Designing and implementing research on a statewide quality improvement initiative: The DIAMOND Study and Initiative

A. Lauren Crain, PhD,
HealthPartners Research Foundation PO Box 1524, MS 21111R Minneapolis, MN 55440-1524

Leif I. Solberg, MD,
952-967-5017 phone 952-967-5022 fax leif.i.solberg@healthpartners.com

Jürgen Unützer, MD, MPH,
University of Washington Medical Center 1959 NE Pacific St, BB-1661A Seattle, WA 98195 (206) 543-3128 phone (206) 221-5414 fax unutzer@u.washington.edu

Kris A. Ohnsorg, RN, MPH,
952-967-5011 phone 952-967-5022 fax kris.a.ohnsorg@healthpartners.com

Michael V. Maciosek, PhD,
952-967-5073 phone 952-967-5022 fax michael.v.maciosek@healthpartners.com

Robin R. Whitebird, PhD, MSW,
952-967-5005 phone 952-967-5022 fax robin.r.whitebird@healthpartners.com

Arne Beck, PhD, and
Clinical Research Unit Kaiser Permanente Colorado PO Box 378066 Denver, CO 80237 303-614-1326 phone 303-614-1285 fax arne.beck@kp.org

Beth A. Molitor, MBA
952-967-5174 phone 952-967-5022 fax beth.a.molitor@healthpartners.com

Abstract

Objective—To demonstrate a rigorous methodology that optimally balanced internal validity with generalizability to evaluate a state-wide collaborative that implemented an evidenced-based, collaborative care model for depression management in primary care.

Study Design and Setting—Several operational features of the DIAMOND Initiative suggested that the DIAMOND Study employ a staggered implementation design with repeated cross-sections of patients across clinical settings. A multi-level recruitment strategy elicited virtually complete study participation from the medical groups, clinics and health plans that coordinated efforts to deliver and reimburse DIAMOND care. Patient identification capitalized on large health plan claims databases to rapidly identify the population of patients newly treated for depression in DIAMOND clinics.

Results—The staggered implementation design and multi-level recruitment strategy made it possible to evaluate DIAMOND by holding confounding factors constant and accurately

A. Lauren Crain, PhD – corresponding author 952-967-5354 phone 952-967-5022 fax lauren.a.crain@healthpartners.com.

identifying an intent-to-treat population of patients treated for depression without intruding on or requiring effort from their clinics.

Conclusion—Recruitment and data collection from health plans, medical groups and clinics, and patients ensured a representative, intent-to-treat sample of study-enrolled patients. Separating patient identification from care delivery reduced threats of selection bias and enabled comparisons between the treated population and study sample. A key challenge is that intent-to-treat patients may not be exposed to DIAMOND which dilutes the effect size but offers realistic expectations of the impact of quality improvement in a population of treated patients.

Keywords

stepped wedge design; depression; models, organizational; patient selection; quality improvement; research design; delivery of health care

There has been an increasing focus on research to inform the implementation of evidence-based treatment programs in real world practice settings, but there are few studies with designs or recruitment strategies that support such work. We present as a case study an evaluation of a state-wide quality improvement initiative for depression that uses a novel quasi-experimental design and multi-level claims data-based recruitment strategies. The complexity of implementing the quality improvement initiative required a study design that was non-disruptive for participants yet sufficiently rigorous to allow accurate estimation of effectiveness. Similarly, an efficient and unbiased recruitment process was needed that did not over-burden care delivery personnel. This paper describes the logic of the design, multilevel recruitment strategy**,** data collection **and data analytic** plans of the DIAMOND Study, and concludes with a discussion of how its design can inform future studies of natural experiments.

The DIAMOND Initiative

The DIAMOND (Depression Improvement Across Minnesota, Offering a New Direction) Initiative is a state-wide collaborative effort among health care payers and clinicians that is coordinated by an independent quality improvement organization (The Institute for Clinical Systems Improvement, ICSI) with the goal of implementing an evidenced-based, collaborative care model for depression management in primary care. This Initiative has transformed depression care in Minnesota, and is a model for collaborative transformational change for other conditions and regions.

