidlow@staffs.ac.uk

Tel: +44 (0) 1782 294038

Lynne H. Johnston BA MSc PhD

Clinical Psychology

Faculty of Medical Sciences,

University of Newcastle

Ridley Building

Newcastle upon Tyne, NE1 7RU, UK

Email: lynne.johnston@newcastle.ac.uk

Tel: +44 (0) 191 222 6000

Diane Crone BSc PhD

Faculty of Sport, Health & Social Care

University of Gloucestershire

Oxtalls Lane

Gloucester, GL2 9HW, UK

Email: dcrone@glos.ac.uk

Tel: +44 (0) 1242 715161

David V.B. James BSc PhD.

Faculty of Sport, Health & Social Care

University of Gloucestershire

Oxtalls Lane

Gloucester, GL2 9HW, UK

Email: djames@glos.ac.uk

Tel: +44 (0) 01242 715166

✉ Author to whom all correspondence should be addressed

Abstract

In the pursuit of evidence-based practice, the common focus on controlled research within physical activity and health promotion has created a restricted view of the acceptable type of evidence upon which practice should be based. However, to improve our understanding of physical activity (and other behavioural) interventions, a more holistic approach to evaluation is required. In the context of Physical Activity Referral Schemes, the present paper considers the implications of adhering to this narrow definition of 'acceptable' evidence and the importance of recognising alternative evaluative approaches.

Key words

Exercise referral; evidence based practice; evidence informed practice; longitudinal cohort design

Introduction

In the present climate of evidence-based practice, the evaluation of health promotion and physical activity interventions has become increasingly important[1] with randomised controlled trials (RCT) frequently being promoted as the 'gold standard'.[2,3] Systematic reviews, used to guide practitioners, managers and policy makers, often impose strict inclusion criteria, which preclude studies deviating from the RCT model. This creates a limited definition of *acceptable* evidence, which has implications for the development of evidence-based practice and policy.[4,5] These implications are well illustrated within the context of Physical Activity Referral Schemes (also known as Exercise Referral Schemes), and the development of an acceptable evidence base upon which to develop physical activity programmes in the community.

Physical Activity Referral Schemes (PARS), described in more detail elsewhere,[6,7] have proliferated rapidly in the UK since their conception in the early nineties (approximately 800 in 2003[8]). Physical Activity Referral Schemes involve the referral of patients by health professionals to undertake a programme of physical activity under the supervision of qualified exercise professionals (Figure 1). Despite the abundance of these schemes, the continued absence of robust RCT evidence for PARS effectiveness in promoting sustained behaviour change has fuelled ongoing debate regarding their public health value.[9-11] This debate culminated in 2006, when a serious question mark was placed over the future of PARS following recommendations that schemes should only be endorsed as part of controlled research into effectiveness.[12] These recommendations are a good example of how practice can be guided by findings from reviews that adhere to a narrow definition of evidence and are consequently based on a small number of studies.[13]

In the context of the PARS type of physical activity intervention, the present paper questions such a narrowly defined approach to evidence-based practice and considers the importance of alternative evaluative approaches. Building on previously reviewed PARS research[6] and including more recent examples,[14-16] specific aspects of PARS evaluation design and their implications will be considered. Although based in the PARS context, the aim of the present paper is to raise methodological issues that are relevant and applicable to physical activity research and behavioural medicine in general.

Study design in PARS evaluation

PARS evaluations can be broadly broken down into controlled and randomised controlled trials, uncontrolled longitudinal, cohort, and qualitative studies. Important differences between controlled and uncontrolled approaches relate to differences in validity. Controlled experimental research strives to maximise *internal validity*, controlling for the effects of confounding variables by using strict experimental controls.[2,17] Although this enables the detection of intervention effects, it is often at the expense of *external* or *ecological* validity; i.e. the extent to which the research setting replicates practice.[17] The latter is more characteristic of uncontrolled research which, in the context of PARS, often takes place within 'real life' practice settings, thus providing context specific evidence and facilitating the application of findings to practice.[2] The balance between internal and external validity is problematic because one tends to be achieved at the expense of the other. Methods and data collection procedures must have the rigour necessary to produce data of adequate quality, but unless the experimental setting can replicate practice, the degree to which findings can usefully inform practice is questionable.

