ELSEVIER

Journal of Clinical Epidemiology 66 (2013) S5−S11

Journal of Clinical Epidemiology

Implementing randomized effectiveness trials in large insurance systems

Niteesh K. Choudhry*, William H. Shrank

Division of Pharmacoepidemiology and Pharmacoeconomics, Department of Medicine, Brigham and Women's Hospital and Harvard Medical School, 1620 Tremont Street, Suite 3030, Boston, MA, USA

Accepted 22 March 2013

Abstract

Background: The need to identify how best to structure health insurance and to deliver health care services is a central priority for comparative effectiveness research. Studies designed to evaluate these issues are frequently conducted in large insurance systems. We sought to describe the challenges faced when conducting trials in this context.

Methods: Using the Post-Myocardial Infarction Free Rx Event and Economic Evaluation (MI FREEE) trial as an example, we describe the methodological and practical challenges of conducting trials in large insurance systems.

Results: We encountered six key challenges while conducting MI FREEE trial, namely the need to obtain plan sponsor permission to experiment, the desire of plan sponsors to have all of their beneficiaries receive the same intervention, the inaccuracy of claims-based identification methods and the impact of claims lag on the timely enrollment of potentially eligible patients, the reluctance of patients to participate in insurance-based interventions and the potential need for informed consent, the frequent introduction of new cointerventions in real-world delivery systems, and the high rates of loss to follow-up because of insurance ''churn.'' We describe the approaches we used to overcome these challenges.

Conclusions: Studies in insurance settings are a powerful and necessary design for evaluating comparative effectiveness interventions. There are numerous strategies to address the potential logistical and methodological challenges that this research environment uniquely creates.

Keywords: Health insurance; Randomized trials; Cost sharing; Acute myocardial infarction; Adherence; Secondary prevention

1. Introduction

The need to identify how best to structure health insurance and to deliver health care services is a central priority for comparative effectiveness research. However, rigorous prospective studies of innovative policies and benefit designs or large quality improvement interventions occur infrequently, leaving a notable lack of data to support evidence-based policy making. New policies are often implemented in the context of health insurance systems, the primary purpose of which is to administer health care benefits rather than to conduct research. These experiments often are not designed in a manner that promotes rigor, limiting the opportunity to optimally learn generalizable lessons about better policy making.

More rigorous prospective randomized designs conducted in large insurance systems are a rarity [1,2]; and when they are attempted, pragmatic and methodological problems are encountered that are distinct from those seen in more traditional research settings. We are unaware of any systemic description of these problems, and, accordingly, the objective of this article is to describe the challenges that we faced when conducting the recently completed Post-Myocardial Infarction Free Rx Event and Economic Evaluation (MI FREEE) trial and to discuss the strategies we used to overcome them [3]. We focus on issues that are of particular relevance to cluster randomized policy studies with prospective participant recruitment conducted in partnership with large commercial insurers, although we have also recently faced many similar challenges while designing individually randomized comparative effectiveness trials of quality improvement interventions in other insurance settings. We hope that the lessons learned can serve to help develop best practices as researchers increasingly collaborate with commercial insurers to implement and test new payment and delivery system approaches.

Funding: No external funding was received for this work.

Conflict of interest disclosure: Drs N.K.C. and W.H.S. have received research support through unrestricted grant to the Brigham and Women's Hospital from Aetna and CVS Caremark to conduct studies in large insurance systems. Dr N.K.C. is a consultant to Mercer Health and Benefits, Inc. Dr W.H.S. has been a consultant on research methodology to United Healthcare.

* Corresponding author. Tel.: +617-278-0930; fax: +617-232-8602.

E-mail address: nchoudhry@partners.org (N.K. Choudhry).

0895-4356/$ - see front matter
http://dx.doi.org/10.1016/j.jclinepi.2013.03.022

What is new?