The DIAMOND Study was conducted by an external team of researchers in partnership with ICSI. It provided a rare opportunity to evaluate **the DIAMOND Initiative,** a large scale implementation of a fundamental change in the way that care is delivered in primary care settings. The DIAMOND care model departed from usual primary care by incorporating all key components of the evidence-based collaborative care management program including assessing and monitoring depression, care management follow-up and psychiatrist availability to provide treatment recommendations.(1-3) It required close collaboration between the multi-site, multi-disciplinary team of researchers and ICSI, who coordinated the interests of all Initiative and Study stakeholders, to quantify the effectiveness and impacts of this transformation.(1)

Numerous depression trials have identified that collaborative care, particularly the involvement of a care manager and planned psychiatry review, is effective in improving the short-term and possibly long-term outcomes for patients with depression.(2-3) The predominant model of care for depression in primary care does not include these key elements of collaborative care, in part because the reimbursement structure has not

supported it. The DIAMOND Initiative sought to rectify this through a statewide collaboration among numerous stakeholders including health plans, medical groups, and ICSI. The key operational features **of the Initiative** that impacted the DIAMOND Study design were:

1. ICSI identified the specific care model features from the literature, trained clinic teams in those features, and certified their presence prior to clinic eligibility for reimbursement.

2. Clinic training and certification was spread across five sequential periods over two and a half years to aid implementation feasibility.

3. Multiple clinics within a single medical group could join the Initiative in different training and certification sequences, or not join the Initiative at all.

4. Because ICSI membership includes most medical groups in Minnesota and all were offered the opportunity to participate in DIAMOND, it was not feasible to identify a non-participating group of comparison clinics.

5. It was not possible to randomize patients within clinics into new versus traditional care since entire clinics **and its practicing clinicians** participate**d** in the new process while allowing individual clinician and patient choice in using this care approach.

6. As is true for nearly all primary care clinicians in the U.S., patients within a clinic have a variety of insurance coverage sources and products, with no single source for more than 20-30% of their patients.

More complete and detailed descriptions of DIAMOND Initiative components has been provided elsewhere. (1, 4-5)

Payment for DIAMOND services was a flat monthly per-participant amount which was negotiated independently between participating health plans and medical groups, but was billed using a common special billing code. Because only patients with a pre-specified set of diagnostic codes were eligible for DIAMOND payment, it was possible to clearly define a denominator of DIAMOND-eligible and DIAMOND-treated patients from claims data. However, with the important exception of prescription drug claims, the time lag in processing claims prevented their use in identifying DIAMOND-eligible or -treated patients until they were well-into their care. Nor would claims for DIAMOND services enable a denominator of comparison patients to be identified.

METHODS

With these constraints in mind, we sought to develop recruitment and analysis plans that could assess the effects of the Initiative on a variety of patient outcomes and healthcare costs while also learning about the implementation process and its costs. It was necessary that recruitment and data collection be designed and implemented between the time DIAMOND Study research funding was approved by NIH and when ICSI, the health plans, and medical groups were prepared to roll-out the DIAMOND Initiative. This timing problem for research funding before the beginning of a rigorous prospective evaluation was an additional challenge that may be under-appreciated when attempting to study complex natural experiments such as this one.

Study design

The nature of DIAMOND care and the scale of its implementation made patient randomization infeasible and sub-optimal. A compelling argument for staggered

implementation came from the practical constraints imposed by the DIAMOND certification process. The health plans and medical groups agreed that DIAMOND care would only be offered to patients treated in ICSI-trained and certified clinics. The structure of the negotiated certification process and the volume of clinics seeking to offer DIAMOND care made it impossible for ICSI to simultaneously train and certify all interested clinics. To manage this complexity, one sequence of clinics began training and certification with ICSI every six months over a two and a half year period. At the conclusion of the process the sequence would implement DIAMOND as the next sequence began training and certification (Figure 1). Each sequence of clinics continued to deliver usual care for depression to patients prior to implementing DIAMOND until all five sequences of clinics were trained and certified. The clinics in each sequence were determined by ISCI so that each sequence would be similarly sized with respect to the number of clinician FTEs represented (by health plan requirement) and clinic size; and clinics from large medical groups with involvement of many sites were distributed across multiple sequences.

It was not possible to randomize patients to receive DIAMOND care prior to their clinic being certified. Nor was it preferable to individually randomize patients treated for depression in clinics after implementing DIAMOND. Developing a patient randomization protocol would have created logistical challenges that could not be addressed in the time available between finalizing the Initiative, obtaining research funding and Initiative start-up. Even if these challenges were overcome, randomization would likely be subverted by the incidental benefits to non-DIAMOND patients that would result from structural or practice changes that had taken place in the clinic. Patient and physician preferences, pre-existing complications, and other factors influence clinical decision making,(6) and asking clinicians to engage in a randomization process would have interrupted and distorted clinical practice. These distortions to clinical practice may have provided the study with outcome data that estimated the efficacy of DIAMOND but the evaluative objective of the Study was to estimate its effectiveness so that Initiative stakeholders would have externally valid estimates of the benefits and costs of implementing DIAMOND.

