Using IT to improve access, communication, and asthma in African American and Hispanic/Latino Adults
NCT 02086565

June 25, 2018

METHODS AND DATA ANALYSIS

Study Overview

The overall study design has been previously published.[1] We describe methods in full here, and in more detail in the attached appendix (Appendix A). Adults with uncontrolled asthma, generally unfamiliar with the patient portal, were recruited from outpatient primary care and specialty clinics of two health systems that serve predominantly but not exclusively low-income urban neighborhoods. Participants were trained in the use of the portal: Portal Training (PT). Participants were then randomized 1:1 to home visits (HV), from a Community Health Worker (CHW, described below), to encourage use of the portal and to facilitate care coordination or no home visits. The project was approved by the University of Pennsylvania Institutional Review Board and was registered with ClinicalTrials.gov (NCT02086565).

Study design

We chose a randomized controlled trial **(HV + PT versus PT)** because randomization reduces the likelihood of bias from both known and unknown confounders.[2] The longitudinal design supports baseline measures as within-subject controls and the multicenter recruitment also enhances the generalizability to a diversity of social circumstances.

The choices for the randomized trial, portal training (PT) and home visits (HV), were inspired by the RE-AIM conceptualization framework[3] and the need for implementation research.[4] Implementation of portal training and home visits were hypothesized to improve access and communication resulting in improved patient-centered outcomes including asthma control and asthma-related quality of life.

Those randomized to PT+HV worked with the same CHW who conducted both portal training sessions. At each home visit, the CHW ensured care coordination, including obtaining an asthma action plan that linked patients' home and community with the clinic. The CHW provided standard asthma education, promoted communication with providers, encouraged appointment keeping, and facilitated familiarity with health information technology.

Study setting

Study sites were chosen to recruit the most diverse group of patients who satisfied enrollment criteria and whose clinical experience was diverse (primary care versus specialist care, different health practices (University of Pennsylvania and Temple). Participating sites included two family medicine, two general internal medicine, two pulmonary, and one allergy-immunology outpatient practice from the University of Pennsylvania Health System (UPHS). Recruiting from these practices in the past had consistently yielded a population that was more than 65% African American and more than 65% female. More than 1000 patients with asthma, prescribed inhaled steroids, and living in low income neighborhoods were identified in initial screening. Because University of Pennsylvania clinics are not in Hispanic/Latino areas of Philadelphia, we recruited and enrolled patients from Hunting Park Adult Medicine, a primary care practice serving mainly Spanish-speaking patients, more than 85% Hispanic/Latino, and with more than 700 patients with asthma on their problem lists.

Participants

We planned to identify 300 adults living in low-income neighborhoods with moderate to severe asthma requiring urgent care and little previous exposure to the portal.

Inclusion criteria: We recruited adults, only one from a household:

- 18 years or older.
- living in a Philadelphia neighborhood in which at least 20% of households have incomes below the federal poverty level.
- having a doctor's diagnosis of asthma.
- prescribed an inhaled corticosteroid-containing medication.
- who required prednisone or had an ED visit or hospitalization for asthma in the past year.
- who had not previously signed in to a portal more than 3 times.
- who were patients in a participating clinic.

Exclusion criteria

Severe psychiatric or cognitive problems (e.g., obvious mania, schizophrenia, significant mental retardation) that would make it impossible to understand and carryout this protocol.

Recruitment

Using guidance and policies of the supervising institutional review board and the participating health care systems, we recruited patients as follows. After explaining the protocol in staff meetings of participating sites, we received lists of potentially eligible patients (age $\geq$ 18 years, asthma in problem list of the electronic health record, prescribed an inhaled corticosteroid, having an address in one of the low-income neighborhoods) for whom further screening was needed. First we sent "opt-out" letters to clinicians asking for permission to contact their patients. If providers either did not respond to two letters or gave permission, we sent letters to potential participant patients asking to contact them for screening. If patients gave permission, or did not respond, we called the potential participant or approached them at a clinic visit to explain the study and ask for permission to screen for eligibility. According to their preference, we screened potential participants at their clinic or in their home. Data collectors completed a screening form for each patient contacted for recruitment which included a section to document reasons for declining. The following are reasons for declining: inability to commit to study timeline, disinterest in research participation, and patient illness. Participants all signed informed consent to enroll in the study and to undergo Portal Training and data collection. Those randomized to home visits signed a second consent form informing them they were also randomized to home visits and giving permission for these visits.

Randomization and allocation concealment

Randomization was stratified by clinical site. To maintain allocation concealment, we implemented randomly permuted blocks with varying block sizes (2 to 4). For each clinic we prepared opaque envelopes and instructions on how to preserve allocation concealment until the patient gave consent and the envelope was unsealed. Randomization programs and envelopes were prepared under the supervision of one of the project statisticians.

Interventions and Comparators: Portal Training and Home Visit with Community Health Worker

Community Health Worker (CHW)

CHWs live in the community and are familiar with the environment; are acceptable to patients; and connect patients to health education, services, and the health care setting.[5-7] In our Community Asthma Prevention Program (CAPP), we identified and trained CHWs as lay health educators to provide home asthma education and environmental intervention. Over 20 years, CAPP CHWs have visited 3000+ families with children[8] in both African American (West Philadelphia) and Hispanic/Latino (North Philadelphia) communities, the same communities for the current project. These CHWs have a high school education or greater with at least three years' work experience and personal knowledge of asthma. They bring their knowledge as a community resident to suggest local resources to meet daily needs. CHWs are able to establish relationships with participants and their families to promote better asthma self-management, connection with community resources, communication with providers and mitigation of asthma triggers. CAPP's CHWs have functioned not only as home visitors, but also as asthma navigators who have been integrated into the health care team, providing a direct link between families and clinicians. Our experience has allowed us to develop an excellent support system including supervision, training, and systems for scheduling visits and data entry.[9] The retention rate of participating families in CAPP has been above 90%. [10] Home visits for children with asthma often include tailored asthma education and environmental remediation of allergen and irritant environmental exposures.

In the current intervention, each home visit included 1) reinforcement of the use of the portal and 2) care coordination (Table 1). Our activities were unique in utilizing CHWs for home visits for promotion of portal use with the intention of linking the patient to the medical practice, thereby improving patient-physician communication, a known contributor to asthma disparities. The home visits focused on communication and access, the same goals as potentially achievable through portal use.

Table 1. Activities and data collection (DC) for PT+HV and PT groups.

Time	PT+HV	PT
DC- Week 0	Participant meets with a member of the research team for consent, enrollment, randomization, and baseline data collection. This data collector is not the CHW who provides PP1, PP2 or HVs (if participant is randomized to HV).	
PP1 (within 2 weeks of enrollment)	A CHW demonstrates accessing internet if necessary, ensures patient's internet access and request for an activation code for PP, and demonstrates use of PP with a "test" patient. Participants are shown how to complete 7 tasks and how to ask for on-line help and to read FAQs.	
PP2 (within 2 weeks of enrollment)	Patient activates their PP and goes through exercises with the same CHW, using the PP to obtain and ask for information (the 7 tasks demonstrated at PP1). Participants are again shown how to ask for help & to read FAQs. They are given a printed card with a phone number for the Help Line. They are shown how to access medical information through the PP.	
HV1 (within 2-4 weeks of enrollment)	Re-Introductions & conversation- CHW is the same CHW who supervised PP1 and PP2. CHW meets patient and family and asks how they are doing. Care Coordination. The CHW: • asks patient to produce medications and explain doses • asks about smokers in home and offers a QUIT number if needed • reviews participant's baseline environmental survey • asks where emergency medical information is kept, emergency telephone numbers • assists as needed as patient drafts an asthma action plan for review/revision by patient's asthma care provider • asks when next appointment with asthma doctor is & assists with scheduling if necessary • asks about refills needed • asks about patient Care Coordination goals CHW reviews use of PP and how to access it and the internet including access to health information resources • shows how to access asthma education materials with the PP	Email/phone to patient encouraging sign in to PP.
HV2 (between weeks 4-7)	CHW inquires how patient and family are doing. THEN CHW and participant: • review Action Plan and care coordination goals, medications, refills, appointments etc. CHW and patient review use of tablet and PP: • Problems with use • Complete at least 2 tasks with PP • CHW demonstrates how to "google" for information	Email/phone to patient encouraging sign in to PP.
HV3 (between weeks 6-11)	CHW inquires how patient and family are doing. Review care coordination, action plan, goals, medications, and relevant community resources. Review use of tablet and PP: • Problems with use addressed • Complete at least 2 tasks with PP • Show how to email if not known already	Email/phone to patient encouraging sign in to PP.
DC-week 12	Participant called by a member of the research team for data collection.	