Key findings

- We encountered six challenges while conducting the Post-Myocardial Infarction Free Rx Event and Economic Evaluation (MI FREEE) trial, namely the need to obtain plan sponsor permission to experiment, the desire of plan sponsors to have all of their beneficiaries receive the same intervention, the potential inaccuracy of claims-based identification methods and the impact of claims lag on the timely enrollment of subjects, the reluctance of patients to participate in insurance-based interventions and the potential need for informed consent, the frequent introduction of new cointerventions in real-world delivery systems, and the high rates of loss-to-follow-up because of insurance ''churn''.

What this adds to what is known?

- The need to identify how best to structure health insurance and to deliver health care services is a central priority for comparative effectiveness research. Studies designed to evaluate these issues are frequently conducted in large insurance systems.
- There are numerous strategies to address the potential logistical and methodological challenges that this research environment uniquely creates.

What is the implication and what should change now?

- Pragmatic studies in insurance settings are essential to identify strategies to improve health care delivery, and best practices to conduct such studies must be developed.

2. Case example: MI FREEE trial

The MI FREEE trial was a randomized policy experiment designed to evaluate the comparative effectiveness of two insurance benefit designs (full vs. usual prescription drug coverage) for secondary prevention medications prescribed to patients after acute myocardial infarction (MI) [4]. The study was motivated by the consistent observation that rates of long-term use or ''adherence'' to evidence-based cardiovascular medications are extremely low [5]. Of the many factors that contribute to this problem, the costs faced by patients when purchasing their drugs (generally in the form of copayments and coinsurance) are a well-recognized contributor [6]. Considering the proven efficacy of secondary prevention medications after MI, there was reason to believe that efforts to improve adherence by eliminating patient out-of-pocket spending could both improve health outcomes and reduce health care expenditures because of savings from averted hospitalizations and procedures [7,8]. The MI FREEE trial aimed to test this hypothesis.

The design of the MI FREEE trial has previously been published [3,4]. In brief, the trial included individuals recently discharged from hospital after acute MI who received health and pharmacy benefits from Aetna, a large health insurer in the United States. Potentially eligible patients, who were identified using administrative discharge claims submitted by hospitals to Aetna, either received full or usual coverage for secondary prevention therapies (i.e., any prescribed angiotensin-converting enzyme inhibitors, angiotensin receptor blocker, beta-blocker, or statin). Assignment occurred by cluster randomization at the level of the plan sponsor (i.e., the employer, union, government, or association that sponsors a particular benefits package), such that after the first eligible beneficiary of a given plan sponsor had been identified and assigned to a treatment arm, all subsequently eligible beneficiaries of that plan sponsor received the same coverage. The overall trial procedures are summarized in Fig. 1.

During the 34-month study period, 6,768 potentially eligible subjects were identified of whom 913 were not randomized because their plan sponsors opted to not participate. The remaining 5,855 patients (87% of those who were potentially eligible) were randomized an average of 49 days post-MI and followed for a median of 394 days (interquartile range: 201−663 days). Providing full coverage improved rates of adherence to each class of study medication by 4−6% points, reduced rates of first major vascular event (11.0 vs. 12.8 per 100 person-years, hazard ratio: 0.86; 95% confidence interval [CI]: 0.74−0.99) and total major vascular events or revascularization (21.5 vs. 23.3, hazard ratio: 0.89, 95% CI: 0.90−0.99) but did not significantly change not the prespecified primary endpoint, first major vascular event, or revascularization (hazard ratio: 0.93, 95% CI: 0.82−1.04) [3]. Providing more generous coverage led to a nonsignificant reduction in total per capita health care spending ($66,008 for the full-coverage group and $71,778 for the usual-coverage group; relative spending: 0.89; 95% CI: 0.50−1.56) and reduced the patient out-of-pocket costs for drugs and other medical services (relative spending: 0.74; 95% CI: 0.68−0.80).