Control groups

A major advantage of controlled, experimental research is isolation of intervention effects using control groups to account for changes in behaviour that are not related to the intervention. In the absence of a control group, and especially with self-reported behavioural outcomes (e.g., physical activity), it becomes difficult to attribute change to the intervention. For example, a reported increase in physical activity could simply reflect increased motivation as a result of participants' behaviour being measured (Hawthorne Effect), or that participants are reporting behaviour they perceive to be socially desirable.[18] The value of the control group has been illustrated to some extent by most PARS RCTs reporting measurable increases in physical activity in the control group.[9,14,19] This was largely thought to be a result of control groups receiving some baseline intervention (e.g. written information), which increases the chance of contamination. However, it is well acknowledged that failure to offer some kind of baseline intervention could be criticised on ethical grounds.[20]

Recruitment

The method of recruitment has been a weakness of most of the RCTs evaluating PARS,[6] which has implications for the research question.[21] Participants in uncontrolled evaluations of 'real life' schemes tend to be recruited through routine health professional referral (e.g., a GP), thus providing a faithful representation of practice. In contrast however, RCTs have involved researcher recruitment and used randomisation to reduce selection bias and obtain a representative PARS sample.[6] Randomly selecting participants, usually from a practice register (and further randomising to exercise/control groups), can reduce selection bias by the health professional or researcher.

However, it cannot be assumed that a patient who was recommended a physical activity referral by their health professional on health grounds would respond in the same way if invited by a researcher to participate in a physical activity study. This has two important implications. Firstly, it prevents insight into how health professionals might under, or over, refer certain patient groups (i.e., which members of the population PARS are accessible to). Secondly, 'volunteerism' will mean that the resulting sample is not truly representative. A common limitation in physical activity research is the reliance upon volunteers who tend to be more motivated and have often made, or are about to make, decisions regarding initiating exercise.[22] The result is often a study population in which the people most in need of the intervention (i.e., a physical activity programme), and who ideally would be targeted by referring health professionals, are under-represented. Furthermore, if randomised to a RCT control group, the motivated volunteer may well contaminate the group, or simply dropout.[22]

The controlled experimental approach has provided a limitation to knowledge about bias at the point of referral. Information is only provided about the most adherent people, with little known about the priority groups who are the target for these programmes. In this respect, a population based longitudinal study can provide a new and alternative insight. Not only has this alternative approach highlighted differences between characteristics of the 'typical' PARS participant population and that from which they were selected, but also provided information about who fails to engage at various stages of the referral process once referred.[16] Such information is highly desirable because it can feed directly into the development of practice, providing information about appropriate targeting and possible retention strategies.[16,23] Clearly, such information can only be achieved through applied research. Using participants routinely referred by health professionals confers the necessary ecological validity which, through efforts to increase internal validity, is often neglected in controlled research.

Inclusion/exclusion criteria

The use of specific inclusion and exclusion criteria in most PARS RCTs does offer some benefit. Sample homogeneity, in terms of both socio-demographics (e.g. age, gender, socio-economic position) and other potential confounders (e.g. medical conditions, body mass index, physical activity levels), is increased, and certain patient groups may be targeted. Researchers have commonly delimited recruitment to adults of middle-to-early old-age, the age group most commonly referred to PARS.[15,16] Although this approach ensures adequate numbers within a target group, such specific age limits are generally not a true representation of practice. In practice, most PARS simply impose a minimum age limit (≥16 or 18 yrs) and exclude those with high risk medical conditions.[24] The ecological validity is maximised if inclusion/exclusion criterion of the PARS under evaluation can be adhered to. This is observed in PARS research employing a longitudinal cohort design[15,16,23] and a more recent RCT.[14] Again, the advantages of the more representative sample include the potential for insight into health professional referral patterns and the relative participation and progess of different groups typically referred to PARS.

Data collection

Conducting experimental evaluations can produce more complete datasets of higher quality that potentially include a greater range of dependent variables than it is possible for practitioners (i.e., health and exercise professionals) to collect routinely. Such studies are undoubtedly useful for answering specific questions involving intervention outcomes for which data are not routinely collected, or are not feasible to collect (e.g., some physiological outcomes). Although not a common issue in PARS RCTs to date,[6] there is certainly an issue around completeness and quality

of data in applied research that takes place within a practice setting. Where exercise and health professionals are responsible for the collection of data that will ultimately be used for evaluation, there is the need to develop an 'evaluation culture'. To facilitate data collection without compromising quality, those involved with PARS must be made aware of, and become accustomed to, evaluation as an integral component of practice. Data collection processes implemented at the design phase, in addition to regular evaluation/re-evaluation, should facilitate collation of sufficient good quality data that can explain scheme outcomes and, as such, inform and improve practice. Clearly the same cannot be achieved through 'one-off' experimental evaluations that often employ complex and unsustainable data collection methods.