Table 1 (continued). Activities and data collection (DC) for PT+HV and PT groups.

Time	PT+HV	PT
HV4 (between weeks 23-27)	CHW inquires how patient and family are doing. Review care coordination, action plan, goals, medications, and relevant community resources. Review use of tablet and PP: • Problems with use addressed • Review how to "google", email, and how to get weather and latest news on the internet	Email/phone to patient encouraging sign in to PP.
DC – week 28	Participant meets with a data collector for data collection which includes spirometry.	
DC-week 36	Participant called by a member of the research team for data collection.	
DC week 48-50	Participant meets with a data collector for data collection which includes spirometry.	
DC every 12 weeks	Participant called by a member of the research team for data collection.	

CHW= community health worker, DC= data collection, PP= patient portal, PT= patient portal training, PT+HV= patient portal training plus home visits, HV1= first home visit, etc.

Components of the Interventions

Portal Training and Home Visits represent different approaches to improve patient-provider communication and care coordination as well as access. Although vastly different interventions, they provide potential synergistic benefit particularly for vulnerable patients exposed to poverty and with prevalent comorbidities.[11] Although portals have been widely employed and undergo continuous refinement,[12-14] their benefits in adult inner city asthma patients remains for investigation.

We hypothesized that home visits may be especially effective for low-income inner city patients facilitating two-way patient-clinician communication; informing clinicians of important environmental, social, and medical barriers to asthma self-management; and reinforcing medical information from clinicians to patient.[11] We also hypothesized that the home visit may be especially effective for those who speak English as a second language, in our sample, mostly Spanish-speaking patients, or those with low numeracy and health literacy. We had piloted the home visits in families with children[8] and the activities to be performed in home visits in adults.[10]

All participants received Patient Portal Training (PT) and those randomized to home visits additionally were scheduled for four Home Visits by a CHW to take place over 6 months: at 2-4 weeks, 4-6 weeks, 6-11 weeks, and 23-27 weeks following randomization. All participants were to be followed for at least a year (Figure 1).

Figure 1. Planned time course for participation and data collection.

Figure 1. Planned time course for participation and data collection

PT= portal training, PT+HV= Portal Training + Home Visit,
PT1= 1st portal training session, HV1= 1st home visit
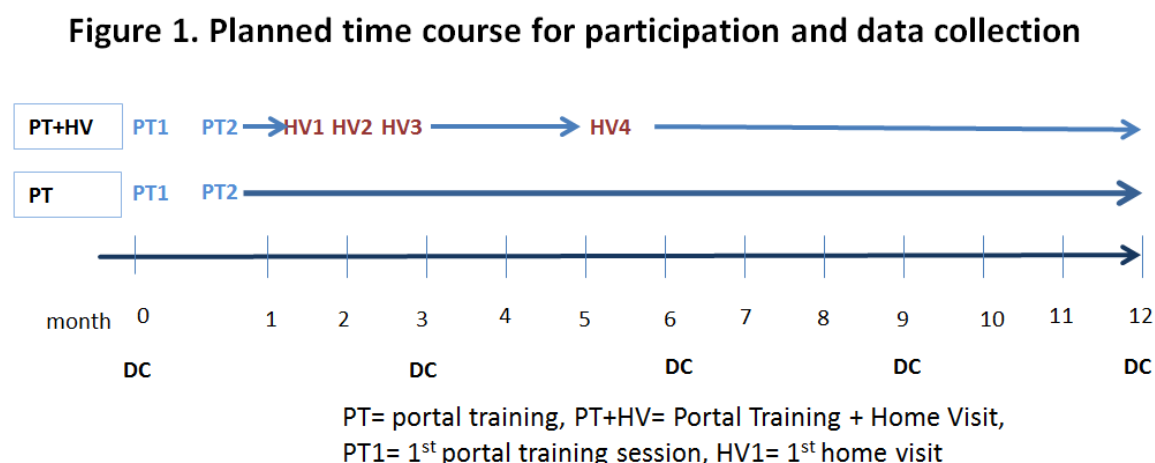
Patient Portal Training (PT)

We ascertained portal familiarity and interest through focus groups of patients and providers exploring their knowledge of and use of the portal.[11] To develop and pilot portal training, we had asked 10 focus-group patients who did not have portal access to register for it with our help.[11]   All were able to use the point and click mechanism within the portal to accomplish 7 tasks: 1) locate a laboratory test result, 2) look up an upcoming doctor's appointment, 3) learn how to schedule an appointment with their provider (the opportunity to actually make the appointment was offered), 4) locate their medication list, 5) locate their immunization record, 6) determine how to request a refill, and 7) send a secure message to their care team. All were positive that they could accomplish the tasks and found them useful. This protocol became the basis for the two portal training sessions, PP1 and PP2 (Table 1). At PP1 a CHW inquired whether internet access by computer, tablet, or smart phone was available at home, work, community hot spots, or participating clinics and together the participant and CHW confirmed which was closest and most convenient (Figure 1).

The CHW described the portal and its functions, and assisted participants in requesting an activation code. PP2, originally planned for a second visit, often followed PP1 on the same day as more convenient for participants. At PP2, working with the same CHW, participants activated and accessed the portal and performed exercises in using it.

Internet access

To ensure internet access at home, we originally contemplated providing a tablet to all participants. Stakeholders concluded that even with vendor discounts, tablets would require costly internet access not affordable to many and would be subject to theft. Our information technology specialists pointed to the challenges of installation in older homes. Although the Philadelphia Housing Authority, a provider of subsidized housing for low-income city residents, provides internet access, not all participants live in these homes. Philadelphia also has high speed internet discounts to low income families with children, but not all participants lived with school-age children. From patient focus groups and our pilot of portal training, we estimated that half of potential participants had computer access at home or work. Of those with access, only half of these checked email regularly.

We insured internet access for participants at primary care and asthma specialty clinics and in the NIH-funded Clinical and Translational Research Center. The medical facilities of the participating practices are "hot spots." At enrollment we recommended ways to access the internet: nearest library, nearest hot spot (recognizing low income communities have fewer), and if desired for those with smart phones, using smartphone apps.

Home Visit (HV) Protocol

Those randomized to PT+HV worked with the same CHW who conducted both portal training sessions. At each home visit, the CHW ensured care coordination, including obtaining an asthma action plan that linked patients' home and community with the clinic. The CHW provided standard asthma education, promoted communication with providers, encouraged appointment keeping, and facilitated familiarity with health information technology. Home visits were intended to link the adult patient to the medical practice through the portal, to provide tailored care coordination and coaching on portal use, and to communicate and hopefully improve asthma control.[9] These activities were selected after conducting focus groups of patients[15] and piloting.[10,16] The CHW was to empower patients to communicate with their health care providers via the portal, to follow asthma guidelines, reconcile medications, facilitate appointment scheduling, and also control environmental exposures (e.g. tobacco).[9] Each home visit had two parts: 1) care coordination using the portal and 2) improving familiarity with health information technology (Table 1). For care coordination the CHW and participant drafted an asthma action plan for review/revision/approval by the participant's asthma care provider and the participant was asked to schedule through the portal an appointment with the provider in which this plan would be discussed. The American Lung Association Asthma Action Plan format was used and included contact information of the providers, preventative and emergency medications .[17] The CHW helped patients identify specific care coordination goals that could improve asthma control and quality of life and that with patients' permission could be shared with providers.

In previous home visits for children living in the same urban Philadelphia neighborhoods, we found that half of families had at least one smoker in the home.[18] Residential exposures to allergens and pollutants contributed to asthma morbidity. CHWs

facilitated patients' communication through the portal with their asthma clinician about exposures to tobacco smoke, pollutants, potentially relevant allergens, and about comorbidities. CHWs were knowledgeable in community resources and had access to a resource database which includes smoking cessation programs and housing opportunities.

At the first home visit, HV1, designed to occur as early as within 2 weeks of the second portal training session, PP2, the CHW became further acquainted with the participant, if possible met the family, collected information, and conducted a needs assessment. The CHW gathered information on medications in the home to coordinate with the participant's clinician, facilitated appointment scheduling, and if applicable encouraged smoking cessation and reduction of secondhand smoke exposure. From the needs assessment the CHW and the patient created care coordination goals. Examples of these goals chosen by patients were to eliminate asthma triggers from the home, to stop smoking or to exercise more. The action plan was then drafted by the participant with the assistance of the CHW for review, amendment, and approval by the participant's asthma care provider. At the second home visit, HV2, intended to take place within 2-3 weeks of HV1, the CHW reviewed the individualized action plan, asthma control, the goals from the previous visit, and relevant asthma education. Use of the portal was reviewed and encouraged. The third visit, HV3, was to occur approximately 2-4 weeks later. At the final visit, HV4, intended to take place approximately 6 months after enrollment, the home visitor and patient reviewed care coordination and all portal and internet skills (Table 1).