3. Challenge 1: Conducting trials in insurance systems frequently requires plan sponsor permission

Large insurers provide benefits to millions of individuals, and thus, in principle, trials conducted in this environment should be more than adequately powered for even relatively rare conditions or outcomes. However, insurers provide and administer benefits on behalf of numerous plan sponsors, the largest of which are ''self-insured'' and for

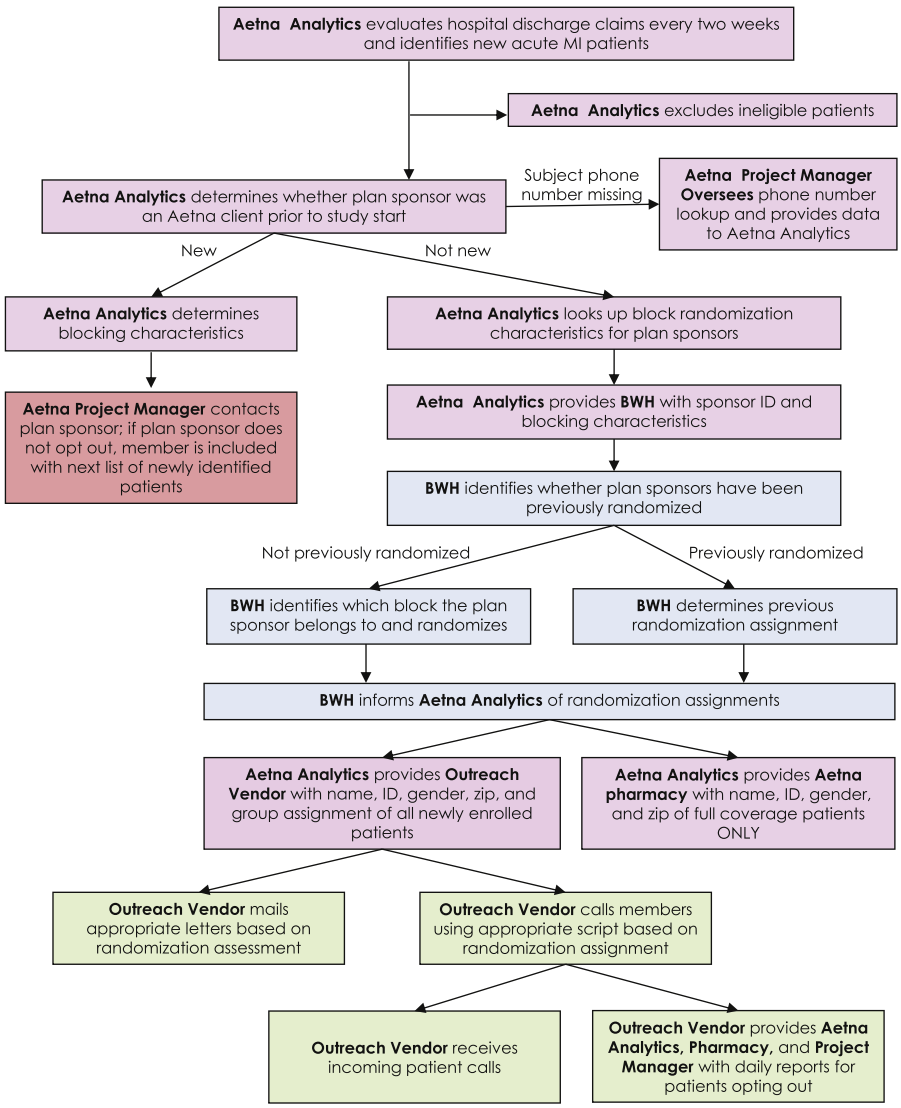

Fig. 1. Patient and plan sponsor recruitment process. MI, myocardial infarction; BWH, Brigham and Women's Hospital.

whom insurers primarily process claims, rather than to insulate the payer against risk (as is done for "fully insured" employers). Because changes in benefit design or other care processes may require additional expense that is borne by the plan sponsor, experimentation requires plan sponsor permission.

Often, even the largest of these plan sponsors will individually have an inadequate number of potentially eligible beneficiaries to allow for a study to reach statistically robust conclusions. For example, in the MI FREEE trial, the median number of enrolled patients per plan sponsor was 1 and 99% of the plan sponsors had 16 or less enrollees, with only one of the almost 3,000 plan sponsors having more than 100 patients enrolled. Furthermore, relying on only a few large plan sponsors for adequate enrollment may influence a study's generalizability as these employers may have different organizational structures, workforce characteristics, industry groupings, geographic locations, and levels of engagement in quality improvement activities than other plan sponsors to whom the results might be applied. Thus, in insurance system–based trials, many different employers must often be approached, which can be resource intensive and, in some cases, infeasible.

