Methods for trials of interventions to enhance patient adherence to medication prescriptions, based on a systematic review of recent randomized trials

Haynes RB1,2, Jeffery RA3, Keepanasseril A1, Wilczynski NL1, Navarro-Ruan T1, Mustafa RA4, Nieuwlaat R1 and the Patient Adherence Review (PAR) Team5

1Health Information Research Unit, Department of Clinical Epidemiology and Biostatistics, McMaster University, Canada
2Department of Medicine, McMaster University, Canada
3Faculty of Medicine, Dalhousie University, Canada
4Departments of Medicine and Biomedical & Health Informatics, University of Missouri-Kansas City, USA
5Nicholas Hobson, Thomas Agoritsas, Niraj Mistry, Alfonso Iorio, Susan Jack, Bhairavi Sivaramalingam, Emma Iserman, Dawn Jedraszewski and Chris Cotoi

Abstract

Background: Low patient adherence to prescribed medications is very common and likely the most important barrier to the implementation of effective health care. Controlled trials of interventions to improve adherence have not found consistent success for any type of approach, but the research methods of these trials can lead to biased and imprecise testing, making it difficult to distinguish poor interventions from poor testing.

Purpose: To outline key methodologic principles for testing adherence interventions and compare these with the methods used in recent randomized controlled trials from a systematic review.

Methods: All recent trials included in an update of a Cochrane review of interventions to assist patients to adhere to prescribed medication were assessed for selection of participants, measurement of baseline adherence, data analysis according to baseline adherence, allocation to study groups, description of interventions, measures of medication adherence and clinical outcomes, and blinding.

Results: Of 109 new trials included in the systematic review update, 51% measured baseline adherence, 5% specifically recruited non-adherent participants, and 10% reported their final results according to baseline adherence. Concealment of allocation to study groups was unclear in 65% of trials. Subjective measures of adherence were used in 48% of studies and 68% did not report on clinically important outcomes. Only 39% of studies reported on adverse effects of interventions and just 11% reported on incremental costs. We recommend remedies for these methodologic limitations.

Conclusion: Recent trials of interventions to assist patients to take prescribed medications fail on key methodologic practices for fair and precise testing. This may be a major reason for the failure to identify effective ways to improve patient adherence and health care outcomes for self-administered treatments. Many of the failings are remediable.

Introduction

Low patient adherence with prescribed self-administered medication regimens is a distressingly common problem that sharply limits the success of medical care. Many patients drop out of care [1] and among those who remain under medical supervision, typical medication adherence rates average about 50-60%, with a wide range among individuals, and within individuals from time to time [2]. In our recently updated Cochrane review of higher quality trials of adherence interventions [3], no consistent evidence emerged that adherence can be substantially improved by practical and ethical means. While the findings to date could simply reflect ineffectual interventions, they may also be due, at least in part, to deficiencies in the design, execution and analysis of trials of interventions to improve adherence.

The imperative to improve the methods of adherence research grows ever greater for 2 key reasons. First, there is no need to do poorly designed and executed research [4] that fails to build on previous studies [5]. Second, with the development of increasing numbers of efficacious medical treatments (i.e., those that do more good than harm when taken), and the ongoing problem of low adherence, the gap continues to widen between health benefits that could be achieved and those that actually are. The dramatic 149% increase in the number of adherence intervention trials we observed in the 6 years covered by the most recent update [3] reveals intense interest on the part of researchers.

Unfortunately, these trials revealed many flaws in design, conduct and reporting that undermine their credibility and conclusions. For...
example, for 50 selected trials, Jeffery et al. [6] showed problems with the measurement of adherence and the selection of participants. In this paper we will describe problems in the state of adherence trial methodology based on recently published trials, and key features of methods that should be considered in future trials. Our objective is to contribute to improving the quality of adherence intervention research. Improvements in research methods would facilitate the testing of innovations for enhancing adherence and could lead to the validation of efficacious, cost-effective, and acceptable interventions to enhance adherence and improve clinical outcomes, potentially transforming the effectiveness and efficiency of health care [7].

