

Debate

Open Access

## Pragmatic controlled clinical trials in primary care: the struggle between external and internal validity

Marshall Godwin\*<sup>1</sup>, Lucia Ruhland<sup>1</sup>, Ian Casson<sup>1</sup>, Susan MacDonald<sup>1</sup>, Dianne Delva<sup>1</sup>, Richard Birtwhistle<sup>1</sup>, Miu Lam<sup>1,2</sup> and Rachelle Seguin<sup>1</sup>

Address: <sup>1</sup>Centre for Studies in Primary Care, Queen's University, Kingston(K7L 5E9), Canada and <sup>2</sup>Department of Community Health and Epidemiology, Queen's University, Kingston(K7L 5E9), Canada

Email: Marshall Godwin\* - godwinm@post.queensu.ca; Lucia Ruhland - ruhlandl@post.queensu.ca; Ian Casson - ric@post.queensu.ca; Susan MacDonald - sem@post.queensu.ca; Dianne Delva - mdd2@post.queensu.ca; Richard Birtwhistle - birtwhis@post.queensu.ca; Miu Lam - lamm@post.queensu.ca; Rachelle Seguin - seguinr@post.queensu.ca

\* Corresponding author

Published: 22 December 2003

Received: 21 August 2003

*BMC Medical Research Methodology* 2003, **3**:28

Accepted: 22 December 2003

This article is available from: <http://www.biomedcentral.com/1471-2288/3/28>

© 2003 Godwin et al; licensee BioMed Central Ltd. This is an Open Access article: verbatim copying and redistribution of this article are permitted in all media for any purpose, provided this notice is preserved along with the article's original URL.

### Abstract

**Background:** Controlled clinical trials of health care interventions are either explanatory or pragmatic. Explanatory trials test whether an intervention is efficacious; that is, whether it can have a beneficial effect in an ideal situation. Pragmatic trials measure effectiveness; they measure the degree of beneficial effect in real clinical practice. In pragmatic trials, a balance between external validity (generalizability of the results) and internal validity (reliability or accuracy of the results) needs to be achieved. The explanatory trial seeks to maximize the internal validity by assuring rigorous control of all variables other than the intervention. The pragmatic trial seeks to maximize external validity to ensure that the results can be generalized. However the danger of pragmatic trials is that internal validity may be overly compromised in the effort to ensure generalizability. We are conducting two pragmatic randomized controlled trials on interventions in the management of hypertension in primary care. We describe the design of the trials and the steps taken to deal with the competing demands of external and internal validity.

**Discussion:** External validity is maximized by having few exclusion criteria and by allowing flexibility in the interpretation of the intervention and in management decisions. Internal validity is maximized by decreasing contamination bias through cluster randomization, and decreasing observer and assessment bias, in these non-blinded trials, through baseline data collection prior to randomization, automating the outcomes assessment with 24 hour ambulatory blood pressure monitors, and blinding the data analysis.

**Summary:** Clinical trials conducted in community practices present investigators with difficult methodological choices related to maintaining a balance between internal validity (reliability of the results) and external validity (generalizability). The attempt to achieve methodological purity can result in clinically meaningless results, while attempting to achieve full generalizability can result in invalid and unreliable results. Achieving a creative tension between the two is crucial.

## Background

Controlled clinical trials of health care interventions are either explanatory or pragmatic. [1-8] Explanatory trials test whether an intervention is efficacious; that is, whether it can have a beneficial effect in an ideal situation. These trials are often conducted in large tertiary care, referral-based, health centres on a homogenous group of patients, who have demonstrated compliance, who are likely to remain in the study, and who often have no medical condition other than the one under treatment.

Pragmatic trials measure effectiveness; they measure the degree of beneficial effect in real clinical practice. Pragmatic trials are conducted on patients who represent the full spectrum of the population to which the treatment might be applied. These patients may demonstrate variable compliance, have a number of co-morbid conditions, and use other medications. If an intervention is shown to have a significant beneficial effect in a pragmatic trial then it has been shown not only that it can work, but also does work in real life.