To improve familiarity with health information technology, CHWs reviewed use of the portal as the second part of each visit. In addition, at each visit they taught a relevant activity such as using the internet to access health information, how to obtain educational materials through the portal, how to "Google," how to email, and how to get a weather report.

At the end of each visit the participant and CHW completed a report to the clinician of topics discussed (e.g. action plan, medications, other topics). Although use of the portal was encouraged, CHWs and patients found it more feasible to use paper, which the CHW gave to clinic personnel for the clinician's mailbox. Goals for the participant to pursue with the asthma doctor to achieve asthma control were also reported.

Quarterly phone calls for the remaining 6 months were planned with information on portal training to be reviewed along with information from the prior home visits. The use of the portal was encouraged.

Study Outcomes

The outcomes selected to be meaningful to patients, were among the outcomes recommended by the recent Asthma Outcomes Workshop (Table 2).[19] The primary outcome, asthma control, reflecting symptoms over the past week, was measured by the 7-item Juniper Asthma Control Questionnaire (ACQ).[20-22] The score is the mean of all responses (0=total control, 6=extremely uncontrolled). The minimally important clinical difference is or change is 0.5. A score >1.5 is considered inadequate control.[23] Asthma-related quality of life (AQOL) was measured with the Mini Asthma Quality of Life Questionnaire (AQLQ).[24-26] This 15-item questionnaire has a 7-point response scale that provides a mean summary score. A 0.5-unit change is considered clinically meaningful.[26] The AQLQ has been shown to be a useful indicator of AQOL in low-income adults.[27] Hospitalizations including ICU admissions, ED visits, urgent medical visits (scheduled < 24 hours in advance), prednisone bursts (a new prescription of prednisone or an increase in an already-prescribed dose for an asthma exacerbation), and other medical visits were obtained by self-report because not all such events occur within our health system. The medical records were examined for documentation as feasible. Spirometry was obtained using American Thoracic Society procedures for FEV1 and FVC.[28]

Table 2. Summary of data collection.

Measure	At enrollment	Quarterly*	12 months
Outcomes			
Asthma control	X	X	X
Asthma-related quality of life	X	X	X
ED visits for asthma or any cause	X	X	X
FEV1++	X	X	X
Prednisone bursts+	X	X	X
Hospitalizations for asthma or any cause	X	X	X
Mediators			
Appointments kept with asthma doctor	X	X	X
Use of patient portal	X	X	X
Inhaler Adherence Scale#	X	X	X
Moderators (Baseline)			
Sociodemographics	X		
Computer at home	X		
Convenience of internet access	X		
Educational attainment	X		
Health literacy –S-TOFHLA, ANQ, eHealth Literacy	X		
Comorbidities	X		
Depressive thoughts	X		
Primary language	X		

*Quarterly data collection occurs at months 3, 6, 9, and 12 and every 3 months after month 12 until the end of the study.

** We collect outcomes quarterly, at 12 months, and every 3 months until the study ends. Final visit is the final visit of data collection that occurs after the 12-month collection.

+Prednisone bursts are defined as a new prescription or an increase in dose of prednisone.

++FEV1 could be collected for only 203 patients owing to non-availability of spirometry equipment at the start of the study.

#The Inhaler Adherence Scale asks participant's report of non-adherence.[29,30] It is a 6-item scale with each item scored 1 (non-adherence) or 0 (adherence. The range is 0-6, with 0 being optimal adherence.

Time frame for the study

   All participants were to be followed for data collection every 3 months for at least 1 year (please see the Home Visitor Protocol sub-section and Figure 1). We chose a 3-month interval for data collection, as the Expert Panel Report uses 3 months for making a change such as "stepping down."[31] We chose the times for home visits, based on Dr. Bryant-Stephens' experience with home visits in children. She found it was important to have home visits early in participation.[32] We wanted the last 6 months to be for observation to assess sustainability. For reasons for withdrawals or becoming lost to follow-up, please see Figure 2.

Data collection and sources

   The IRB-approved consent and all communications were read/spoken in English or Spanish as preferred by the potential participant. Data collection was accomplished by a CHW, called a Data Collector, who did not act as the home visitor if the patient was so assigned. This choice of Data Collector reduced any pressure on the patient to answer in a way to please the researcher CHW. Baseline questionnaires assessed socio-demographics, asthma severity, health literacy, and comorbidities (Table 2). The first and last data collection visits were in person so that spirometry could be measured. For the other data collection times, we offered a phone or email or in-person visit to minimize the burden of and effect on participation in a study. The Data Collector entered responses either directly into a secure, encrypted, web-based database REDCap (Research Electronic Data Capture) or on paper for transfer to REDCap later.[33] Data Collectors used tablets (iPads) equipped with Wi-fi but internet access was difficult at times and CHWs noted many patients were not comfortable with having their information entered onto an iPad.

   For each patient we collected their contact information and that of three persons. We also were able to read the electronic record to determine when patients had appointments and to meet them there. Additionally, we wrote letters, visited homes, and left postcards.

Data management

REDCap (http://www.project-redcap.org/) allows data attribution and audit capabilities, integrity checks, real-time validation, data storage and backup, and export functions.[34] Whenever any response of the patient required elaboration, the interviewer entered comments into this database.

At weekly meetings, the community health workers, project manager, the staff who entered data into REDCap, the project statisticians, and the principal investigators discussed data collection problems. The principal investigator then reviewed all reported adverse events. These meetings then led to the development of project rules for interpreting any of the questions that presented difficulties for patients to answer.

A set of 30 tables stored in the REDCap database contain data from interviews and phone assessments of patients over time (Table 2 and Appendix A1).

Analytical and Statistical Approaches

Patient Portal usage

Portal usage data came from two sources: Penn Medicine and Temple Health. Details on our data management approach to the Portal system appear in Appendix A2.

Patient activity within Patient Portal Episodes

After combining the two EPIC Portal databases from the two participating institutions, there were 61 activities that appeared for the patients in our sample (Appendix A2, Tables 1 and 2). These activities we then combined into a set that corresponded to the basic Portal usage that participating patients were trained to use (Appendix A2, Table 3). Finally, we calculated the frequency of Portal usage episodes during the duration of the patient's study involvement from randomization to the last data collection (subtracting off any Portal usage that coincided with patient connecting with the community health workers or study personnel). We also calculated the frequency of all portal usage by the rate of usage over time.

Mediation

Mediators explain how the interventions influence asthma outcomes. For this study we designed three candidate mediators to measure communication, access to the health care system, and improved inhaler usage: (1) the rate of portal use over time, (2) appointments kept within 6 months

of a data collection time, and (3) inhaler use according to best practices.[30] To determine whether mediation might be present, we first estimated the association between the randomized assignment (portal only versus portal + home visits) and the hypothesized mediator. In the absence of an association, there could be no indirect effect and no mediation. Details appear in Appendices D3 and D4.

(1) ***Portal usage*** as a potential mediator. Because patient follow-up varied depending on the time from randomization to last data collection interview, portal use was based as a rate or intensity over time. We hypothesized that the home visitor would lead to greater use of the portal, and that with greater portal use, the patient would have better outcomes.

(2) ***Appointments kept***. For health care usage, we hypothesized that access to medical care would improve in that patients assigned to home visitors would be more likely to make and keep regular appointments. According to prevailing guidelines,[31] a patient with moderate or severe asthma should have regular appointments, and at least every 6 months. To effect that standard using data on appointments, we merged the data on appointments with the dates of the data collections and then determined for each collection interview whether the patient had kept an appointment in the 6 prior months. Using appointments kept as a time-varying exposure, we estimated the difference in expected value of asthma control at 0 months (date of randomization) and at 12 months assuming alternatively that patients had kept regular appointments versus had never kept regular appointments. Details of these analysis (and results) appear in Appendix A4.

(3) ***Use of inhaler***. The inhaler adherence scale[30] is scored from 0 through 6. We fit a longitudinal model similar to the one for our outcomes of asthma control and quality of life to estimate the association of treatment assignment and inhaler adherence scale over time. Details appear in Appendix A5.

Heterogeneity of treatment effect—identification of subgroups: Effect modifiers are baseline variables or groups of factors, gleaned from the literature, hypothesized to affect the intervention-outcomes relationships.[35,36] Recent literature on heterogeneity of treatment effect suggests that one-by-one testing or estimation of candidate effect

modifiers leads to underpowered contrasts, excessive reliance on p-value testing, inadequate pre-specification of candidates and their rationale, and no attention to multiple comparisons.[37] To address these criticisms, we grouped candidate measures into several clearly pre-specified themes from which profiles or latent classes are derived. We implemented a latent class analysis for each group of candidate measures to distinguish profiles of common responses to the set of variables.[37,38] Then the degree of effect modification of the association of the intervention (home visitors) and the key outcome (asthma control) was estimated for each candidate modifier group by testing interactions between the candidate moderator variable and the randomization indicator. For those candidate effect modifiers that approached conventional levels of statistical significant, we estimated the intervention effect and 95% confidence intervals for each level of the effect modifier. The nine pre-specified candidate effect modifier groupings appear in Table 3. Details appear in Appendix A6.