Methods

We analyzed methodologic features of randomized clinical trials (RCTs) included in a recently published systematic review of interventions to enhance patient adherence with prescribed medications [3], and formally assessed potential sources of bias, using the Cochrane ‘risk of bias’ tool [8]. We included unconfounded RCTs of interventions to improve adherence with prescribed medications, measuring both medication adherence and clinical outcome, with at least 80% follow-up of each group studied and, for long-term treatments, at least 6 months follow-up for studies with positive findings at earlier time points. Thus, these trials all met basic criteria for methodologic merit and represent the best trials to date.

Methods for assessing patient recruitment methodology and measures of adherence were described in detail in a recent report [6] for 50 selected trials. Scale dimensions derived from the literature review included validity, reliability, objectivity, unobtrusiveness and longitudinality (Table 1). Here, this analysis is extended to 109 trials added to the review update, which includes all eligible trials retrieved in multiple bibliographic database searches for the period from February 1, 2007 to January 11, 2013 [3]. Except for historical comparisons (for example, success in avoiding bias over time), we focused on these 109 recently conducted RCTs to assure that our findings and recommendations are consistent with current practices in adherence intervention trials. This decision is supported by a modest reduction in the risk of bias over time among all 182 trials in the 35 year span of publication in the cumulative review [3] (Pearson correlation 0.156, 2-tailed p=0.035).

In our approach, we begin with methodological principles for features of design for intervention trials intending to test interventions to improve adherence with prescribed medications, and then assess the extent to which trials in the review observed these principles. We end with prioritized recommendations for improvement of adherence research.

Results

Participant recruitment

Testing of interventions to improve adherence is most efficiently based on selecting patients who are showing low levels of adherence to begin with, or who have a prior track record of low adherence. Two key reasons for this are as follows. First, patients whose adherence is high may not need assistance to sustain this state; use of scarce resources for a “non-problem” is wasteful and potentially harmful, both directly (costs for the providers and risks for the patient) and indirectly (lost opportunities for patients in greater need). Second, including patients who are highly adherent volunteers substantially limits the likelihood of achieving important benefits (“ceiling effect”) [9] and accordingly reduces the power of the investigation, resulting in the need to recruit

Table 1. Measurement qualities of all types of adherence measures from a sample of randomized controlled trials of interventions to increase patient adherence published since 2007.

| Measurea | Frequency of use in 109 trials (n) | Median validity score (max 2)b | Median reliability score (max 2)c | Median objectivity score (max 2)d | Median unobtrusive score (max 2)e | Median longitudinality score (max 1)f | Median total score (max 9) |
|----------|----------------------------------|------------------------|-------------------------------|---------------------------------|-------------------------------|-------------------------------|------------------------|
| Attendance | 3 | 1 | 1 | 0 | 1 | 1 | 1 | 4 |
| Clinician judgment | 1 | 1 | 1 | 0 | 1 | 0 | 3 |
| Direct observation | 1 | 1 | 1 | 2 | 0 | 0 | 4 |
| Drug concentration in body fluid | 4 | 1 | 1 | 2 | 1.5 | 1 | 6 |
| Medication Event Monitoring System | 33 | 1 | 1 | 2 | 0 | 1 | 5 |
| Pharmacy refill record | 19 | 1 | 1 | 2 | 2 | 1 | 7 |
| Pill count | 16 | 1 | 1 | 2 | 0 | 1 | 5 |
| Self-report questionnaire | 61 | 2 | 2 | 0 | 0 | 0 | 4 |
| Self-report diary | 5 | 1 | 1 | 0 | 0 | 1 | 3 |
| Self-report interview | 18 | 1 | 1 | 0 | 0 | 0 | 2.5 |
| Therapeutic response | 2 | 1 | 1 | 2 | 1.5 | 1 | 6.5 |

a Measures are not meant to be compared against each other as the nature of each quality category (i.e., validity, etc.) is not comparable between measure types, as criterion standards can differ. Some measures are from the same study as some studies used multiple measures of adherence.