Which approach to evaluate PARS and other behavioural interventions?

The issues discussed make a case for using a range of methodologies to address the various types of research question that are important within PARS, physical activity research and behavioural medicine more generally. The internal validity associated with the RCT approach enables the study of *efficacy* of interventions in promoting outcome changes (e.g., in physical activity or associated physiological or psychological outcomes).[20] However, limited external or ecological validity for experimental research suggests that alternative complimentary approaches are required. Uncontrolled population-based cohort/longitudinal studies within an applied setting can provide an insight into a range of other areas. In the PARS context, these include identification of: potential bias at the point of referral (in terms of referral behaviour and patient choice); which participants attend or dropout; and differential engagement and participation with the intervention.

In addition, there is an important, and often undervalued, role for qualitative research. By exploring the perceptions and experiences of all those involved (i.e. participants, referring health professionals, exercise professionals, and scheme coordinators) qualitative research may improve understanding of exercise referral schemes **(Crone et al., 2005**). Furthermore, qualitative research might provide improved insight into *why* interventions work or do not work, *why* health professionals under/over referral certain groups, *why* some groups are more suited to a particular intervention than others, and so on. Quantitative approaches are useful for identifying and quantifying patterns and differences such as, differences in referral rate between health professionals, differential referral and attendance of certain patient groups, or the degree of subsequent behaviour change. However, to gain a better understanding of *why* such differences exist and therefore how schemes can be improved, or more appropriately targeted, requires an in-depth understanding of the referral process, participants' journeys through it, and how this is influenced by the perceptions and attitudes of health and exercise professionals involved. There has been an increasing recognition of the value of qualitative methods as a means of exploring process issues within interventions,[29,30] although there remains a dearth of good quality qualitative research within the PARS literature.

Recognition of the need for a multifaceted holistic approach to evaluating public health interventions is not new. In the PARS context, published guidance several years ago identified roles for controlled trials, audit (including applied longitudinal research), and reflective practice (qualitative by nature). Indeed, opposition to the sole use of the RCT-orientated approach to pursuing evidence-based practice exists across physical activity and public health promotion.[5,25,31-33] Yet the evidence hierarchy remains dominant. This paper highlights the importance of different evaluative approaches in this area, which are appropriate for answering different types of research question.[20] A case is made for the recognition of the differing approaches within the research

community to guide and develop evidenced-based practice. Through the use of a greater range of approaches we can be more confident that public health promotion and policy is informed by the richest available evidence base.

1. Webb D. Current approaches to gathering evidence. In: Perkins ER, Simnett I, Wright L, ed. Evidence-based health promotion. Chichester: John Wiley & Sons Ltd, 1999.
2. Evans D. Hierarchy of evidence: a framework for ranking evidence evaluating healthcare interventions. J Clin Nurs. 2003;12:77-84.
3. Grossman J, MacKenzie GJ. The randomized controlled trial: gold standard, or merely standard? Perspect Biol Med. 2005;48:516-534.
4. Riddoch C, Puig-Ribera A, Cooper A. Effectiveness of physical activity promotion schemes in primary care: a review. London: Health Education Authority, 1998.
5. Blamey A, Mutrie N. Changing the individual to promote health enhancing physical activity: the difficulties of producing evidence and translating it into practice. J Sports Sci. 2004;22:741-754.
6. Gidlow C, Johnston L, Crone D, James D. Attendance of exercise referral schemes in the UK: a systematic review. Health Educ J. 2005;64:168-186.
7. Crone D, Johnston L, Grant T. Maintaining quality in exercise referral schemes: a case study of professional practice. Prim Health Care Res Dev. 2004;5:96-103.
8. Wright Foundation. Title|. Conference Name|; Year of Conference| Date|; Conference Location|. Publisher|.
9. Harland J, White M, Drinkwater C, et al. The Newcastle exercise project: a randomised controlled trial of methods to promote physical activity in primary care. BMJ. 1999;319:828-832.
10. Kerse N, Walker S, Petrovic M, et al. Rapid responses to: The Newcastle exercise project. BMJ. 1999;320:1470-.
11. Hillsdon M, Foster C, Naidoo B, Crombie H. The effectiveness of public health interventions for increasing physical activity among adults: a review of reviews. London: Health Development Agency, 2004.
12. National Institute for Health and Clinical Excellence. Four commonly used methods to increase physical activity: brief interventions in primary care, exercise referral schemes, pedometers and community-based exercise programmes for walking and cycling: Public Health Intervention Guidance no. 2. London: NICE, 2006.
13. Sheldon TA, Sowden AJ, Lister-Sharp D. Systematic reviews includes studies other than randomised controlled trials. BMJ. 1998;316:703a-.
14. Harrison RA, Roberts C, Elton PJ. Does primary care referral to an exercise programme increase physical activity one year later? A randomised controlled trial. 2005;27:25-32.
15. Harrison RA, McNair F, Dugdill L. Access to exercise referral schemes - a population based analysis. J Public Health Med. 2005;27:326-330.
16. Gidlow C, Johnston L, Crone D, et al. Sociodemographic patterning of referral, uptake and attendance in Physical Activity Referral Schemes. 2007;doi:10.1093/pubmed/fdm002.
17. Rothwell PM. External validity of randomised controlled trials: "To whom do the results of this trial apply?" Lancet. 2005;365:82-93.
18. Shephard RJ. Limits to the measurement of habitual physical activity by questionnaires. Br J Sports Med. 2002;37:197-206.
19. Taylor AH, Doust J, Webborn N. Randomised controlled trial to examine the effects of a GP exercise referral programme in Hailsham, East Sussex, on modifiable coronary heart disease risk factors. J Epidemiol Community Health. 1998;52:595-601.
20. Dugdill L, Graham RC, McNair F. Exercise referral: the public health panacea for physical activity promotion? A critical perspective of exercise referral schemes; their development and evaluation. 2005;48:1390-1410.
21. Tai SS, Iliffe S. Considerations for the design and analysis of experimental studies in physical activity and exercise promotion: advantages of the randomised controlled trial. Br J Sports Med. 2000;34:220-224.