The explanatory clinical trial remains the standard for assessing drug efficacy and is required for drug licensure in most countries. Testing the efficacy of medications and other interventions in explanatory controlled clinical trials is critical to the progress of medicine. It is necessary, but it is not always sufficient. The overall effectiveness of a treatment is best assessed by carefully designed and well conducted pragmatic randomized trials. Pragmatic trials inform practitioners and health care planners on the most clinically effective and cost effective treatments. However there is a critical issue that needs to be considered in pragmatic trials; that is the balance between external validity (generalizability of the results) and internal validity (reliability or accuracy of the results). Whereas the explanatory trial seeks to maximize the internal validity by assuring rigorous control of all variables other than the intervention, the pragmatic trial seeks to maximize external validity to ensure that the results can be generalized. However the danger of pragmatic trials is that internal validity may be overly compromised in the effort to ensure generalizability.

We are conducting two pragmatic randomized controlled trials on interventions in the management of hypertension in primary care. We will use these two studies as examples to help the discussion around achieving a balance between internal and external validity in pragmatic trials conducted in a practice setting.

### ***Trial 1: a randomized controlled trial of the effects of home blood pressure monitoring (HBPM) on blood pressure control***

This is a randomized controlled clinical trial of HBPM (intervention group) compared to usual care (control group). Randomization is by physician. The intervention involves once a week measurements of blood pressure at home, and the reporting of these measurements to the family doctor at each office visit. The primary outcome is blood pressure control measured by 24-hour ambulatory blood pressure monitoring (ABPM) at entry, and after 6 and 12 months. Secondary outcomes include the number of office visits for hypertension, change in lifestyle behaviors by patient, lifestyle counseling by the physician, quality of life of the patient, medication compliance, and intensity of treatment. (Figure 1)

### ***Trial 2: an intensive scheduled management (ISM) strategy for increasing blood pressure control in patients in primary care***

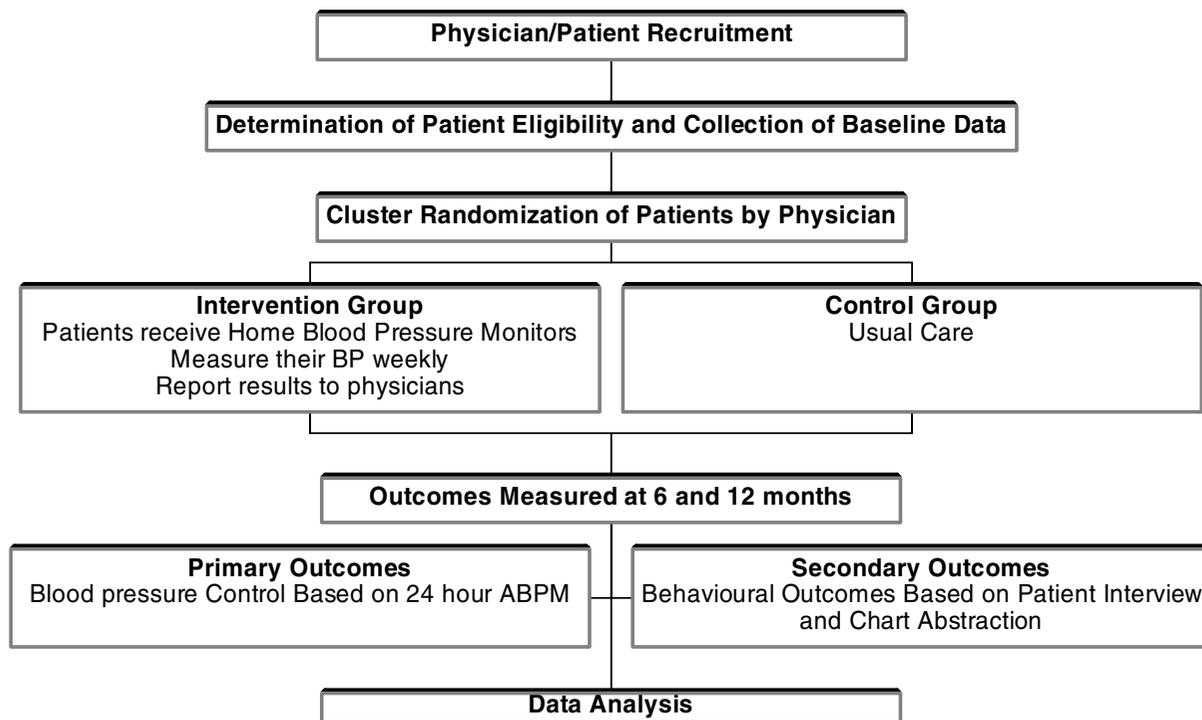
This is a randomized controlled trial, in the primary care setting, comparing intensive scheduled management of hypertension (aggressive achievement of target blood pressure over 16 weeks) with usual management of hypertension. Randomization is by physician. The primary outcome is blood pressure control on 24-hour ABPM; it is measured at 16 weeks to determine the short-term effect of the intensive approach to therapy and again at 1 year to determine if the blood pressure control is maintained. Secondary outcomes include patient quality of life, physician compliance with the intensive protocol, patient compliance with medication, and adverse effects. (Figure 2)