b Validity is the extent to which a measure accurately assesses what it purports to, generally in relation to a criterion standard. This scale ranged from 0 to 2, where 0 indicated that the trial stated the measure was not valid; 1 indicated it was unclear in the trial if the measure was valid, and 2 indicated the measure was stated to be valid in the trial.

c Reliability refers to a method’s reproducibility when applied repeatedly during the same state of adherence. This scale ranged from 0 to 2, where 0 indicated that the measure was not reliable, 1 indicated it was unclear in the trial if the measure was reliable, and 2 indicated the measure was stated to be reliable in the trial.

d Objectivity refers to a method that is independent of the perception of the observer. This scale ranged from 0 to 2, where 0 indicated that the measure was subjective without appropriate blinding, 1 indicated the measure was subjective with appropriate blinding of the observer to the study group of the participants, and a score of 2 indicated the measure was objective.

e Unobtrusiveness refers to whether the participant was unaware of the measure, reducing the chance of deceit or evasion by the participant. This scale ranged from 0 to 2, where 0 indicated the measure was obtrusive or obvious to participants, 1 indicated it was unclear based on trial information whether the measure was obtrusive or not, and 2 indicated the measure was unobtrusive to participants.

f Longitudinality refers to whether the measure provides adherence findings over a period of follow up. This scale ranged from 0 to 1, where 0 indicated the measure did not assess adherence over a period of more than one week. A score of 1 indicated that the measure was able to assess adherence for a period of over one week.
many more participants or risk the consequences of indeterminate trial results (i.e., no difference in outcomes, with a low power to detect a difference).

Our analysis of 109 newly added RCTs in the Cochrane review update [3] found that just 5 (5%) recruited participants based on their adherence status, that is, recruited only non-adherent patients into the trial.

Measuring adherence at baseline and stratifying according to baseline adherence

In drug efficacy trials, assessing patient adherence before randomization (“faintness of heart period”) is a common practice, to select patients with high adherence [10]. The same procedure can be used in trials testing the efficacy of adherence interventions, but in this instance the purpose is to select, or at least identify, nonadherent patients for recruitment. Measuring adherence at baseline should ideally be done after prescription of the targeted medication(s) but before randomization to adherence interventions. In the absence of a run-in period, previous adherence rates are strong predictors of future adherence [11].

Even if the decision is made to include adherent as well as nonadherent patients, assessing adherence at baseline and stratifying randomization by adherence level will more likely result in balanced groups for this variable and analysis can be stratified to determine whether initially nonadherent patients benefit, as well as whether the intervention helps initially adherent patients to maintain high adherence. Having data on individual change in adherence during the course of the trial also assures maximal power to detect differences in adherence between study groups being compared.

Of the 109 trials in the Cochrane review update, 56 trials (51%) measured baseline adherence with medications that had been prescribed before the study began, but only 11 (10%) reported their final results according to baseline adherence levels. Thirty-eight trials (35%) tested an intervention for a newly started medication, with no measure of baseline adherence.

Reporting the intervention

Many studies do not clearly report important details of the intervention they tested, making it difficult for practitioners to emulate the intervention, or investigators to assess potential risk of bias or replicate the findings. Hoffmann et al. [12] have provided detailed recommendations for how to improve completeness of reporting interventions and intervention replication, using a 12-item checklist. Applying this checklist to the 109 adherence trials revealed a number of key problems. A descriptive name of the intervention was typically available, but an intervention rationale was explicitly supported by a theoretical framework in just 42 RCTs (39%) and only 36 (33%) involved key stakeholders in study design. Any resources or procedures that were not considered part of the intervention, but could have an effect on adherence or the clinical outcome of interest, are co-interventions and thus potential confounders. For example, in addition to applying an adherence intervention, nurse educators in Piette et al. [13] were also allowed to recommend dose changes to prescribers based on their intensive monitoring of patients, but it was not described how often this occurred and it was not possible to estimate how this might have obscured the adherence intervention’s effect on clinical outcomes.