All of these considerations pointed to the value of a staggered implementation,(7-8) or stepped wedge,(9-10) design with repeated cross-sections of patients across clinics settings This approach is valued for its reduction in threats to internal validity for quality improvement research,(8) as an excellent way to encourage "practical clinical trials,"(7, 10-12) in which hypotheses and study designs are tailored to meet the needs of decision makers, and as being "consistent with a decision making process used by a wide range of policymakers, educators, and health officials in their periodic examination of administrative or surveillance data."(13)

Staggered implementation strengthened the study design by reducing internal validity threats that would have resulted from patient randomization. The two year implementation period and five implementation sequences made it less plausible to attribute pre-to post-implementation changes in depression care to secular trends since there was temporal variation in the clearly defined pre- and post-implementation periods across clinics. It also meant that a patient identification protocol had to be developed that could be implemented identically during the pre- and post-implementation periods in each clinic.

Multi-level recruitment

The DIAMOND Study requested active participation from ICSI, health plans, medical groups and clinics, and patients treated for depression, beginning as **sequence 1** was about to start training. Recruitment at all of these levels was undertaken to ensure enrollment of **a sample of patients** that was representative of all patients initiating treatment for depression in DIAMOND-participating clinics.

ICSI agreed to identify and schedule training for all potential clinic participants prior to **training sequence 1**, provided Study personnel with contact information for all **participating** clinics, and regularly updated information regarding participating clinics and their sequence assignments. It also provided data about the cost of implementing its part of the Initiative. Health plans provided the Study with cost estimates for DIAMOND implementation; enacted identification, notification and opt-out procedures for potential study-eligible health plan members; and provided contact information about these members as well as claims data for those who consented. Medical group and clinic study roles entailed completing cost estimates for DIAMOND implementation and operation; completing leadership surveys regarding readiness for organizational change and implementation of depression systems prior to implementing DIAMOND and again 1 and 2 years post-implementation; periodically updating clinician lists; allowing the Study team to contact patients being treated for depression in participating clinics; and providing emergency care for patients identified by the Study with suicide risks. Patients completed screening, baseline and **6 month** follow-up telephone surveys, and most consented for their health plan to provide the DIAMOND Study with their health care utilization data. All protocols were reviewed and approved by the HealthPartners Institutional Review Board.

Health Plan Databases for Patient Identification

The aims of the DIAMOND Study were to evaluate the effectiveness of DIAMOND implementation for patients treated for depression in participating primary care clinics. The Study protocol required:

1. Identifying and attempting to recruit all patients initiating treatment for depression in a DIAMOND clinic,

2. Enrolling patients who were representative of the treated population demographically, clinically and with respect to health plan, clinic, and clinician,

3. Use of a common identification method for patients treated in pre- and post-implementation settings, and

4. Identifying a sufficiently large number of patients to achieve the recruitment goal of N=2715 study-enrolled patients in 36 months.

One common approach to identifying patients in studies of improved depression care has been to place study personnel in a clinical setting or to enlist the assistance of clinicians so that potentially study-eligible patients may be identified at the time of clinic visits. The DIAMOND Study team instead worked with the health plans to develop a patient identification protocol that would capitalize on the efficiency of large claims databases to rapidly identify patients who had a high probability of being newly treated for depression.

The primary means of identifying patients relied on pharmacy claims information that was gathered by health plan personnel every 7 or 14 days for the duration of the 36 month enrollment period. Each participating health plan used submitted pharmacy claims to identify members who in the past 7 or 14 days had filled a prescription for an anti-depressant medication that had been ordered by a clinician practicing in all pre- or post-implementation DIAMOND clinics. Removed from this list were members who had additional anti-depressant fills in the previous 4 months and were therefore unlikely to be initiating depression treatment, who did not have continuous medical and pharmacy coverage in the previous 4 months so that other recent fills might not be observed, **who had previously been identified** or who had opted out of all research participation. The health plans notified by mail each member who remained on the list that unless the member contacted the health plan within 7 or 10 days, the health plan would forward to the Study

team a minimally necessary set of fields pertaining to the anti-depressant fill and contact information.

Following this 7-10 day opt-out period, health plan personnel sent the information to the Study team via a secure file transfer protocol. The study team recorded the health plan from which each patient in the sampling frame was received, and used the prescribing clinician on the pharmacy claim to assign each patient to the clinic and medical group where he or she was being treated for depression. Clinic assignment was used to classify each patient into one of the five implementation sequences; in turn, implementation sequence, anti-depressant fill date and sequence implementation dates will be used in the primary analysis to categorize each patient as a pre- or post-implementation patient.