## Discussion

### ***Inclusion/exclusion criteria***

A hallmark of pragmatic trials is that the participants reflect the population for which the treatment is intended. Exclusion criteria are kept to a minimum. In both our studies the patient population (inclusion criteria) consists of "adults, age 18 years and older, who have been diagnosed with essential hypertension and who have not reached their target blood pressure level." Exclusion criteria common to both studies are patients with "a diagnosis of secondary hypertension; pregnancy; hypertension management primarily by a consultant; inability to provide informed consent". Particular to the HBPM study is the exclusion criteria "a physical or mental disability that precludes use of a home blood pressure monitor".

All adults with uncontrolled essential hypertension constitute the target population. Patients are not excluded for co-morbid conditions or if medication is already being used. We accept the family physician's diagnosis of essential hypertension without doing any confirmatory tests, because this is what happens in practice, and because



**Figure 1**  
Schematic for the HBPM study

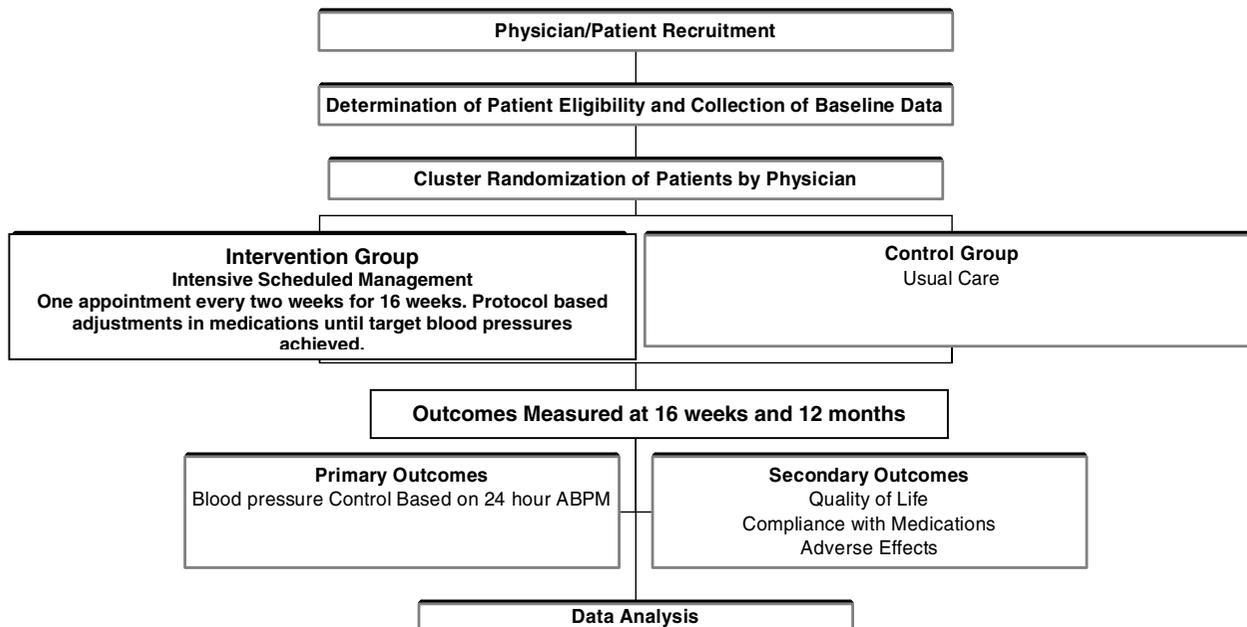
essential hypertension is, by far, the most likely diagnosis in primary care. Children and pregnant women are excluded because they are not part of the population for which the interventions are intended. And finally, since this is a trial of the primary care management of hypertension, patients whose blood pressure is managed by hypertension specialists are excluded.