To overcome this, the MI FREEE trial used an "opt-out" rather than the more traditional "opt-in" approach. All of Aetna's plan sponsors were contacted by mail before the initiation of the study or as soon as they began providing benefits through Aetna and were given the opportunity to choose not participate in the trial. As a result, only 913 of 3,893 (23.5%) identified plan sponsors chose not to participate, and a high proportion of potentially eligible patients were included in the study and underwent randomization. This is in contrast to typical randomized trials of investigational drug products where it is common for only a small fraction of initially screened individuals to undergo randomization. In the case of the MI FREEE trial, plan

sponsors who agreed to participate and who opted out were very similar with regard to their size, the type of insurance coverage they offered (i.e., full vs. self-insurance), their industry grouping, and the generosity of the health benefits they offered before randomization. Thus the opt-outs did not meaningfully influence the generalizability of the trial results, likely because decisions to decline participation were generally based on the unwillingness of plan sponsors to be involved with research, the unionization of their workforce, or their planned termination of Aetna coverage, rather than systematic characteristics of patients, the generosity of their insurance at baseline, or beliefs about the potential effectiveness of the intervention being evaluated. Of course, not all insurers are willing to use an opt-out approach to recruit plan sponsors and doing so in the MI FREEE trial required very high-level support from Aetna.

4. Challenge 2: Plan sponsors prefer that all of their beneficiaries receive the same intervention

Many employers are willing to participate in low-risk effectiveness studies that seek to maximize the value of health services; however, others are hesitant because of concerns of employee satisfaction or perceptions of equity for beneficiaries with other conditions. Even plan sponsors who are willing to participate in insurance-based trials generally prefer that all of their beneficiaries receive the same intervention, especially for interventions involving changes in insurance benefit design. Employers generally try to avoid the ''water cooler'' problem, where one employee reports receiving an intervention that a coworker did not. This can make patient-level randomization difficult. In such cases, cluster randomization may be used so that all employees of a given employer receive the same intervention. In the case of the MI FREEE trial, patient treatment allocation was determined by random assignment of their plan sponsor to either full or usual insurance coverage. When the first eligible patient of each plan sponsor was identified, plan sponsor randomization occurred and all subsequent patients of that plan sponsor were assigned to the same group.

Although cluster randomization by plan sponsor helps address the ''water cooler'' problem in the short run, it may introduce selection bias for trials with long enrollment periods if beneficiaries who know that they are potentially eligible for an intervention become aware of the group assignment of their plan sponsor and change employment as a result. Furthermore, although plan sponsor clustering reduces contamination in a policy trial, it may not do so for other interventions. For example, in a comparative effectiveness trial of a quality improvement program, it is possible that patients from different plan sponsors who are enrolled in the same trial may interact at their doctor's office or in another health care setting.

As with more traditional clinical trials, the use of cluster randomization introduces a ''design effect'' (i.e., the product of the intercluster correlation and the average size of the cluster), which can reduce statistical power and requires the use of statistical methods that account for the correlations between observations when analyzing study results [9]. In the MI FREEE trial, we assumed a design effect of 1.10 (i.e., 10%) when conducting a priori power calculations based on other published literature [9], although in post hoc analysis, we estimate that the actual design effect was only 1.016 (1.6%), thereby making our pretrial assumptions quite conservative. An alternative strategy, especially when conducting trials in large insurance systems, would have been to use historical claims data to estimate the design effect using methods proposed by Glynn et al. [9] Of course, cluster randomization has the potential to create important imbalances between treatment groups if several of the largest sponsors are randomly allocated to the same arm, especially when there are few participating employers. This concern can be addressed using a block randomized design. In the case of the MI FREEE trial, plan sponsors were categorized into blocks on the basis of whether or not they were nationally based (defined as a Fortune 500 company with more than 3,000 employees or a governmental plan sponsor) and the baseline average copayments required for study medications, as these factors were identified a priori to predict the number of enrolled patients and the outcomes we sought to evaluate. Accordingly, analyses in the MI FREEE trial adjusted for clustering using a robust sandwich estimate for the covariance matrix [10] and blocking factors were used for sample stratification. Stratification factors for other trials will almost certainly be different and may include considerations such as the industry of the plan sponsor (e.g., manufacturing, banking, or health care) as this may reflect knowledge, attitudes, and/or behaviors of the workforce.