Further, it is very important to clarify who applied the intervention, especially regarding their expertise, study training, or whether they were specifically hired for the study. This could improve intervention fidelity, while interpretation of the findings will depend on whether the purpose of the study was “explanatory” with specifically employed personnel ("can this intervention work, under relatively ideal circumstances?") or “pragmatic” with routinely available personnel (“does this intervention work under usual clinical circumstances?”) [14]. Details on timing and intensity of the intervention can be used to determine if the effect can be replicated with available resources, and to estimate whether different timing and intensity could lead to a similar effect. Most of the studies in our review (104 studies, 95%) reported sufficient information to determine when the adherence intervention was started or applied (eg, at medication initiation, following a clinical event, or without these triggering events).

Tailoring an intervention to the patient’s characteristics, preferences, or developing clinical status can improve the efficacy of the intervention [9]. However, if this is planned in a study, but it is not reported how tailoring was actually done, it will be hard to replicate the intervention and others applying the non-tailored version of the intervention might not achieve the same benefit. In our review, we found that just 14 studies (13%) assessed the reasons for low medication adherence at baseline, which could be used for both general tailoring of the intervention to group-level characteristics, or to individuals. It is conceivable that adherence interventions were modified during these studies, as they are targeting a complex and moving target. Reporting and explaining such changes will help to understand the intervention effects, as well as aid implementation of the optimal iteration of the intervention.

Finally, evaluating the degree of adherence to the intervention and fidelity of using the intervention will help to determine how feasible it will be to deliver the intervention beyond the study setting. For this, investigators must describe how and by whom intervention adherence and fidelity were assessed, and what was planned to enhance these [11], as well as what the actual findings were [12]. We found that 99 RCTs (91%) reported the level of participation with the intervention.

Allocation to study groups

Random assignment of participants to study groups with concealed allocation is widely accepted as the least biased way to form study groups [15]. All studies in our review employed random assignment, as this was a condition for inclusion. However, concealment of allocation until the exact time of randomization for individual patients was unclear in 71 trials (65%) and not done in 3 (3%). Concealment of allocation is best done through central randomization once the patient has agreed to participate, or by using sequentially numbered, opaque, sealed envelopes that remain unopened until the patient’s consent to participate has been obtained.

Bias in group assignment can be further reduced, especially in smaller studies, by randomization within strata, if there are key characteristics of patients that could affect the outcome. Adherence is such a variable and stratification according to baseline adherence, then randomization with blocking (eg, after 2, 4 or 6 patients) will improve balance.

If the intervention is to be administered by health professionals directly involved in the care of the patients, then consideration should be given to cluster randomization [16], with the unit of randomization being the health care practitioner or the clinical team caring for the patient. This reduces the risk of the control group participants also receiving one or more components of the intervention.
Medication event monitoring, with electronic pill container caps, can recruit patients into a study [19]. Self-report has validity compared to other methods that can be readily obtained as part of the initial consent to participate at the beginning of the study. No ideal method exists, but monitoring prescription refills is readily obtained as part of the initial consent to participate at the beginning of the study. These are not “fatal” flaws in this type of research, as long as an objective measure of adherence is used.

Those who administer the adherence intervention cannot be blinded even using distracters, although caregivers could be blinded if they are not implementing the intervention themselves. Outcome assessors, data analysts and authors are also candidates for blinding, and measures should be taken to do so, especially for outcome assessors and analysts [18]. A desirable way to do this for outcome assessments is to use objective measures of adherence (eg, prescription refill monitoring) and objectively measured outcomes (eg, objectively defined patient important events or at least objective surrogates such as glycated hemoglobin for diabetes, or laboratory values for dyslipidemia). An alternative would be to have outcome assessors for physiological assessments, such as blood pressure, who are unaware of the study group of patients.

Blinding of the data analysts should also be considered. The best method is to blind the data analysts by coding the study groups until the pre-specified analyses have been completed.

In the 109 new studies of the review, 9 (8%) reported blinding of patients or any attempt to disguise the adherence intervention, and just 7 (6%) reported blinding of those who administered the intervention. These are not “fatal” flaws in this type of research, as long as an objective measure of adherence and an objective or blinded clinical outcome assessment are used.