**The interventions**

For the HBPM study the intervention consists of the provision of a home blood pressure monitor to each patient in the intervention group, patient instruction on its proper use, a recording frequency of at least once a week, and the reporting of the results of the HBPM to the physician at each office visit. The patient may use the HBPM more often than weekly if they wish; the frequency of visits to the physician is determined by the physician and the patient; and the physicians may choose to use the HBPM information presented to them by the patients in the manner they feel most appropriate. We postulate several

mechanisms by which the intervention may improve blood pressure: bringing in the HBPM measurements to the physician may lead to enhanced discussion between the physician and patient about the value of controlling high blood pressure and the lifestyle changes that can lead to improved blood pressure control and health outcomes; the patients' knowledge of their blood pressure may influence them to adhere to their medication regime and make appropriate life style changes; it may also influence the patient and the physician's management of uncontrolled hypertension, by increasing the intensity of treatment.

For the ISM study, the intervention includes both the eight visits (every two weeks for 16 weeks) to see their family doctor and the adjustment of medications at each visit to achieve target blood pressure levels. This study is designed to test the premise that guidelines, to be effective in primary care, need to be operationalized for the practitioner. To this end we have devised a protocol where medications are initiated at the recommended starting dose



**Figure 2**  
Schematic for the ISM study

and increased by one recommended increment before adding the next drug. The goal is to increase the medications over the 16-week period such that, if necessary to reach the target blood pressure level, a patient is on a medium dose of three different antihypertensive agents (medium dose is being defined as one recommended increment above the recommended starting dose).

The pragmatic nature of this trial allows the choice of medications to be decided by the physician, but it is recommended that guideline choices be followed. As well, in keeping with real life and generalizability, the physician, when making decisions about increasing a patient's drug regime intensity, will be able to take into account medication compliance, side effects, and life events that may cause transient elevations of the patient's blood pressure. (Figure 3)

**Recruitment and randomization**

The intervention in both trials is dependent on physician behavior. There is a risk that if an individual physician had patients in both the intervention and control groups, the modification of behaviour caused by the intervention would affect the management of patients in the control

group. To prevent this we use cluster randomization; the physician is randomized to the intervention or control group rather than the individual patients. All patients of a given physician are then enrolled into the group to which their physician has been assigned. Because we want the research assistant who collects the enrollment data to be blinded to the patient group allocation, we have to enroll all patients of a given physician before randomizing that physician. This means we cannot enroll a physician's patients in an ongoing manner as they become available but have to identify them all at one time, seek consent, and do all the enrollment work before assigning them to a group.