5. Challenge 3: Claims-based identification methods and claims lag make the timely and accurate identification of patients difficult

Studies conducted in insurance systems can leverage administrative claims data (i.e., diagnosis, procedure, or encounter or prescription drug information) to identify potentially eligible patients for inclusion. This method of recruitment is appealing because it leverages existing data, at little to no cost, and can be easily scaled to large populations. However, such records are generated based on typical interactions between patients and the health care system for billing and administrative purposes, rather than to support research, and inaccurate claims-based algorithms can reduce the ability to identify patients of interest. This may necessitate substantial additional effort to screen subjects for enrollment and/or undermine the potential benefits of the intervention being evaluated by unintentionally offering it to patients who may benefit less. When choosing between claims-based identification strategies, relying on those with high specificity and/or positive predictive values

will maximize the inclusion of patients of interest but do so at the expense of excluding some potentially eligible subjects. The MI FREEE trial aimed to enroll patients recently discharged from hospital after MI; and for this outcome, there are fortunately algorithms with extremely high specificity and positive predictive value, which also have excellent sensitivity [4,11].

In addition, insurance claims may take months from the time services are rendered to be submitted, processed, and become available for research purposes. Although the resultant ''claims lag'' is well appreciated by investigators working with these data sources, it is particularly problematic in trials that seek to enroll patients prospectively. Furthermore, depending on the nature of the study design and intervention, patients may need to be randomized and contacted by mail and/or phone to obtain consent even after identification, which introduces a further delay (''randomization lag''). In the case of conditions such as acute MI, where recurrent events take place shortly after an initial event, claims and randomization lag could prevent the timely deployment of interventions to the patients who may benefit most from them and may generate a null result for strategies that may have been effective (or harmful) if more rapidly deployed.

The difficulties of claims lag can be partially overcome by querying the data warehouses of insurers on a frequent basis. However, the volume of claims processed by large insurers and the frequency with which analytic databases that can be used for research purposes are refreshed may make this very resource intensive and, therefore, practically challenging. In the MI FREEE trial, we queried Aetna's data warehouse every 2 weeks to identify newly discharged patients and had a very simple enrollment process. As a result, study group assignment occurred at a mean of 49 days after hospital discharge, and 95% of patients were assigned within 100 days after discharge, which is the time point after which nonadherence to post-MI medications begins to become frequent [12].

6. Challenege 4: Patients may be skeptical of insurer-based interventions and may need to consent to participate

Patients may be skeptical of interventions being delivered by health insurers as compared with those administered by direct health care providers, and therefore may be less likely to agree to participate or adhere to the study protocol. This even occurred in the MI FREEE trail, albeit to a very small extent, as several patients randomized to full coverage chose not to participate. This was surprising, especially when considering that patients in the intervention group received their already prescribed medications for free and had to do nothing to receive the benefit (because insurance authorization codes were changed such that after randomization patients paid nothing for the study drugs at the point of care). Nevertheless, it is likely that patient opt-outs may be more frequent in interventions that are more onerous, and this risk may be mitigated with a carefully scripted recruitment process. In the MI FREEE trial, hundreds of standardized responses to potential inquires from patients were generated before the study start.