Table 3. Relevant patient subgroups for targeting the initiative and their data elements.

Relevant subgroup	Data Elements
(1) Primary language	Primary language is Spanish
(2) Primary care vs specialty practices	Internal medicine or family medicine versus allergy-immunology or pulmonary
(3) Age	18-39 years, 40-49 years, $\geq$ 50 years
(4) Skills that would support use of portal and asthma self-management	Numeracy* Literacy* Education Computer literacy* Inhaler technique
(5) Social Community barriers	Food or clothing inadequacy MOS Social Support** Violence
(6) Trust of medical personnel	Patient portal preserves privacy
(7) Depression and chronic disease load	Depression Diabetes Hypertension High cholesterol Obesity Cancer Current Smoker
(8) Asthma severity	Hospitalizations ED visits Intubation Years taking ICS Prednisone (Days/week)
(9) Home environment	Crowding at home (number of rooms, number of people at home) Been without housing in last 6 months Utilities shut off in the last 6 months Times moved in last 12 months Exposure to second hand smoke

*Three literacy measures were used: Asthma Numeracy Questionnaire,[39] the Short Test of Functional Health Literacy in Adults,[40] and Electronic Health Literacy.[41] Electronic health literacy was measured with the eHEALS, the eHealth Literacy Scale, an 8-item measure of knowledge, comfort, and perceived skills at finding and evaluating electronic health information.[41] It has been used in low-income patients.[42] Each item is measured with a 5-point likert scale and the eHEALS score is the mean of each item.

**The Medical Outcomes Study Social Support Survey.[43,44]

Our working hypothesis was that patients with comorbidities and with limited computer literacy might (a) not benefit from portal education at all, and (b) might have additional benefit with the home visitor intervention. Further details appear in Appendix A6.

Statistical Analysis plan

Testing within-group effects over the time of the study allowed us to assess the effect of each intervention. In brief, the within-group change in the PT group allowed measurement of PT alone. Subtracting any change over time in the PT+HV group from the PT group allowed estimates of the additional effect of home visits.

Intention to treat approach –primary analysis

The primary analysis was "as randomized" (intent to treat), with the assumption that any dropout visits were missing completely at random (MCAR). The estimand of interest was the difference between the home-visitor versus portal-only groups in the change in outcomes over time. We attempted to use mixed effects models, which make less strong missing at random (MAR) assumptions, with varying success. (Details appear below and in the analysis-specific appendix, Appendix A).

Modeling approach – flexible modeling of time

Irregular data collection times (see results) required models that used outcome time as a continuous measure. For that reason, analysis models implemented spline based marginal (GEE) (MCAR assumption) and mixed effects (MAR assumption) longitudinal models, both identity/Gaussian, log/Poisson, and log/gamma with marginal splines to estimate expected values of outcomes at 0 months and 12 months and outcome changes between 0 and 12 months. We used all data collection times and the actual times. We did not artificially categorize data collection times to equate to the preplanned 0, 12, 24, 36, and 48 weeks from randomization. The contrast (estimator) of interest was the difference between the treatment groups in the changes in expected values from 0 months to 12 months. Examples in the statistical literature support this approach.[45] All confidence bounds were estimated using 999 bootstrap resamples (percentile-based). By using 999 samples, confidence bounds are from the 25th and 975th order statistics without interpolation. Details appear in Appendix A7.

Model form

We used both marginal and mixed effects models, with the marginal model being the primary approach owing to convergence problems with mixed effect models (noted below). [46] By including splines for time and group-by-time interaction terms, these models could produce the expected values at 0 and 12 months and predicted values of interest.

For the primary outcome of asthma control, the skewed distribution of outcomes necessitated that we use log gamma models. The treating clinic was a stratifying factor in all models to account for the stratified randomization.

Alternative approaches for estimating asthma control

Differences between intervention groups in expected values at given follow up times, although statistically appropriate for comparing groups, lacks an immediate connection to the manner in which clinicians might want to interpret results. For that reason, we pre-planned two methods for translating findings into clinically useful metrics.

(1) Fraction of patients who have achieved asthma control (a level below 1.5 on the scale):

We avoided dichotomizing data, an ad hoc approach, but instead used mixed effects linear model and predictions to report the change in the fraction of patients who achieved adequate control. This prediction accounts for patient variation about the expected value and in theory represents a better estimate of the patient's true outcome (asthma control) than does the patient's observed value. Once we estimated each patient's predictions at baseline and at 12 months, we dichotomized these model-based predicted values to estimate the fraction of patients who achieved asthma control at baseline and at 12 months in each group. This is not the primary analysis, however, but represents an alternative demonstration of intervention effectiveness. Details appear in Appendix A8.

(2) Minimally important improvement: For estimates of the improvements at the patient level on the asthma control (Juniper) scale, the minimally important difference is 0.5.[23,47] Again, the model-based predicted values are the best estimates at the individual level of the level of asthma control at a given time and account for measurement error as well as variation across individual patients. We compared changes in individual predictions of asthma control at 0

months and 12 months and counted the number of predictions that improved by at least 0.5. Methodological details appear in Appendix A8.

Sensitivity analyses

We planned the following sensitivity analyses, and pre-specified the endpoint for these analyses to be the between- group differences of the change over time of asthma control from baseline to 12 months.

Sensitivity analysis #1. Selection bias from loss to follow-up – covariate adjustment

In keeping with recommendations, [48,49] we included in our longitudinal analysis model (described previously), in addition to the randomized treatment assignment and the pre-planned stratification variable (clinical site), the covariates that might be related to dropout and the primary outcome. Appendix A9.

Sensitivity Analysis #2. Nonignorable timing of data collection -- Global sensitivity analysis to irregular visit times and dropout

The sensitivity analyses we have described (#1) make assumptions that the dropout and visit times are not related to the values of outcomes, i.e., that dropout is at random and that visit times are ignorable. Our dropout was limited, but irregular interview times were common. Global sensitivity analyses are now available but only when visit times are regular.[50] Methods for irregular visits are awaiting methodological development (Daniel Scharfstein, personal communications 04/2017). Details appear in Appendix A7.

Sensitivity Analysis #3. Non-adherence to home visit protocol

As randomized analyses do not answer the question of the effect of an intervention if the patients adhere to the protocol, simplistic approaches to per-protocol analysis can be biased. Although weighting methods can generate unbiased estimate of the effect of a treatment under full adherence, [51-53] we could not adopt these methods because they assume regular observation times. Use of weighting approach for irregular outcome times awaits methodological development. For that reason, we implemented an alternative based on instrumental variables.

Instrumental variable methods adjusted for differences in the characteristics of patients who received all home visits and those who did not.[54] In our case the instrument was

randomization at baseline. These methods require several assumptions.[55] Details along with the specific method we used (two- stage residual inclusion) appear in Appendix A10.

Missing Data

Missing baseline data

Missing covariate data were infrequent, but some patients were missing an element of a questionnaire, or responded "did not know".   We created latent classes of covariates for both confounder control and for estimating effect modification (please see below).

Missing data for effect modifiers

Latent class modeling was used both to reduce the number of covariates used to adjust for baseline factors and to examine possible moderators of the exposure outcome relationship (Table 3). If we were treating each covariate as a potential confounder or as an individual effect modifier, we would need to perform formal imputation of missing data elements of an item that contributes to a latent class or drop patients who had one or more missing items. By contrast, latent class modeling does not require complete data on every factor that contributes to a class. Mixture models generated latent classes and probabilities of class membership for all patients which were used to assign patients to these classes even if some data elements used in the mixture model were missing. This approach assumed missingness completely at random (MCAR), or at least missingness at random given the values of other covariates. In addition, the latent class method allows for the high degree of collinearity among the candidate factors. In short, latent class modeling both accomplished data reduction (too many covariates for adjustment) and facilitated identification and testing of patient subgroup differences in treatment effect.

Statistical power

Estimates of power for longitudinal data must account for the added power of following individuals over time. To estimate power of our somewhat complex design, we resorted to statistical simulations, implemented in Stata v 13 (StataCorp, College Station, TX), in which we varied the correlations of outcome within individual over time. Using the asthma control measure as an example, we estimated that power of a balanced design with 300 evaluable subjects was adequate (from 0.82 to 0.87) to detect a difference between groups of 1/3 of a

standard deviation of asthma control. According to Juniper's work on asthma control, this detectable difference is less than the minimally important clinical difference.[22,47] Additional simulations, for effect modification, suggested good power (0.9) to demonstrate a greater improvement among low literacy patients with home visitors. For all results, confidence intervals reflect the post hoc power.