Monitoring adherence

Many measures of adherence exist but few meet the criteria of reliability, validity, objectivity, unobtrusiveness and longitudinality [6], as well as being applicable to any medication regimen and affordable. No affordable ideal method exists, but monitoring prescription refills meets most criteria, and is increasingly available through electronic records. It requires specific patient consent which, in our experience, is readily obtained as part of the initial consent to participate at the recruitment phase of a study [19]. Self-report has validity compared with more objective measures, but low sensitivity (around 50%) [20]. Medication event monitoring, with electronic pill container caps, can provide detailed records, but is expensive and impractical for usual care settings, and may be “reactive”, increasing adherence rates for all study groups, making a difference between groups more difficult to detect.

Overall 163 measures of adherence were included in the analysis from 109 RCTs (Table 1). The median quality of these measures was 4 on a scale from 0 to 9. Only 14 studies (13%) used a measure of adherence that was valid, reliable, objective, unobtrusive and longitudinal. Fifty-two studies (48%) used only subjective adherence measures.

Measuring clinical outcomes

Ethical standards for adherence research require that attempts to improve adherence should be judged by their clinical benefits [21]. The outcome measures reported in adherence trials are determined by the clinical condition in question, and range from patient-reported subjective outcome measures (eg, health-related quality of life, pain, depression scales) and intermediate biological outcomes (eg, HbA1c levels for diabetes, CD4 levels for human immunodeficiency virus) to clinical events (eg, death, stroke, emergency department visits, hospitalization). Intermediate biological or physiological measurements can also have a low risk of bias but have limited correlation with patient-important outcomes, are sensitive to measurement techniques and, if not objective (eg, blood pressure), require blinded assessment. Patient reported measures, on the other hand, although very appropriate in many conditions, are associated with high risk of bias, especially in adherence intervention research, where blinding of participants is impossible or at least seldom done. From a methodological standpoint, clinical events are preferable because they are important to patients and often objectively verifiable, thus less likely to be biased.

Of the 109 new intervention trials, just 34 (31%) reported patient-important clinical events. Most adherence intervention studies, 73 (67%), reported biological outcomes and 71 (65%) provided patient reported outcomes. Some studies collected care-giver-reported symptoms, which can also be prone to bias. Another problem limiting the comparison of interventions targeted even for the same disease is the wide variety of outcome measures used. For instance, while studying the effectiveness of interventions to increase adherence to asthma medications, Otsuki et al. [22] Charles T et al. [23] Chan DS et al. [24] and van der Meer V et al. [25] used 6 different clinical outcomes (4 patient-reported measures and 2 biological measures).

Monitoring for adverse effects and costs

Adherence interventions are like any other type of clinical intervention, with potential for good and harm, and, inevitably, invoking costs. For example, if adherence increases with a medication that has adverse effects, such as insulin, one can expect that the frequency of adverse effects, such as hypoglycemia, will increase. Quality of life may also be adversely affected as well. Further, for those wishing to implement the findings of successful trials of adherence interventions, data must be provided concerning the incremental costs/resources consumed.

In the review of 109 adherence intervention trials, 43 (39%) reported on adverse effects and 12 (11%) reported on incremental costs. Only 7 studies (7%) reported on both adverse events and incremental costs. Of the studies that reported on both adverse events and incremental costs, 6 studies simply reported that adverse events were checked but found to be non-significant, while one published a separate paper specifically on the effect of adverse events on antidepressant non-adherence [26]. Likewise, 5 studies reported incremental costs within the paper, while Fortney et al. [27] and Pyne et al. [28] published supplementary papers...
on, respectively, costs and effects of a rural telemedicine program, and a collaborative care program for depression.

Discussion and priorities

We matched key methodologic standards for trials of adherence interventions with the methods reported by recent trials in this field and found very low compliance of studies with most standards. We used selection criteria to include just the best designed studies available, so our findings represent just the “tip of the iceberg” for problems with adherence research. We conclude that poor research methods may be an important reason why there has been little or no progress in finding and validating effective interventions to promote patient adherence. This is particularly troubling because low patient adherence is arguably the most important barrier to realizing the benefits of all self-administered interventions, regardless of their efficacy among people who do adhere (Table 2).