Similar to more conventional trials, patients may be required to provide informed consent, which may create an additional barrier to participation. Consent may not always be necessary from either a regulatory or ethical perspective for ''low-risk research when soliciting consent is not practicable and consent would not provide meaningful protection for subjects'' [13], conditions that may be common for trials conducted in the context of large insurance systems. In the case of the MI FREEE trial, no specific patient-level written informed consent was sought because all patients, at a minimum, received their usual level of prescription drug coverage, the intervention involved a change in insurance benefits, which plan sponsors can make without the permission of their beneficiaries, and the intervention aimed to promote adherence to guideline-recommended therapies that were prescribed, independent of the study, by patients' treating providers. As a result, in cases where the risk to patients is negligible, the opportunity to forgo informed consent can facilitate the enrollment process.

7. Challenge 5: Cointerventions are frequently introduced in real-world delivery systems

Because health services interventions occur in the context of real-world health benefit and delivery systems, cointerventions are likely to be introduced during the conduct of a trial and are more likely to occur than in the highly structured environments of typical clinical trials. For example, benefit changes, such as changes in formulary status, levels of cost sharing, or the nature of covered services, occur on an almost annual basis. This may be particularly true for patients who are able to choose from more than one insurance product (offered by the same insurer). Other quality improvement strategies, such as disease management or wellness initiatives, could be introduced on the basis of newly generated research evidence or the interest of insurers and plan sponsors in achieving performance measurement bench marks, such as Healthcare Effectiveness Data and Information Set scores [14]. Payers may be unwilling to withhold these programs from the trial participants.

Unanticipated cointerventions of these sorts could certainly confound the exposure−outcome relationship of interest. Their potential impact can be addressed with a randomized study design with or without post hoc adjustment. In the case of randomization, cointerventions would generally bias results toward the null if applied equally to all treatment groups, especially when using intention-to-treat principles. Of course, producing conservative results may still lead to incorrect inferences about the potential benefits

or harms of an intervention. For anticipated or known cointerventions, preassignment stratification could be implemented, although this may add complexity to the study design. In the MI FREEE trial, the baseline generosity of insurance benefits was one of the stratification factors used to minimize confounding by the change in cost sharing experienced by patients. Furthermore, we prospectively recorded cointerventions for patients enrolled in the trial and conducted sensitivity analyses adjusting for these variables.

8. Challenge 6: Insurance "churn" makes loss to follow-up a frequent occurrence

Achieving sufficient sample size may also be challenged by insurance ''churn,'' or the fact that large numbers of beneficiaries will drop coverage every year because of a change or loss of employment or because the employer has changed the insurer that provides benefits to its employees. In the United States, an estimated 15−20% of patients enrolled in typical (often called ''commercial'') insurance plans change insurers (or drop insurance altogether) every year, with even higher rates among Medicaid enrollees [15]. This threatens both eligible pool of subjects (when there is a lag between identification and recruitment) and follow-up time, and may introduce bias if these beneficiaries are systematically different than those who remain in plans.

In the case of the MI FREEE trial, approximately 5% of patients who were identified as potentially eligible lost insurance eligibility in the period between hospital discharge and randomization. Because we conducted our analysis using intention-to-treat principles, these patients were included in the denominator of our outcome rates although they contributed no follow-up data. To address this information loss, we accounted for a noninformative censoring rate of 15% in our a priori sample power calculations [4]. Fortunately, our trial focused on patients with MI for whom events occur soon after hospital discharge (and therefore for whom long-term follow-up may not meaningfully threaten power). For studies interested in long-term outcomes, investigators should consider insurance environments that provide more stable coverage, like Medicare or the Veterans Affairs system or other countries with less fragmented health care systems than in the United States. Alternatively, other data collection methods, such as direct patient and/or provider outreach or even the use of national registries, such as the National Death Index [16], could be used to supplement standard claims-based approaches and facilitate long-term follow-up, albeit for a more limited set of outcomes.