Programs and software

Analyses were performed in SAS v9.4 (SAS Institute Cary NC), Stata v 15.0 and 15.1 (College Station Texas), and M Plus v7.0 (Los Angeles CA).

Trial monitoring

An external Data Safety Monitoring Board reviewed progress and data every six months and monitored adverse events and serious adverse events (unexpected ED visits and hospitalizations). Because patients had moderate or severe asthma, it was expected that ED visits and hospitalizations for these and patients' co-morbid conditions would occur but perhaps change as a result to the interventions.

Changes to the original study protocol

There were no substantive changes to the original study protocol. Small changes made it easier for patients to adhere to the protocol. For example, if it were easier for the patient, we went to their home for enrollment. For some patients the first portal training session and the second were given on the same day for those for whom a single training session was more convenient.

REFERENCES

1.   Apter AJ, Bryant-Stephens T, Morales KH, et al. Using IT to improve access, communication, and asthma in African American and Hispanic/Latino Adults: Rationale, design, and methods of a randomized controlled trial. Contemporary clinical trials 2015;44:119-28.
2.   Rothman K, Greenland S. Modern Epidemiology. Philadelphia: Lippincott Williams and Wilkins; 1998.
3.   Glasgow RE, Vogt TM, Boles SM. Evaluating the public health impact of health promotion interventions: the RE-AIM framework. American journal of public health 1999;89:1322-7.
4.   Glasgow RE, Vinson C, Chambers D, Khoury MJ, Kaplan RM, Hunter C. National Institutes of Health approaches to dissemination and implementation science: current and future directions. American journal of public health 2012;102:1274-81.
5.   Singh N, Squier C, Sivek C, Wagener M, Nguyen MH, Yu VL. Determinants of compliance with antiretroviral therapy in patients with human immunodeficiency virus: prospective assessment with implications for enhancing compliance. AIDS Care 1996;8:261-9.
6.   Peretz PJ, Matiz LA, Findley S, Lizardo M, Evans D, McCord M. Community health workers as drivers of a successful community-based disease management initiative. American journal of public health 2012;102:1443-6.
7.   Krieger JW, Takaro TK, Song L, Weaver M. The Seattle-King County Healthy Homes Project: a randomized, controlled trial of a community health worker intervention to decrease exposure to indoor asthma triggers. American journal of public health 2005;95:652-9.
8.   Bryant-Stephens T, Kurian C, Guo R, Zhao H. Impact of a household environmental intervention delivered by lay health workers on asthma symptom control in urban, disadvantaged children with asthma. American journal of public health 2009;99 Suppl 3:S657-65.
9.   Krieger JW, Philby ML, Brooks MZ. Better home visits for asthma lessons learned from the Seattle-King County Asthma Program. American journal of preventive medicine 2011;41:S48-51.
10.   Apter AJ, Wan F, Reisine S, et al. Feasibility, acceptability and preliminary effectiveness of patient advocates for improving asthma outcomes in adults. J Asthma 2013;50:850-60.
11.   Black HL, Gonzalez R, Priolo C, et al. True Meaningful Use: Technology meets both patient and provider needs. Am J Managed Care 2015;21:e329-e37.
12.   Goel MS, Brown TL, Williams A, Cooper AJ, Hasnain-Wynia R, Baker DW. Patient reported barriers to enrolling in a patient portal. Journal of the American Medical Informatics Association : JAMIA 2011;18 Suppl 1:i8-12.
13.   Goel MS, Brown TL, Williams A, Hasnain-Wynia R, Thompson JA, Baker DW. Disparities in enrollment and use of an electronic patient portal. Journal of general internal medicine 2011;26:1112-6.
14.   Seidman J, Wolf MS. The role of health literacy in health information technology. Innovations in Health Literacy Research. Washington, DC: The National Academies Press; 2011:23-32.
15.   Black HL, Priolo C, Akinyemi D, et al. Clearing clinical barriers: enhancing social support using a patient navigator for asthma care. J Asthma 2010;47:913-9.
16.   Black HL, Gonzalez R, Priolo C, et al. True "Meaningful Use": Technology Meets Both Patient and Provider Needs. The American journal of managed care 2015;in press:In press.
17.   Asthma Action Plan. American Lung Association, 2015. (Accessed July 8, 2015, at http://www.lung.org/associations/states/colorado/asthma/asthma-action-plan.html.)

18.   Bryant-Stephens T. Asthma disparities in urban environments. The Journal of allergy and clinical immunology 2009;123:1199-206; quiz 207-8.
19.   Busse WW, Morgan WJ, Taggart V, Togias A. Asthma outcomes workshop: overview. The Journal of allergy and clinical immunology 2012;129:S1-8.
20.   Juniper EF, O'Byrne PM, Roberts JN. Measuring asthma control in group studies: do we need airway calibre and rescue beta2-agonist use? Respiratory medicine 2001;95:319-23.
21.   Juniper EF, O'Byrne PM, Ferrie PJ, King DR, Roberts JN. Measuring asthma control. Clinic questionnaire or daily diary? American journal of respiratory and critical care medicine 2000;162:1330-4.
22.   Juniper EF, O'Byrne PM, Guyatt GH, Ferrie PJ, King DR. Development and validation of a questionnaire to measure asthma control. Eur Respir J 1999;14:902-7.
23.   Juniper EF, Bousquet J, Abetz L, Bateman ED. Identifying 'well-controlled' and 'not well-controlled' asthma using the Asthma Control Questionnaire. Respiratory medicine 2006;100:616-21.
24.   Juniper EF, Guyatt GH, Cox FM, Ferrie PJ, King DR. Development and validation of the Mini Asthma Quality of Life Questionnaire. Eur Respir J 1999;14:32-8.
25.   Juniper EF, Buist AS, Cox FM, Ferrie PJ, King DR. Validation of a standardized version of the Asthma Quality of Life Questionnaire. Chest 1999;115:1265-70.
26.   Juniper EF, Guyatt GH, Willan A, Griffith LE. Determining a minimal important change in a disease-specific Quality of Life Questionnaire. Journal of clinical epidemiology 1994;47:81-7.
27.   Leidy KN, Chan KS, Coughlin C. Is the asthma quality of life questionnaire a useful measure for low-income asthmatics? American journal of respiratory and critical care medicine 1998;158:1082-90.
28.   Standarization of Spirometry, 1994 Update. American journal of respiratory and critical care medicine 1995;152:1107-36.
29.   Morisky DE, Green LW, Levine DM. Concurrent and predictive validity of a self-reported measure of medication adherence. Medical care 1986;24:67-74.
30.   Dolce JJ, Crisp C, Manzella B, Richards JM, Hardin JM, Bailey WC. Medication adherence patterns in chronic obstructive pulmonary disease. Chest 1991;99:837-41.
31.   Expert Panel Report 3: Guidelines for the diagnosis and management of asthma Bethesda, MD: National Institutes of Health, National Heart, Lung, and Blood Institute; 2007.
32.   Bryant-Stephens T, Reed-Wells S, Canales M, et al. Home visits are needed to address asthma health disparities in adults. The Journal of allergy and clinical immunology 2016;138:1526-30.
33.   Research Electronic Data Capture. 2015. (Accessed 3/26/2015, at http://project-redcap.org/.)
34.   Harris PA, Taylor R, Thielke R, Payne J, Gonzalez N, Conde JG. Research electronic data capture (REDCap)--a metadata-driven methodology and workflow process for providing translational research informatics support. Journal of biomedical informatics 2009;42:377-81.
35.   Holmbeck GN. Toward terminological, conceptual, and statistical clarity in the study of mediators and moderators: examples from the child-clinical and pediatric psychology literatures. J Consulting and Clinical Psychology 1997;65:599-610.
36.   VanderWeele TJ. Explanation in Causal Inference: Methods for Mediation and Interaction. Oxford: Oxford University Press; 2015.
37.   Alosh M, Huque MF, Bretz F, D'Agostino RB, Sr. Tutorial on statistical considerations on subgroup analysis in confirmatory clinical trials. Statistics in medicine 2017;36:1334-60.

38.   McCutcheon AL. Latent Class Analysis. Beverly Hill, CA: Sage University Press Series on Quantiative Applications in the Social Sciences; 1987.