Priorities for sound methods for adherence research appear in Table 2. Investigators should consider recruiting only non-adherent patients. These patients will be more difficult to recruit than adherent patients, for both logistical reasons (they will be less likely to attend appointments) and personal reasons (they may simply decline to take their medications or participate in a study that is focused on increasing their adherence to medication). Thus, investigators must be prepared to approach patients who miss appointments or ‘shy away’ from participation, with a clear and persuasive, but non-coercive ‘pitch’ concerning the potential benefits of their participation. This approach will need to be described, both rationally and operational details, to ethics committees for prior approval. The evidence that low adherence is an independent risk factor for mortality [29], and that the treatment on offer has been found to be efficacious when taken, should be informative in such discussions.

Measurement is the key to all research, and our review shows that this is the Achilles’ heel of many intervention trials (and no doubt observational research in this field as well). With the exception of short-term treatments for acute medical conditions (eg, a 1-week course of antibiotic for a urinary tract infection), trialsists should always measure adherence in some way at baseline, and adherence should be measured continuously during the course of the investigation if possible (eg, through prescription refill monitoring at pharmacies) or periodically if not. Findings should be analyzed according to baseline adherence. Clinical outcomes should be measured, as objectively as possible, along with adverse effects.

Interventions should build on previous research and must be described in full and replicatedly executed. As we have little evidence to date of interventions with consistent benefits, a case can be made for “kitchen sink” interventions, designed to maximize the size of effects on adherence. However, practical innovation is needed if interventions are to be useful in usual care settings, so success with complex interventions should be followed by trials of their components. Further, a strong case can be made for testing the tailoring of interventions, from a defined set of available components, for the needs and preferences of individual patients, rather than a “one size fits all” approach. This would be consistent with the emerging principles of patient engagement and shared decision-making [30].

The remaining issues in Table 2 are standard for clinical trials of all health care interventions and need not be reiterated here except to point out that most of these “best in class” adherence clinical trials fell short of meeting them.

Management of adherence in clinical trials that are testing interventions not for the purpose of improving adherence is discussed elsewhere [31].

Improved adherence to efficacious, self-administered medications could have a galvanizing effect on the effectiveness and efficiency of modern health care. Studies that follow the methods that are described in this article are more likely to lead to such advances.

Funding source

Partially supported by Canadian Institutes of Health Research Knowledge Synthesis Grant KRS 262115.

References

1. Tamblyn R, Eguale T, Huang A, Winslade N, Doran P (2014) The incidence and determinants of primary nonadherence with prescribed medication in primary care: a cohort study. J Intern Med 160: 441-450. [Crossref]
2. Naderi SH, Bestwick JP, Wald DS (2012) Adherence to drugs that prevent cardiovascular disease: meta-analysis on 376,162 patients. Am J Med 125: 882-887. [Crossref]
3. Nieuwlaat R, Wilczynski N, Navarro T, Hobson N, Jeffery R, et al. (2014) Interventions for enhancing medication adherence. Cochrane Database Syst Rev 11: CD009011. [Crossref]
4. Ioannidis JP, Greenland S, Hlatky MA, Khoruy MJ, Macleod MR, et al. (2014) Increasing value and reducing waste in research design, conduct, and analysis. Lancet 383: 166-175. [Crossref]
5. Chalmers I, Bracken MB, Djulbegovic B, Garattini S, Grant J, et al. (2014) How to increase value and reduce waste when research priorities are set. Lancet 383: 156-165. [Crossref]
6. Jeffery RA, Navarro T, Wilczynski NL, Isman EC, Keepanasseri A, et al. (2014) Adherence measurement and patient recruitment methods are poor in intervention trials to improve patient adherence. J Clin Epidemiol 67: 1076-1082. [Crossref]
7. Cochrane AL (1972) Effectiveness and efficiency. Random reflections on health services. The Nuffield Provincial Hospitals Trust, Burgess and Son, Abingdon, Berks, UK.
8. Higgins JP, Altman DG, Gotzsche PC, Juni P, Moher D, et al. (2011) The Cochrane