9. Conclusions

Despite the potential barriers described previously, randomized studies conducted in real-world insurance settings are a powerful and necessary design for evaluating comparative effectiveness interventions. As we attempt to transform the health care system to deliver higher quality care at lower cost, there is great urgency to identifying and testing payment strategies to achieve these goals. Pragmatic trials are essential to test new benefit designs and payment methodologies in real-world settings and can provide actionable evidence for policymakers that can be rapidly expanded. In fact, Aetna announced that they would scale a variant of the MI FREEE trial benefit to their fully insured beneficiaries after the results of the trial were published. Such prospective randomized controlled trial evidence is rarely available to policymakers when determining the structure of coverage and benefit designs. More such studies, prospectively and rigorously evaluating the effect of new payment strategies, are essential if we hope to implement evidence-based policies.

More generally, the strategies and methods used in the MI FREEE trial to overcome challenges of trials conducted in the context of large insurance systems could be applied to a broad range of other comparative effectiveness questions beyond benefit design, and the experience implementing this trial offers meaningful insight and guidance to others. As experience with such trials increase, standard methods will develop to facilitate their ongoing conduct.

References

[1] Finkelstein A, Taubman S, Wright B, Bernstein M, Gruber J, Newhouse JP, et al. The Oregon health insurance experiment: evidence from the first year. Q J Econ 2012;127:1057−106.
[2] Newhouse JP, Manning WG, Morris CN, Orr LL, Duan N, Keeler EB, et al. Some interim results from a controlled trial of cost sharing in health insurance. N Engl J Med 1981;305:1501−7.
[3] Choudhry NK, Avorn J, Glynn RJ, Antman EM, Schneeweiss S, Toscano M, et al. Full coverage for preventive medications after myocardial infarction. N Engl J Med 2011;365:2088−97.
[4] Choudhry NK, Brennan T, Toscano M, Spettell C, Glynn RJ, Rubino M, et al. Rationale and design of the Post-MI FREEE trial: a randomized evaluation of first-dollar drug coverage for post-myocardial infarction secondary preventive therapies. Am Heart J 2008;156:31−6.
[5] Choudhry NK, Setoguchi S, Levin R, Winkelmayer WC, Shrank WH. Trends in adherence to secondary prevention medications in elderly post-myocardial infarction patients. Pharmacoepidemiol Drug Saf 2008;17:1189−96.
[6] Goldman DP, Joyce GF, Zheng Y. Prescription drug cost sharing: associations with medication and medical utilization and spending and health. JAMA 2007;298:61−9.
[7] Choudhry NK, Avorn J, Antman EM, Schneeweiss S, Shrank WH. Should patients receive secondary prevention medications for free after a myocardial infarction? an economic analysis. Health Aff (Millwood) 2007;26:186−94.
[8] Choudhry NK, Patrick AR, Antman EM, Avorn J, Shrank WH. Cost-effectiveness of providing full drug coverage to increase medication adherence in post-myocardial infarction Medicare beneficiaries. Circulation 2008;117:1261−8.
[9] Glynn RJ, Brookhart MA, Stedman M, Avorn J, Solomon DH. Design of cluster-randomized trials of quality improvement interventions aimed at medical care providers. Med Care 2007;45:S38−43.

[10] Lin DY, Wei LJ. The robust inference for the proportional hazards model. J Am Stat Assoc 1989;84:1074−8.
[11] Petersen LA, Wright S, Normand SL, Daley J. Positive predictive value of the diagnosis of acute myocardial infarction in an administrative database. J Gen Intern Med 1999;14:555−8.
[12] Benner JS, Glynn RJ, Mogun H, Neumann PJ, Weinstein MC, Avorn J. Long-term persistence in use of statin therapy in elderly patients. JAMA 2002;288:455−61.
[13] Miller FG, Emanuel EJ. Quality-improvement research and informed consent. N Engl J Med 2008;358:765−7.
[14] National Committee for Quality Assurance. HEDIS & quality measurement. Available at http://www.ncqa.org/tabid/59/default.aspx. Accessed August 15, 2012.
[15] Sommers BD, Rosenbaum S. Issues in health reform: how changes in eligibility may move millions back and forth between medicaid and insurance exchanges. Health Aff (Millwood) 2011; 30:228−36.
[16] Center for Disease Control and Prevention. National Death Index. Available at http://www.cdc.gov/nchs/data_access/ndi/about_ndi.htm. Accessed August 15, 2012.