39.   Apter AJ, Cheng J, Small D, et al. Asthma Numeracy Skill and Health Literacy. J Asthma 2006;43:705-10.

40.   Baker DW, Williams MV, Parker RM, Gazmararian JA, Nurss J. Development of a brief test to measure functional health literacy. Patient education and counseling 1999;38:33-42.

41.   Norman CD, Skinner HA. eHEALS: The eHealth Literacy Scale. Journal of medical Internet research 2006;8:e27.

42.   Knapp C, Madden V, Wang H, Sloyer P, Shenkman E. Internet use and eHealth literacy of low-income parents whose children have special health care needs. Journal of medical Internet research 2011;13:e75.

43.   Ware JE, Jr., Sherbourne CD. The MOS 36-item short-form health survey (SF-36). I. Conceptual framework and item selection.PG - 473-83. Medical care 1992;30.

44.   Sherbourne CD, Stewart AL. The MOS social support survey. Social science & medicine (1982) 1991;32:705-14.

45.   Howe LD, Tilling K, Matijasevich A, et al. Linear spline multilevel models for summarising childhood growth trajectories: A guide to their application using examples from five birth cohorts. Statistical methods in medical research 2016;25:1854-74.

46.   Fitzmaurice GM, Laird NM, J.H. W. Applied Longitudinal Analysis, 2nd Edition. Hoboken, NJ: John Wiley & Sons; 2011.

47.   Juniper EF, Svensson K, Mork AC, Stahl E. Measurement properties and interpretation of three shortened versions of the asthma control questionnaire. Respiratory medicine 2005;99:553-8.

48.   Council NR. The Prevention and Treatment of Missing Data in Clinical Trials. Panel on Handling Missing Data in Clinical Trials. Committee on National Statistics, Division of Behavioral and Social Sciences and Education. Washington, DC: The National Academies Press.; 2010:106.

49.   Council NR. The Prevention and Treatment of Missing Data in Clinical Trials Washington, DC: The National Academies Press.; 2010:ISBN: 978-0-309-15814-5.

50.   Scharfstein DO, McDermott A, Olson W, Wiegand F. Global Sensitivity Analysis for Repeated Measures Studies With Informative Dropout: A Fully Parametric Approach. . Statistics in Biopharmaceutical Research 2014;6:338-48 Taylor and Francis

51.   Toh S, Hernan MA. Causal inference from longitudinal studies with baseline randomization. The international journal of biostatistics 2008;4:Article 22.

52.   Toh S, Hernandez-Diaz S, Logan R, Robins JM, Hernan MA. Estimating absolute risks in the presence of nonadherence: an application to a follow-up study with baseline randomization. Epidemiology 2010;21:528-39.

53.   Toh S, Hernandez-Diaz S, Logan R, Rossouw JE, Hernan MA. Coronary heart disease in postmenopausal recipients of estrogen plus progestin therapy: does the increased risk ever disappear? A randomized trial. Annals of internal medicine 2010;152:211-7.

54.   Nagelkerke N, Fidler V, Bernsen R, Borgdorff M. Estimating treatment effects in randomized clinical trials in the presence of non-compliance. Statistics in medicine 2000;19:1849-64.

55.   Jo B. Statistical power in randomized intervention studies with noncompliance. Psychological methods 2002;7:178-93.

Appendix A: Additional description of data analyses and accompanying tables.

Appendix A1. PCORI Visit Schedule and list of case report forms used in REDCap.

PCORI Visit Schedule

Visit form	PP2	V1 baseline	V2 week 12	V3 week 24	V4 week 36	V5 week 48	Home Visits A	B	C	D
Asthma Control (ACON)/ Spanish		x	x	x	x	x				
Adverse Events and Serious Adverse Events (AE)			x	x	x	x				
Asthma History (AST)/Spanish		x								
Center for Epidemiologic Studies Depression Scale (CESD)/Spanish		x								
Concominant Medication (CMED)		x	x	x	x	x				
General Information (GEN)/Spanish		x								
Eligibility (Inclusion/Exclusion) Worksheet (ELIG)		x								
Numeracy (NUM)/Spanish		x								
MOS Social Support Survey Questionnaire (MOS)/Spanish		x								
Randomization (RAND)		x								
Quality of Life (QOL)/Spanish		x	x	x	x	x				
STOFHLA (STOFHLA)/Spanish (administrative)		x								
STOFHLA Score (SCORE)		x								
Spirometry (SPIR)		x				x				
Study Stop (SSTOP)						x				
Inhaler Techniques - MDI (MDI)		x								
Smoke History (SMOKHX)		x								
Standard Visit (STV)/Spanish		x	x	x	x	x				
Inhaler Adherence Scale (INH)/Spanish		x	x	x	x	x				
E-Health Literacy Scale (HLS)/Spanish*		x				x				
Contact Information (CONTACT) (administrative)		x					x	x	x	x
Patient Portal Training 2 (PT2)	x	x								
Patient Portal Training 1 (PT1) (administrative)		x								
Home Questionnaire (HOME)		x								
Home Needs Assessment (HVISIT)							x			
Asthma Care Coordination Goals (GOALS)							x			
Asthma Action Plan (administrative)							x			
Report to Clinician (REPORT)							x	x	x	x
Portal Use Baseline Survey (PORTAL)		x								
Information Technology Survey (INFOTECH)		x								

V5.20150408

Appendix A2. Patient portal in EPIC - Data management details.

The Patient Portal in the Epic electronic health record represents an administrative database that underlies the functioning of the patient portal. To our knowledge it has not previously been used for research purposes similar to those we describe in this report. For that reason, we had to devise an approach to capture these raw data and transform them into information for research purposes.

We first had to determine the correspondence among the items in the Portal menu that the patient typically sees on a computer or iPad or smart phone, the selection of these items, and the data records ultimately stored in the underlying database. To understand the Portal process, we established a simple protocol of menu items to select in a given order and at preset times. Then one investigator (ARL), using his own patient Portal account and this protocol, systemically executed each step with pre-specified time delays (in seconds) between each menu selection. The investigator repeated this process for both the EPIC Portal designed for the personal computer, and the EPIC Portal designed for the iPad and smartphone. Then, we notified the investigator who had access to the Portal database (JTH) to retrieve the database transactions. Finally, we compared the protocol-based menu selection with the resulting data entries to identify what menu selection result in different recorded activities.

The Clarity database is a large subset of data that comes from the PennChart (Epic) application. The Portal data from the University of Pennsylvania system entered the PennChart Chronicles dataset. We transferred the Portal data from the PennChart Chronicles database to Clarity, a Microsoft SQL Server database comprised of over 18,000 tables. After several trials, the "MS data only" option seemed to work best in Crystal Reports in creating a smaller report that can be used as data for analysis.

A special challenge with these types of data arises when the medical record number, which ideally should remain unchanged, does in fact change during the course of the study. Although our study followed patients for a limited time (almost all patients for less than 2 years), we found that medical record numbers did change, and communications with the information technology departments requested multiple searches of data for individual patients.

Appendix A2 (continued). Patient portal in EPIC - Data management details.

Each line of data contained the following fields: Patient name, patient medical record number, date-time of a transaction, the transaction activity, and the transaction action (read or write). The raw data consisted of more than 71,000 lines of data from the two databases across Penn Medicine and Temple Health. All data were linked by medical record number and patient name to a crosswalk file that contained the study identifier, at which time the direct identifiers were dropped for further analysis. All portal data were then merged with the records of home visits and Portal practice visits, and we then excluded any Portal data that could have been associated with training and practice during a visit with the patient.

Sixty-one activities appeared for the patients in our sample (Appendix A2, Tables 1 and 2). These activities we then combined into a set that corresponded to the basic Portal usage that participating patients were trained to use (Appendix A2, Table 3).

The portal data file proceeded through several programs to define four groups of Portal usages:

(a)   Pre-study portal activity – prior to the date of randomization

(b)   Training portal activity – Portal usage that coincided with a study training session in which the home visitors were working with the patients to show them how to use the portal.

(c)   Portal usage during the study – from randomization up to the last visit date (acon_date: the date of an Asthma Control Questionnaire was considered the date of an associated data collection date)

(d)   Portal usage after the study – After the last visit date (acon_date)

Two types of data share the same database in EPIC portal functions. Portal data and Welcome data exist on the same database at Penn. Welcome data result from questionnaires given in the clinics as part of a clinic visit. These might be on pain, for example. The questionnaires are administered at a clinic kiosk. Our analysis identified and excluded Welcome data in two ways. First, we inspected the raw portal data to identify examples in which patients' data in the portal file corresponded with their visits to the health clinics. We discovered that during those visits the patients' portal data listed the activity field of "questionnaire". This

Appendix A2 (continued). Patient portal in EPIC - Data management details.

activity field then became a flag for data entries that corresponded to the Welcome function of the database as contrasted with the Portal system. The existence of these Welcome data did not become evident until well after the start of the study and as part of our investigation of the Portal data. Welcome data do not represent patient-initiated Portal use. These data were excluded for analysis.