**Blinding**

As is often true for pragmatic trials, it is not possible to blind either the physicians or the patients in either of the studies. However attempts have to be made to 'blind' as much of the process as is possible. The main outcomes are based on 24-hour ambulatory blood pressure recordings that provide an independent measurement of blood pressure; the biostatistician analyzing the data will be blinded to the group assignment; and enrollment data will be

VISIT #1

Pragmatic controlled clinical trials

1. BP #1: \_\_\_\_\_ BP #2: \_\_\_\_\_ BP target achieved (2<sup>nd</sup> reading)?  YES  NO
2. Do you have reason to believe today's BP is not representative of patient's usual BP?  
(e.g. acute stressful situation, life crisis, home BP measurement, etc.)  YES  NO
3. Does the patient state that they have taken their antihypertensive medications 80% of the time?  YES  NO
4. Is the patient having **significant** or unusual side effects from the medication?  YES  NO

**ACTION:**  
Both diastolic and systolic BP targets are reached?  
 NO  YES →  I will recheck the BP in two weeks.

↓  
If NO (one or both targets are not reached), did you increase medication?  
 YES  NO

**IF MEDICATION INCREASED:**

- Increased an existing antihypertensive agent by one recommended increment. (see table of drug dosages)
- Started a new antihypertensive agent at the recommended initial dose.

**IF MEDICATION NOT INCREASED:**

- Patient not compliant with current medications. Compliance discussed and patient encouraged to take medications as prescribed.
- BP not representative due to stress etc. Will recheck BP in 2 weeks before changing medication.
- Medication side effects.
  - Medication discontinued; new medication started.
  - Medication discontinued; new medication NOT started.
  - Medication decreased; new medication started.
  - Medication decreased; new medication NOT started.
  - Medication not changed. Will wait and see if side effects settle.

**ANTIHYPERTENSIVE MEDICATION REGIME AT END OF THIS VISIT**

DRUG#1:  
 DRUG#2:  
 DRUG#3:

Please see this patient again in two weeks

**Figure 3**  
Sample page from the visit form of the ISM study

collected by the research assistant prior to the group assignment.

**Contamination**

Our main strategy for reducing contamination in both studies is the use of cluster randomization so that all enrolled patients under a given physician's care are either in the intervention group or the control group. Contamination is still possible, especially for the HBPM study. We hope to reduce it by asking patients in the control group not to use a home blood pressure monitor and ask control group physicians to refrain from recommending HBPM to their patients during the 12 months of the study. At the time of the 24-hour ABPM, patients will be asked to complete a questionnaire regarding blood pressure measurement outside the physician's office. Regarding contamination in the ISM study, it is possible that control group physicians will take it upon themselves to start an

intensive scheduled approach to management of blood pressure, but we believe this is unlikely.

**Compliance with the intervention**

To a great extent, compliance with the intervention is one of the most important outcomes of pragmatic trials. Unlike explanatory trials where compliance with the intervention must be ensured in order to know that the intervention can work, in pragmatic trials compliance with the intervention is measured as an outcome. If physicians and/or patients do not complied with the intervention, then it doesn't matter that it can work in the ideal world, because in the real world it doesn't. Of course, reasonable attempts are made to encourage compliance with the intervention but these must not go beyond what could be expected in the normal course of practice. We are assessing compliance with home blood pressure monitoring by having patients record their activity in a diary and

bring it to their physician. We are assessing compliance with the intensive scheduled management of hypertension intervention by having the physician record whether they followed the protocol and if not their reasons for it. (figure 3)