Patient Eligibility and Enrollment

The study team attempted to contact by phone each patient in the sampling frame within 21 days of receiving contact information to determine study eligibility. Patients were study-eligible if they were at least 18 years of age on the date of the anti-depressant fill, confirmed that the medication was prescribed for the treatment of depression by a clinician in a DIAMOND-participating clinic, and at the time of the telephone screen had a score of at least 7 on the Patient Health Questionnaire (PHQ-9), a self-reported measure of depression severity. The DIAMOND Study used a lower PHQ-9 criterion than the Initiative (PHQ-9 10) to identify patients eligible for DIAMOND services in order to include patients whose depression severity may have improved during the time delay. Eligible patients were asked for their consent to complete a baseline telephone survey and a follow-up survey 6 months later, as well as their consent for their health plan to provide the DIAMOND Study with their health care utilization data for the 12 month period following the qualifying anti-depressant fill. A subset of patients was asked to complete two additional surveys at 3 and 12 months to better assess patterns of productivity and quality of life over time.

In each of the 36 months of the enrollment period, health plan members who had received care in all DIAMOND participating clinics, both pre- and post-implementation, were identified via pharmacy claims and screened for eligibility on a continuous basis. The screening protocol ensured that the patient acknowledged being treated for depression at a DIAMOND clinic, and that the depression severity was sufficient to make the patient eligible for DIAMOND care. Each enrolled patient contributed a baseline survey at enrollment and was asked to complete a 6 month follow-up survey.

ENROLLMENT RATES

Clinics and Medical Groups

By the time **training and certification sequence 5** ended in March 2010, n=7 payors (6 private health plans and the Minnesota Department of Human Services), n=22 medical groups and n=84 clinics participated in the DIAMOND Initiative, while n=7 payors, n=25 medical groups and n=96 clinics participated in the Study (some of which did not implement DIAMOND). Strong study participation from the population of DIAMOND health plans, medical groups and clinics has helped guarantee an accurate understanding of DIAMOND implementation at all levels, but more importantly ensured that patients targeted for recruitment were representative of the population of patients for whom DIAMOND was developed.

Table 1 summarizes characteristics of the medical groups and clinics that participated in the DIAMOND Study. The participating medical groups ranged in size from one care site with relatively few clinicians to nearly fifty sites with hundreds of clinicians including dozens of mental health clinicians. Care manager credentials varied across the medical groups, as did

the payer mix of clinic patient populations. Some groups provided care mostly to patients with commercial insurance, while others had a larger share of patients whose medical coverage was provided by Medicare, Medicaid, or another non-commercial entity. The participating clinics were nearly evenly split between those located in the Twin Cities metropolitan area, and those in the smaller cities of Rochester and Duluth or in rural Minnesota. The number of participating medical groups and clinics, and the breadth in their characteristics, will enable an assessment of whether these characteristics are related to DIAMOND implementation (e.g., patient enrollment rates) and effectiveness (e.g., proportion of patients in remission within 6 months).

Patients

During the patient recruitment period (February, 2008 to January, 2011), the health plans identified and sent the study team contact information for N=24,065 patients who were being treated for depression in a clinic that implemented DIAMOND, as well as the clinics that had expressed an intention to ICSI to implement DIAMOND but did not do so. The study team was able to make contact with 60% (N=14,453) of these patients within 21 days of receiving the contact information (Table 2). Lack of contact was due in almost equal parts to not having received correct information or repeated attempts without reaching the intended patient. Of those who were contacted, 66% (N=9,498) agreed to be screened for study eligibility, and about 1 in 4 of those screened (26%, N=2,435) met the eligibility criteria. The most common reason for ineligibility was that the patient did not report severe enough depression symptoms (i.e., PHQ-9 < 7, N=3,561) or reported that the antidepressant medication they received was not for the treatment of depression (N=2,860). The remainder reported being treated in a clinic that we determined was not scheduled to implement DIAMOND (N=440) or that they had not filled a prescription for antidepressant medication (N=202). Virtually all of the study-eligible patients consented to participate (N=2,428), completed the baseline survey (N=2,423), and consented to have their claims data used for cost analysis (N=2,315).

Comparisons of patient characteristics through the stages of enrollment demonstrate that the study participants departed in only small ways from the sampling frame, for example, in having a higher proportion of women (3.2%) and being 1.1 years younger (see Table 3). Relative to those screened for eligibility fewer participants held a college degree (5.4%), were living with a significant other (6.9%), or were enrolled in a commercial insurance plan (5.1%).

Enrolled patients were similarly distributed across the five clinic training sequences to the sampling frame, the only small difference being that a smaller proportion of **patients treated in sequence 2 clinics** enrolled than had been in the sampling frame (4.3%). Some of the clinics that had expressed interest in implementing DIAMOND did not complete training. Because all clinics had been invited to participate in the DIAMOND Study, health plans provided the study team with contact information for patients treated in these 'non-implementation' clinics. Non-implementation patients were just as likely to enroll as were the pre- or post-implementation patients, and as a group, were similar in size to the training sequences. Although the study was not designed with this intention, the non-implementation patients will be available as a secular trend comparison group for ancillary analyses.