We then had a mixture of binary outcomes (used the patient portal vs did not use) and for those who used the Portal, we had distilled the information into "episodes."

Episodes of use of the Patient Portal

An "episode" was characterized by the following criteria:

(1) It occurred between the randomization data and the last data collection time

(2) It did not occur during a visit day

(3) It did not occur at a clinic visit (Welcome data)

(4) It includes all activities beginning with a log on and ending with a log out (or a lapse of 30 minutes). We determined lapses by looking at the data ordered by the system date-time stamp.

Using this information, we counted the number of episodes of Portal usage for each patient over time. Then we calculated the number of episodes per 3 months. Finally, we categorized episode usage as follows.

(1) no usage at all

(2) < 1 per quarter

(3) 1 per quarter up to 1 per month

(4) 1 per month or more

This approach to the analysis of raw portal transaction data resulted in 4 categories with acceptable numbers of patients in each category.

Appendix A4. Appointments kept among asthma patients.

Background

The goals of this analysis were: first, to use treatment assignment as the exposure of interest, in an as-randomized analysis, to estimate the association of treatment assignment and appointments kept over time; and second, in an exploratory observational analysis, to examine the possible association of the time-varying factor of having regular appointments (with treating physicians) and asthma control.   The underlying hypothesis was that the home-visitors might be a positive influence in reminding patients to keep regular appointments, and that regular appointments might in turn lead to improved asthma control among patients.

Data Management Methods

We obtained data on appointments from the two participating health systems (Penn Medicine and Temple Health) on the enrolled patients for the time during their study participation.   With these data, we then identified all appointments that were made and kept by the patients and the dates of those appointments. Then we determined whether a patient had kept an appointment within 6 months of the date of interview.   This variable was binary (0/1) and repeated for each patient over time.   We allowed a kept appointment to satisfy the condition of "appointment within 6 months" for more than one interview date.

Statistical Methods

The proportion of patients with 0 to 5 data collection interviews with a kept appointment within 6 months was compared across intervention groups using a Chi-square test of association.

Using a longitudinal model similar to that used for the main analysis of the association of treatment assignment (home visitor versus portal only) we adjusted for baseline covariates (including treatment assignment). An appointment kept within six months (yes/no) was the time-varying exposure of interest. Time in the model was again fit with marginal splines. Together with a main effect for kept appointment, we added time-by-kept appointment interaction terms. We used a marginal model (generalized estimating equations) with a log link and a gamma family, to be consistent with the other models for asthma control.

Appendix A4 (continued). Appointments kept among asthma patients.

Our estimates of interest were the expected values of asthma control at 0 months and 12 months. Because of the irregular visit times, few patients had visits at exactly 12 months. For that reason, we used predictions from the longitudinal model as of 12 months and then compared them with predictions as of 0 months (date of randomization).

To answer the question of the effect of kept appointments we compared the two scenarios: kept appointments within 6 months of each interview versus did not have any appointments kept within 6 months of the interview. This analysis reflects the more common paradigm for comparing "always exposed" versus "never exposed", where scheduling appointments serves as the exposure. We accomplished this contrast by augmenting the original dataset with two pseudo-datasets, one with all binary values for appointments kept set to 0 and the second with all binary values of appointments kept set to 1. This approach, using data augmentation, is a form of marginal standardization (or predictive margins) as outlined by Korn and Graubard (1999) and as further described by Vittinghoff (2012).

This simplistic approach does not account for several possible sources of confounding. First, we do not factor in time-varying confounders – we adjust only for baseline covariates. Second, we assume that the direction of causation is always from appointment kept to asthma control. The direction of causation could reverse; poorer control could lead to increased need for appointments. We attempted to address that potential problem in two ways. (a) We use the appointment kept from a prior (lookback) period in evaluating it as an exposure for the subsequently reported level of asthma control. (b) We use only a binary indicator for having kept one appointment in a 6 month period rather than a count or continuous measure of appointments kept. A count or continuous measure might be more likely to reflect the degree of prior illness of the patient as the reason for an appointment, while the binary measure more should reflect the anticipated behavior of all asthma patients, who should be seen at least every 6 months. In this fashion we attempted to distinguish patients by their adequate versus inadequate level of ongoing care, and then to determine the association of that measure with asthma control.

Appendix A5. Association of treatment assignment (home visitor versus portal only) and adherence to recommended regimens for inhaler adherence.

Background

One predefined longitudinal measure of the effectiveness of the home visitor versus portal only intervention was inhaler adherence as estimated by the IAS scale (Dolce 1991, Table 1).   This validated scale rates adherence based on the following six items.

1. During the last 3 months, have you at times been careless about using your [insert name of inhaled steroid, e.g. Advair, Flovent, Pulmicort, Beclovent, Vanceril, Aerobid Azmacort, QVar]?

2. During the last 3 months, have you ever forgotten to use your [insert name of inhaled steroid]?

3. During the last 3 months, have you ever stopped using your [insert name of inhaled steroid] because you felt better?

4. During the last 3 months, have you ever used your [insert name of inhaled steroid] less than the doctor prescribed because you felt better?

5. During the last 3 months, have you ever stopped using your [insert name of inhaled steroid] because you felt worse?

6. During the last 3 months, have you ever used your [insert name of inhaled steroid] more than the doctor prescribed because you felt you were having breathing problems?

A no response to an item receives a score of 1, and a yes a score of 0, the items are summed, and the total score ranges from 0 (worst) to 6 (best) adherence.

The pre-specified hypothesis was that the intervention might operate on asthma control through improved adherence to inhaler use, i.e. that inhaler use was a potential mediator of the effect of the intervention.

Statistical methods

Using this 7-point scale as the outcome in a longitudinal model that we used for both asthma control and quality of life (time measured by splines and baseline covariates added to the model), we estimated the differences between intervention groups (home visitor versus

Appendix A5 (continued). Association of treatment assignment (home visitor versus portal only) and adherence to recommended regimens for inhaler adherence.

portal use only) in the change in expected values (means of adherence scores) from 0 months to 12 months. As with the other analyses, we estimated confidence bound using the percentile method of 999 bootstrap replications of resampling the data at the patient level and with replacement. We used a marginal model, as implemented by generalized estimating equations with a log link, a gamma family, and a independence working correlation matrix.

Reference

Dolce JJ, Crisp C, Manzella B, Richards JM, Hardin JM, Bailey WC. Medication adherence patterns in chronic obstructive pulmonary disease. Chest. 1991; 99(4):837-41. [PMID: 2009784]

Appendix A6. Effect modification and identification of subgroups.

Over several team meetings, candidate effect modifiers were discussed and a list was generated based on prior knowledge and clinical expertise. The candidate measures include self-reported: Spanish as primary language; specialty or primary care practice site; age group (18-39 years, 40-49 years, > 50 years); education attainment; comorbidities (diabetes, hypertension, high cholesterol, and cancer); hospitalizations; ED visits; intubation; years taking ICS; prednisone usage; food or clothing inadequacy; crowding at home; exposure to violence; lack of housing; lack of utilities; housing instability; smoking status; exposure to second hand smoke; numeracy measured with the ANQ (Apter et al., 2006) reading comprehension measured with S-TOFHLA (Baker et al., 1999) inhaler technique measured using a 7-point scale for a metered dose inhaler and a 6-point scale for a dry powder inhaler, testing the patient on the inhaler used for their inhaled steroid; social support assessed with the Medical Outcome Study Social Support Survey (Sherbourne et al., 1991) privacy concerns with the patient portal assessed using the Portal Use Baseline Survey; depression as measured by the Center for Epidemiologic Studies Depression Scale, a validated 20-item scale (Radloff, 1977) and obesity measured by body mass index. Electronic health literacy was measured with the eHEALS, the eHealth Literacy Scale, an 8-item measure of knowledge, comfort, and perceived skills at finding and evaluating electronic health information (Norman et al., 2006). It has been used in low-income patients (Knapp, 2011). Each item is measured with a 5-point likert scale and the

eHEALS score is the mean of each item. The Scale was administered at the baseline and final data collection. Computer literacy was measured using one question from this scale ("I feel confident in using information from the internet to make health decisions").

Since one-by-one testing or estimation of candidate effect modifiers has been heavily criticized, the team decided upon groupings of candidate effect variables according to themes (see main report, Table 3). For each theme or group of candidate measures, we implemented a latent class analysis with the goal of reducing the data and the number of candidates.