#### **Follow-up and outcomes measurement**

In both studies the patients are followed for one year. Data is collected at baseline and then again at 6 and 12 months for the HBPM study and at 16 weeks and 12 months for the ISM study. There are three methods of outcome assessment: 24 hr ambulatory blood pressure monitoring, patient interview, and chart abstraction. For both studies the mean daytime systolic and diastolic blood pressures from the 24 hour ABPM will be used as the primary outcome. Patient interview/questionnaire completion will be used to collect information on quality of life (SF36), lifestyle risk factors (using the Short Lifestyle Indicator Questionnaire or SLIQ which we have developed for the HBPM study), and patient reporting of compliance. For the HBPM study only, the intensity of treatment, the degree to which lifestyle issues were addressed by the physician, and the frequency of visits for hypertension will be assessed chart abstraction

#### **The struggle between external and internal validity**

The need to balance external validity and internal validity is ever-present in pragmatic randomized controlled trials. These two trials provide examples of the issues that are often encountered and how they can be addressed. External validity, or generalizability, was addressed in both studies by having very few exclusion criteria. Hypertension patients are a heterogeneous group, and patients with multiple and varying co-morbidities and medication usage were all included. One of the strengths of a randomized trial is that the diversity of the study population is distributed between the two groups and thus helps maintain internal validity. In addition to minimizing the exclusion criteria, we enhanced external validity of the HBPM study by allowing the patients to use the home blood pressure monitor as often as they wanted with the requirement that they use it at least once a week. This ensures that the minimum intervention occurs and allows for the pragmatic fact that patients will use the machine at variable frequencies; it reflects what happens in practice when a patient obtains a home BP monitoring device. In a similar vein, in the HBPM study physicians were free to manage the hypertension and to use the results of the home blood pressure monitoring as they saw fit. In the ISM study, while the management regime is much more regulated, the physician has the freedom to choose a course of action different from the protocol if it is felt to be in the best interest of the patient. This is fundamentally pragmatic and has to be permitted if the results are to be accepted as generalizable. Again, conservation of internal validity is

achieved through the random distribution of the different management approaches between the intervention and control groups. While this assumption of 'equal distribution of differences by randomization' is not absolute, it remains the best way to achieve internal validity in a trial that maximizes external validity. It underscores the importance of randomization in pragmatic trials.

Internal validity is the extent to which differences between the intervention and control groups can be confidently attributed to the intervention and not due to some alternate explanation. To achieve this, confounding factors and bias must be reduced to a minimum. In a pragmatic trial this is more difficult because of the competing demand of external validity. We worked to increase the internal validity of the study using multiple strategies. Cluster randomization was used to decrease the likelihood of contamination bias by ensuring that physicians did not apply strategies used in the intervention patients to control patients. To further prevent contamination bias in the HBPM study we are enlisting the cooperation of the patients in the control group by asking them not to begin using home blood pressure monitors during the study period. Unable to blind patients or physicians, we are using the techniques of baseline data collection before randomization, automation of outcome assessment, and blinding of the data analysis to decrease observer and assessment bias.

It is possible that our efforts to overcome the effects of pragmatism on internal validity will be only partially successful. Patients in the control group of the HBPM study may start using home blood pressure monitoring despite asking them not to; or they may start checking their blood pressure in pharmacies. We can deal with this only by measuring it (asking them) and attempting to control for its effect in the analysis. We also do not know ahead of time whether control group physicians in the ISM study will increase the intensity of treatment for their patients because they have volunteered to be in a study and they have heard in general terms about the intervention. While we request that they do not change their usual practice during the study, we will measure changes and control for it in the analysis. The possibility that randomization will fail to distribute factors equally between groups and the occurrence of contamination due to lack of true blinding must be considered when planning the sample size and analysis of a pragmatic trial. Hence a plan for analysis must be in place where there is sufficient sample size to allow for regression analysis to determine the independent effects of the intervention taking potential confounders into account.