ANALYSIS PLAN

The primary Study analyses will assess the effectiveness of DIAMOND implementation by comparing clinical processes and outcomes between patients treated in clinics following DIAMOND implementation and patients treated in those same settings prior to implementation. The analytic models will compare patients' 6 month reports of received

depression care processes and secondary outcomes including depression symptoms, work productivity and quality of life. It is possible that measured and unmeasured factors that vary by clinic (e.g., readiness for change, depth of depression system implementation, size, location, patient or staffing mix) are related to patient outcomes at 6 months and introduce dependence among outcomes measured from patients treated in the same clinic. The primary analyses will therefore accommodate intraclass correlation in patient outcomes.

The implementation status of each patient's clinic at the time the patient enrolled in the study will be the key predictor in these analyses. DIAMOND was implemented in a staggered fashion across clinics over a 2½ year period. Patients from all clinics were enrolled into the study throughout this period so that some enrolled prior to their clinic implementing DIAMOND and some following implementation. Implementation status therefore varies across patients who were classified as 'pre-implementation' if their clinic had not yet adopted DIAMOND when they enrolled in the study or 'post-implementation' if they enrolled after their clinic implemented DIAMOND.

It was not possible for a pre-implementation patient to receive DIAMOND care. Following implementation, the intention was that all clinicians would treat patients who met clinical criteria with the DIAMOND care model, although clinicians had the latitude to offer it to individual patients as they deemed appropriate. Consistent with the objective to estimate effectiveness, the primary analysis will take an intent-to-treat approach that compares outcomes of post-implementation to pre-implementation patients, whether or not post-implementation patients received DIAMOND care. A secondary, as-treated approach will compare outcomes among patients to whom DIAMOND care was not yet available at study enrollment, those for whom DIAMOND care became available prior to the 6 month follow-up, those for whom it was available but were not treated with DIAMOND care, and DIAMOND care-treated patients.

General linear mixed model regression will be used to predict individual patient outcomes from their clinic implementation status. Patients will be clustered within clinics, and a random clinic intercept estimated, to minimize the risk of a Type I error resulting from dependence among patient outcomes within clinics. A fixed effect that denotes each patient's clinic sequence will estimate secular trends in patient outcomes. The key predictor will be the fixed effect that denotes the implementation status of each patient's clinic at the time the patient enrolled in the study. The interaction between sequence (or time elapsed since sequence 1 implementation) and implementation status can assess variability in DIAMOND effectiveness over time. Similarly, quantifying each patient's study enrollment relative to DIAMOND implementation at their clinic (e.g., months before or after implementation date) can identify clinic-specific trends in depression care relative to DIAMOND implementation. Of conceptual interest will be including clinic level factors and characteristics so that their contributions to patient outcomes, or their moderating effects on DIAMOND effectiveness, may be estimated. A more detailed treatment of analytic options is presented in Hussey and Hughes (2007).(12)

CONCLUSIONS

Many of the elements of DIAMOND care had demonstrated efficacy in clinical trials, but whether redesigning primary care to deliver these elements improved care among patients initiating depression treatment in real life has not been assessed. The DIAMOND Study is an effectiveness trial that will assess the effects of DIAMOND implementation on patient outcomes. Design decisions for this quasi-experimental trial were made with the goal of minimizing differences between pe- and post-implementation patient groups. The staggered implementation design and multi-level recruitment strategy made it possible to hold constant

factors other than DIAMOND implementation that may have impacted patient outcomes and accurately identify those patients who were treated for depression in a clinical setting that would implement DIAMOND.

Implementing this design required the coordinated action of medical groups, clinics and clinicians to deliver an intervention to patients. The DIAMOND Study team then recruited and collected data from three distinct yet inter-related sources, each necessary for ensuring that the Study recruited a representative, intent-to-treat patient population. Each of the DIAMOND medical groups and clinics described readiness for organizational change and implementation of depression systems throughout the implementation process, and cooperated with our efforts to recruit a representative sample of its patients being treated for depression. These details will provide insight into how participating clinics operationalized DIAMOND so that the impact of the variation in implementation can be quantified for each clinic population. The health plans that provided reimbursement for DIAMOND care provided the Study team with the contact information for a representative sampling frame of patients being newly treated for depression. The result is that the Study team obtained virtually complete data from the higher level populations of DIAMOND-participating health plans and clinics, and from samples of all patients initiating care for depression within each of these units.

The patients of primary interest were those who had initiated care for depression in a primary care setting that had implemented DIAMOND. The most appropriate comparator was patients receiving standard depression care in non-transformed primary care settings. Our collaboration with the medical groups and clinics that implemented DIAMOND, and with the health plans that administered DIAMOND claims, made it possible for us to compare patients being treated in a DIAMOND clinic to those treated for depression in the same clinics but in the absence of DIAMOND implementation.