22. Hardman AE. Physical activity intervention studies with health-related outcomes: some issues. J Sports Sci. 1999;17:685-687.
23. Sidford A, Johnston L, Crone D, et al. Factors associated with physical activity referral uptake and participation. J Sports Sci. in press.
24. Department of Health. Exercise referral systems: a national quality assurance framework. London: The Stationery Office, 2001.
25. Hammond JM, Brodie DA, Bundred PE. Exercise on prescription: Guidelines for health professionals. Health Promot Int. 1997;12:33-41.
26. Jackson C, Bell F, Smith RA, Dixey R. Do adherers and non-adherers to a GP exercise referral scheme differ in their long-term physical activity levels? (conference communication). 1998;16:84.
27. Lord JC, Green F. Exercise on prescription: does it work? 1995;54:453-464.
28. Martin C, Woolf-May K. The retrospective evaluation of a general practitioner exercise prescription programme. J Hum Nutr Diet. 1999;12:32-42.
29. Britton A, Thorogood M, Coombes Y, Lewando-Hunt G. Quantitative outcome evaluation with qualitative process evaluation is best. 1998;316:703-704.
30. McNair F, Graham RC, Dugdill L, et al. Determining the indicators of success in Exercise Referral Schemes. J Sports Sci. 2005;23:222-223.
31. Dugdill L, Graham R. Promoting physical activity: building sustainable interventions. In: Gormley J, Hussey J, ed. Exercise in the prevention and treatment of disease. Oxford: Blackwell, 2004.
32. Puska P. Commentary: physical activity promotion in primary care. Int J Epidemiol. 2001;30:815-816.
33. Barreto ML. Efficacy, effectiveness, and the evaluation of public health interventions. J Epidemiol Community Health. 2005;59:345-346.

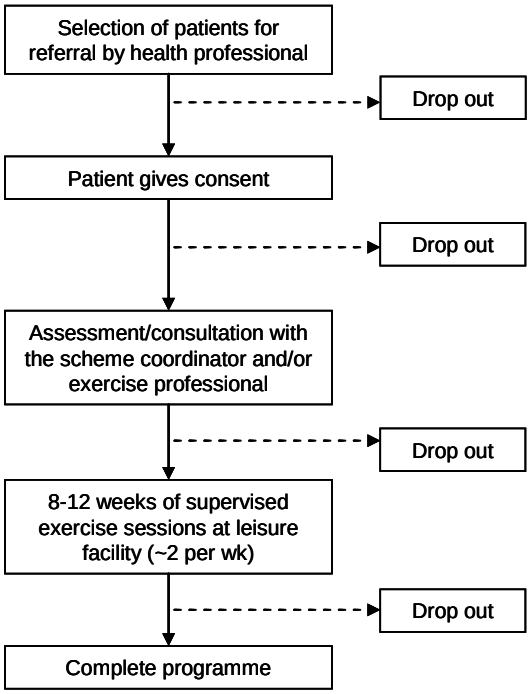

Figure 1. Physical activity referral process