Table 1 (see Additional file 1) provides a summary of the differences between explanatory and pragmatic trials. To a

great extent the conduct of pragmatic trials is a recent phenomenon. While one of the earlier descriptions of pragmatic versus explanatory trials was by Schwartz in 1967 [7] and again by MacRae in 1989 [8], most of the published editorials considering pragmatic trials as a methodology have been since 1998 [1-6]. These commentaries do not differ substantially from this article in tone and concept. However we have extended the information in the literature by addressing how the various issues can be dealt with when designing a pragmatic trial and by using two examples to shed further light on the process. The recent increase in published pragmatic trials indicates the relevance of this methodology, particularly in primary care. A Medline search using pragmatic as a title word and limited to randomized controlled trials yielded 34 articles reporting on pragmatic clinical trials. All 34 were published since 1995 and 26 of them were published since 2000. Of the 34 pragmatic trials 13 (38%) were conducted in primary care practice settings. Investigators using this method understand its strengths and weaknesses.

### Summary

Clinical trials conducted in community practices present investigators with difficult methodological choices related to maintaining a balance between internal validity (reliability of the results) and external validity (generalizability).

To maintain generalizability i) exclusion criteria must be kept to a minimum, and ii) physicians and patients (who actually deliver the intervention in the course of regular physician patient encounters) must be given a large degree of freedom in making choices about the delivery of the intervention in order to keep it pragmatic and to ensure that patients are not harmed.

To maintain internal validity i) randomization is critical, ii) cluster randomization is often needed to deal with contamination issues, and iii) data collection and data analysis must be blinded whenever possible since standard double-blind strategies are often not possible.

Compliance is not something you necessarily struggle to maintain but rather something you measure as an outcome. Lack of compliance in the 'real world' frequently renders an efficacious intervention ineffective.

### List of Abbreviations

ABPM = ambulatory blood pressure monitor

HBPM = home blood pressure monitor

ISM = Intensive scheduled management

SF36 = Short Form 36

### Competing Interests

None declared.

### Authors' Contributions

MG wrote the initial draft of the manuscript. All authors contributed to the design of the studies which included discussion around the issues dealt with in the manuscript. All authors reviewed the manuscript, provided editorial comment, and approved the final draft.

### Additional material

#### Additional File 1

*Comparison of Design Issues in Explanatory and Pragmatic Trials. This table provides a side by side comparison of the design features and issues in explanatory and pragmatic trials*

Click here for file

[<http://www.biomedcentral.com/content/supplementary/1471-2288-3-28-S1.doc>]

### Acknowledgements

This research is being supported by the Heart and Stroke Foundation of Ontario Grant Numbers NA4882 and NA4884

### References

1. Pringle M, Churchill R: **Randomized controlled trials in general practice.** *BMJ* 1995, **311(7017)**:1382-1383.
2. Helms PJ: **'Real world' pragmatic clinical trials: what are they and what do they tell us?** *Pediatr Allergy Immunol* 2002, **13(1)**:4-9.
3. Hotopf M, Churchill R, Lewis G: **Pragmatic randomised controlled trials in psychiatry.** *Br J Psychiatry* 1999, **175**:217-223.
4. Roland M, Torgerson DJ: **Understanding controlled trials: What are pragmatic trials?** *BMJ* 1998, **316(7127)**:285.
5. Moore H, Summerbell C, Vail A, Greenwood DC, Adamson AJ: **The design features and practicalities of conducting a pragmatic cluster randomized trial of obesity management in primary care.** *Stat Med* 2001, **20(3)**:331-40.
6. van der Windt DA, Koes BW, van Aarst M, Heemskerk MA, Bouter LM: **Practical aspects of conducting a pragmatic randomised trial in primary care: patient recruitment and outcome assessment.** *Br J Gen Pract* 2000, **50(454)**:371-4.
7. Schwartz D, Lellouch J: **Explanatory and pragmatic attitudes in therapeutic trials.** *J Chronic Dis* 1967, **20(8)**:637-48.
8. MacRae KD: **Pragmatic versus explanatory trials.** *Int J Technol Assess Health Care* 1989, **5(3)**:333-9.

### Pre-publication history

The pre-publication history for this paper can be accessed here:

<http://www.biomedcentral.com/1471-2288/3/28/prepub>