Efficacy trials often impose stringent inclusion criteria on potential research subjects to rule out plausible alternate explanations for efficacy other than the agent under study. The resulting study-enrolled subjects may be relatively homogeneous with respect to key features, but are likely to differ meaningfully from the clinical population to which a treatment would be applied in real life. The relative homogeneity and size of efficacy trial samples also may preclude well-powered tests for differential efficacy across sub-populations of subjects. For both of these reasons, a proven treatment may not have known efficacy with respect to a more representative treated population or specific sub-populations of interest. The eligibility criteria for the DIAMOND Study were based on clinician diagnoses rather than screening, and on the clinical criteria used by clinicians to determine eligibility for enrollment in the DIAMOND Initiative. This decision increased the likelihood that conclusions drawn from the primary analyses would generalize to a treated population and that there would be ample representation of patient sub-populations to assess differential effectiveness.

Our reliance on large administrative databases to identify patients initiating treatment for depression rather than recruitment in the clinics possessed several clear benefits. There was a clear separation between delivery of DIAMOND care and patient identification and recruitment. Aside from feasibility considerations, such as the relative expense of quickly identifying large numbers of patients, disentangling intervention delivery from its evaluation reduced threats of selection bias by treatment group, care delivery site, or health plan, and provided a sampling frame for a large and well-defined population of patients with quantifiable attributes.

Identical procedures could be used to identify patients being treated for depression in pre- and post-implementation clinics, being treated across all participating health plans and clinics, and regardless of their exposure to DIAMOND care. Except for those administered by a pharmacy-benefit management company, all pharmacy claims are submitted promptly to health plans for reimbursement, creating a well-defined population of all patients with anti-depressant medication within 1-2 days of their prescription fill. From this population, patients can be identified according to study eligibility criteria to obtain a population-representative sampling frame with known characteristics. Furthermore, study enrollees and non-enrollees may be compared on these characteristics to identify key dimensions on which they may differ.

Relative to diagnosis or procedure codes, pharmacy claims rely on coding schemes that uniquely identify the medication that was dispensed, essentially eliminating variation in coding practices, and are processed within days of being submitted. Each health plan was therefore able to identify the population of their members who received care for depression at any of their contracted DIAMOND clinics, and who had filled a prescription for anti-depressant medication, with a high degree of certainty that the patient had recently initiated that medication. The quick processing of pharmacy claims helped the study team to identify and contact study subjects early in their treatment. Pharmacy claims were the most timely, comprehensive means of identifying the population of patients treated for depression and at risk of exposure to the DIAMOND care model, and therefore an intent-to-treat population of patients.

The active involvement of health plans in patient recruitment had the added benefit of asking that they provide the study team with claims data from consenting patients. This will enable the study to estimate healthcare costs for pre- and post-implementation patients and share this secondary outcome information with all participating health plans.

In spite of the clear advantages to using pharmacy claims for identifying patients, there were also some problems with this approach. The negotiation required to secure study participation from all health plans and clinics participating in the Initiative, and then to assist with patient identification, required a cooperative environment that may not be feasible in other settings. Nor was the study team able to control the quality of contact information provided by the health plans.

One obvious drawback to identifying patients via pharmacy claims is that it misses patients whose clinicians are managing the condition without medication, those who did not fill their prescription, those without pharmacy coverage, and those with pharmacy claims not managed by the health plan. It also includes patients being treated with antidepressants for a non-depression condition. In light of the high proportion of patients with depression who are treated with medication, the relatively low likelihood that those treated without medication had severe enough symptoms to be eligible for DIAMOND care, and the high proportion of members that have pharmacy claims accessible to their health plan, we decided that the benefits of using this approach (i.e., sensitive, accurate, timely, no selection) outweighed these limitations.

Another drawback is the time that may elapse between the initiation of treatment and contact with the study participant. Even though health plans provided claims data for antidepressant medication fills relatively quickly, as many as 6 weeks could elapse between the fill and completion of the baseline survey due to the need to mail opt-out postcards and allow time for response prior to initiating recruitment which added to the typical delays associated with using multiple contact attempts to reach patients. This is a potential problem because improvement in depression symptoms is typically the greatest in the acute treatment phase,

with many of the gains made in treatment by 6-10 weeks after treatment initiation. This may reduce the ability to quantify the full improvement in depression symptoms experienced by study participants who are initiated on treatment, although this is a limitation that affects pre- and post-implementation study participants equally.

The most serious concern with this approach, however, is that patients treated with medication in post-intervention clinics were not necessarily exposed to DIAMOND care (e.g., patients without insurance coverage for DIAMOND care) so that the post-implementation patient sample will be a mixture of DIAMOND care and standard care patients. Clinic-based surveillance data from the DIAMOND Initiative suggested that only 31% of patients who were clinically eligible for DIAMOND care may have been activated into DIAMOND care management, thus putting the Study at risk of enrolling too few per-protocol DIAMOND patients. When this threat was recognized, a supplemental patient identification protocol was developed whereby health plans also began to identify patients for study recruitment who had been assigned the DIAMOND procedure code, signifying enrollment into DIAMOND care. Regardless of identification method, the patients in post-implementation clinics who were exposed to the DIAMOND care model can be determined by the presence or absence of DIAMOND-specific procedure codes in their utilization data so that exposure to DIAMOND may be quantified.

Even if DIAMOND care is highly efficacious, the population-level effect size will be diluted by the presumably smaller effect size observed among intent-to-treat patients not exposed to DIAMOND. In the absence of substantial efficacy, then, an intent-to-treat analysis is potentially underpowered to detect a significant difference between post- and pre-implementation patients. A highly efficacious intervention that suffers from poor reach may produce inconclusive results regarding the effects of the intervention, yet promote more tempered and realistic expectations of the impact of quality improvement in a population of treated patients. In addition, variation in how deeply DIAMOND care reaches into each clinic's population of patients being treated for depression may be important for consideration of the relative value of this Initiative for medical groups and health plans contemplating investment in DIAMOND.

The 42 month period during which outcomes were measured put the study at risk for external events to impact patient outcomes (history threat to internal validity). The staggered implementation of the intervention should ameliorate some of this risk because at any point in time outcome data were being gathered from patients treated in pre- and post-implementation clinics. As such, historical events would impact outcomes in both patient groups with the relative impact varying as a function of the increasing ratio of post- to pre-implementation patients over the measurement period.

To take advantage of the opportunity offered by the DIAMOND Initiative to learn about the effectiveness of proven clinical trial strategies when adapted into every-day practice, we needed a study design and recruitment process designed around the implementation strategy and feasible within research funding and implementation timelines. Collaboration with Initiative leaders from the early planning stages allowed us to help shape strategy and timelines in a way that was compatible with research and to submit a proposal to NIH in time to be funded before the initiative began. We believe these challenges are common to prospective evaluations of natural experiments and others facing these challenges may benefit from our deliberations and experience. We were fortunate to have a natural experiment with staggered implementation. However this feature is also helpful to those implementing a large scale change in care, so researchers may be able to suggest staggered implementation when it is not already being considered. Finally, our success in recruitment

demonstrates to others that complex recruitment in prospective evaluations of natural experiments is feasible.

Acknowledgments

References

1. Solberg LI, Glasgow RE, Unützer J, et al. Partnership Research: A Practical Trial Design for Evaluation of a Natural Experiment to Improve Depression Care. Med Care. 2010; 48:576–582. 510.1097/MLR.1090b1013e3181dbea1062. [PubMed: 20508531]

2. Gilbody S, Bower P, Fletcher J, et al. Collaborative care for depression: a cumulative meta-analysis and review of longer-term outcomes. Arch Intern Med. 2006; 166:2314–2321. [PubMed: 17130383]

3. Williams JW Jr. Gerrity M, Holsinger T, et al. Systematic review of multifaceted interventions to improve depression care. Gen Hosp Psychiatry. 2007; 29:91–116. [PubMed: 17336659]

4. Korsen N, Pietruszewski P. Translating evidence to practice: two stories from the field. J Clin Psychol Med Settings. 2009; 16:47–57. [PubMed: 19238525]

5. Williams MD, Jaeckels N, Rummans TA, et al. Creating value in depression management. Qual Primary Care. 2010; 18:327–333. [PubMed: 21114913]

6. Strom BL. Methodologic challenges to studying patient safety and comparative effectiveness. Med Care. 2007; 45:S13–15. [PubMed: 17909371]

7. Glasgow RE, Magid DJ, Beck A, et al. Practical clinical trials for translating research to practice: design and measurement recommendations. Med Care. 2005; 43:551–557. [PubMed: 15908849]

8. Speroff T, O'Connor GT. Study designs for PDSA quality improvement research. Qual Manag Health Care. 2004; 13:17–32. [PubMed: 14976904]

9. Brown CA, Lilford RJ. The stepped wedge trial design: a systematic review. BMC Med Res Methodol. 2006; 6:54. [PubMed: 17092344]

10. Handley MA, Schillinger D, Shiboski S. Quasi-experimental designs in practice-based research settings: design and implementation considerations. J Am Board Fam Med. 2011; 24:589–596. [PubMed: 21900443]

11. Tunis SR, Stryer DB, Clancy CM. Practical clinical trials: increasing the value of clinical research for decision making in clinical and health policy. JAMA. 2003; 290:1624–1632. [PubMed: 14506122]

12. Hussey MA, Hughes JP. Design and analysis of stepped wedge cluster randomized trials. Contemp Clin Trials. 2007; 28:182–191. [PubMed: 16829207]

13. Mercer SL, DeVinney BJ, Fine LJ, et al. Study designs for effectiveness and translation research :identifying trade-offs. Am J Prev Med. 2007; 33:139–154. [PubMed: 17673103]
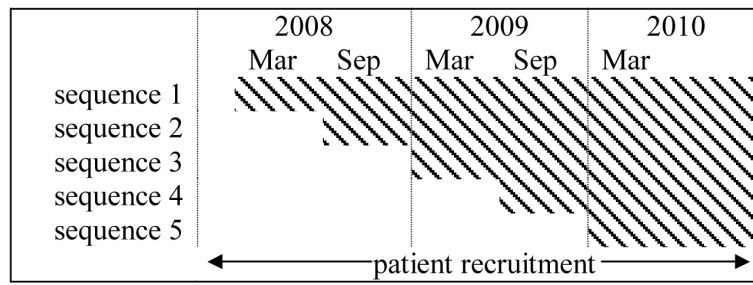
Figure 1.
Patient recruitment period, clinic pre-implementation and shaded post-implementation periods by training sequence.

Table 1

Characteristics of the medical groups and clinics that participated in the DIAMOND Study.

medical group characteristics	min	25th %ile	median	75th %ile	max
patient care sites	1	3	8.5	15	48
primary care sites	1	3	5.0	12	38
MDs	5	17	67.5	136	1700
primary care MDs	5	14	37.0	105	300
psychiatrists	0	0	0	4	38
MH therapists	0	0	1	11	77
patient payer mix					
commercial insurance	4%	39%	50%	67%	80%
Medicare	10%	14%	22%	29%	38%
Medicaid	3%	5%	9%	15%	55%
uninsured	0%	2%	2%	5%	10%
other	0%	1%	4%	16%	73%
care manager credentials					
RN			52%		
LPN			12%		
CMA			32%		
other			28%		
ownership					
health system			60%		
physicians			28%		
other			4%		

clinic characteristics	min	25th %ile	median	75th %ile	max
MDs	1	5	8	17	102
primary care MDs	1	5	7	14	70
PA, NPs	0	1	2	3	20
RNs	0	1	2	6	49

54%

Twin Cities metro

Table 2

Cascade of patients from sampling frame to study enrollment.

sampling frame	24065
made contact	14453
no contact due to:	
incorrect information	5077
not reached in 21 days	4535
screened	9498
not screened due to:	
refused	4074
unable	881
eligible	2435
not eligible due to:	
PHQ9 < 7	3561
not treated for depression	2860
not a DIAMOND clinic	440
no anti-depressant fill	202
consented	2428
enrolled	2423

Table 3

Characteristics of patients in the sampling frame, assessed for eligibility, study eligible and study enrolled.

	sampling frame	eligibility assessed	study eligible	enrolled
n	24065	9498	2435	2423
% of sampling frame		39.5	10.1	10.1
female (%)	69.5	70.7	72.7	72.7
age, M (SD)	45.6 (16.6)	46.5 (15.5)	44.6 (14.9)	44.5 (14.9)
education (%)				
high school or less		26.6	30.8	30.8
some college, technical		36.8	38.0	38.0
4 year degree or more		36.6	31.2	31.2
relationship status (%)				
married		53.6	46.3	46.2
living as married		6.6	7.1	7.1
separated, divorced, widowed		20.6	23.7	23.7
never married		19.2	23.0	23.0
PHQ-9 (%)				
9			34.2	34.3
10-14			37.8	37.7
15-19			18.9	19.0
20+			9.1	9.0
Twin Cities metro (%)	62.3	64.8	64.4	64.6
insurance type (%)				
commercial	66.8	71.1	66.0	66.0
state programs	21.7	17.2	25.1	25.1
Medicare	8.5	9.1	7.0	7.0
other	3.0	2.7	1.9	1.9
clinic training sequence (%)				
1	10.2	11.3	11.4	11.5
2	27.6	27.1	24.4	24.3
3	12.6	12.0	12.7	12.8
4	13.7	13.9	14.6	14.5
5	15.7	15.6	16.8	16.7
non-implementation	20.3	20.1	20.1	20.2
clinic implementation (%)				
pre-implementation	26.7	30.5	29.9	29.6
post-implementation	50.3	49.8	51.4	51.7
non-implementation	20.0	19.7	18.6	18.7