Estimating latent classes

Latent class analysis was implemented for each group of candidate measures which will distinguish profiles of common responses to the set of variables (McCutcheon, 1987). Latent class regression or finite mixture modeling is concerned with deriving information about a categorical latent variable from observed multivariate response patterns. The method takes advantage of the full data likelihood; therefore, subjects with missing elements contribute to the classification model unless all elements are missing.   Full information maximum likelihood, the underlying algorithm, functions with missing data with comparable effectiveness as multiple imputation. The amount of missing data in baseline covariates was small.

Data analysis was performed using Mplus version 7.4, which uses an efficient estimation-maximization algorithm for maximum likelihood estimation (Dempster et al., 1997). The number of latent classes was determined through examination of fit indices and in relation to clinical interpretation of results. Specifically, the Lo-Mendell-Rubin Adjusted Likelihood Ratio Test of model fit (Lo, Mendell, Rubin, 2001) offered a formal comparison of a model with its reduced form. Model estimation yields predicted probabilities of class membership for each individual, which was used to assign an individual to a class or group associated with the highest predicted probability of class membership. The indicator representing the class membership became the moderator variable. Then, the degree of effect modification of the association of the intervention (home visitors) and the key outcome (asthma control) was estimated for each candidate modifier group by testing interactions between the group indicator and the randomization indicator, and time. These preliminary analyses were based on Wald tests for 3-way interactions of time (fit with marginal splines), intervention (home visitor

versus patient portal only), and the candidate effect modifier.   Further investigations used separate regressions (of intervention*time interactions) for each level of the potential effect modifier, while controlling for the baseline covariates based on latent classes of factors (see description of latent classes). We used a common p-value for testing interactions of $p<0.1$. Our working hypothesis is that patients with comorbidities and with limited computer literacy could face additional challenges in the use of the patient portal. They also might need a more intensive (in frequency and in home education) home visitor program than the 4-visit program our intervention offered.

References

1.   Apter AJ, Cheng J, Small D, et al. Asthma Numeracy Skill and Health Literacy. J Asthma 2006;43:705-10.

2.   Baker DW, Williams MV, Parker RM, Gazmararian JA, Nurss J. Development of a brief test to measure functional health literacy. Patient education and counseling 1999;38:33-42.

3.   Sherbourne CD, Stewart AL. The MOS social support survey. Social science & medicine (1982) 1991;32:705-14.

4.   Radloff LS. The CES-D Scale: A self-report depression scale for research in the general population. Applied Psychological Measurement 1977;1:385-401.

5.   Norman CD, Skinner HA. eHEALS: The eHealth Literacy Scale. Journal of medical Internet research 2006;8:e27.

6.   Knapp C, Madden V, Wang H, Sloyer P, Shenkman E. Internet use and eHealth literacy of low-income parents whose children have special health care needs. Journal of medical Internet research 2011;13:e75.

7.   Dempster AP, Laird NM, Rubin DB. Maximum likelihood from incomplete data via the EM algorithm with discussion. J R Stat Soc B. 1977;39:1-38.

8.   Lo, Y., Mendell, N.R. & Rubin, D.B. (2001). Testing the number of components in a normal mixture. Biometrika, 88, 767-778.

9.   McCutcheon AL. Latent Class Analysis. Beverly Hill, CA: Sage University Press Series on Quantiative Applications in the Social Sciences; 1987.

Appendix A7. Methods for handling Irregular data collection interviews - Unified modeling approach.

Background

This appendix details our approach and rationale for handling irregular data collection interviews. Although not specifically detailed in our original PCORI proposal, the methods we outline below we pre-specified before considering analyses that incorporate any outcomes.

Typically reports of longitudinal studies use time as a categorical factor. When all patients are measured at the same time, this analytic approach is entirely appropriate. For such studies as ours, however, this approach was both not feasible and theoretically unsound. For that reason we modeled actual data collection times using a flexible approach – marginal splines.

As with many studies, and in particular as with studies that involve mostly disadvantaged patients who also have underlying comorbidities that limit their mobility, contact and meeting availability with research personnel for the preplanned data collection times, this difficulty presents challenges. The preplanned times were 0, 12, 24, 36, and 48 weeks from randomization. But patients could not be contacted and interviewed at these times due to the previously mentioned challenges. This reality led to measurements at times that differed from, but mostly followed, the preplanned times. With irregular data collection times, we therefore used the following method of analysis.

Methods for irregular data collection times

(1) For analysis, we used all data whenever obtained. No interviews were considered to be out of range. We used all interviews even for those patients who dropped out after baseline interviews but before the completion of the fifth and final interview.

(2) Unlike many studies, we modeled time as a continuous measure rather than as a categorical factor. We recorded and used for each patient the actual time (months from randomization) of the interview.

(3) With continuous time, we needed then to allow for non-linear trajectories of outcomes over time. We made no assumptions about linear trends in health improvement arising out of the

Appendix A7 (continued). Methods for handling Irregular data collection interviews - Unified modeling approach.

study intervention. To that end, we modeled time using flexible splines. This approach is well-described in the statistical literature and implemented in standard statistical software packages. (4) With measures occurring over a continuous time line, and with final measures happening sometime before and some after the pre-planned 12-month goal, we compared the two randomized groups of patients as of the same, pre-planned 12-month time compared to baseline (date of randomization). We accomplished this task by estimating from our models the expected values or mean outcomes as of 0 months and 12 months. This task is routine with typical longitudinal data models, whether those models implement mixed effects or population averaged approaches. In either case, we estimated expected values at times 0 months and at 12 months. The contrast of interest in estimating the effects of the intervention versus control becomes a difference between groups of the within-group differences over time from 0 months to 12 months. This approach used all available data, distinguished, for example, between a final measurement at 11 months versus another at 14 months, and permitted standard methods of covariate adjustment for covariate imbalance arising out of the loss to follow up at any time during the study.

We consider our approach to be superior to some alternatives and especially to ad hoc approaches. First, we do not agree with conventional approaches that simply ignore discrepancies between pre-planned and actual interview times. If over time an intervention is effective in reducing asthma symptoms, then a patient should have a lower (better) asthma control score if he or she is interviewed at month 7 or 8 rather than at month 6. Second, we do not feel that treating time as categorical factor and adding a covariate that indicates whether the interview was inside or outside of a pre-specified window leads to unbiased estimates. There is little reason for this approach when time can be analyzed as a continuous measure. Third, approaches that use a patient's final measurement, even if it occurs before 12 months (pre-specified end date), as the final, 12-month measurement is simply another version of last observation carried forward (LOCF) approaches, which have been thoroughly discredited in numerous analyses, reports, and editorials. A final measurement at month 8 in theory will

Appendix A7 (continued). Methods for handling Irregular data collection interviews - Unified modeling approach.

differ from the same patient's measurement at month 12. Fourth, although multiple imputation is a well-accepted option for missing outcomes when visit times are categorical, the approach is not practical when, as in our case, visit times can and do occur throughout the 12-month follow-up. Multiple imputation, as currently programmed, would necessitate that we impute all 12 months of data for each patient – a very large amount of missing observations to impute and thus a large fraction of monthly outcomes are missing by design.   Finally, examples in the statistical literature support this suggested approach (Howe 2016).

Model form

   The primary outcome is improvement in asthma control in both the intervention (PT + HV) and the control PT (Portal only) groups. The question of interest then is whether asthma control improved over time more in the PT + HV group compared to the control group, and did the control group improve over time. More formally, as the primary comparison of this primary outcome, the corresponding statistical estimand is the difference in expected values between the PT + HV group and the PT (control) group from baseline to 12 months after randomization date. To test the within group and across group effects of the interventions on outcomes we implemented mixed effects models (using restricted maximum likelihood and adaptive quadrature) and marginal models (using generalized estimating equations (GEE)) as a unified approach for the analysis of the longitudinal data with baseline randomization. For within-group outcomes, these models use all available data to estimate changes over time. Both the mixed effects and marginal models can be used for continuous, binary, and count outcomes using linear, logistic, and log linear models that can compare changes over time in the outcomes of asthma control, quality of life, ED visits, hospitalizations over one year. (Fitzmaurice, Laird, Ware 2011). By including group-by-time interaction terms, these models can then be adapted to compare improvement over time between the patient portal plus home visit group and the portal only group.

   Each type of model has its strengths and weaknesses. While the GEE method benefits from simplicity and robustness, it requires stronger assumptions about dropout (dropout

Appendix A7 (continued). Methods for handling Irregular data collection interviews - Unified modeling approach.

completely at random). Planned analyses considered the effects of both dropout and nonadherence (failure to complete the scheduled home visits), irregularity of the data collection and outcome ascertainment, as sensitivity analyses.   Details are below.

We selected the model form based on the particular form of the outcome of interest. For the primary outcome of asthma control, we found that the responses were heavily skewed. For that reason we implemented log gamma models. In each case, we include in the base model as a stratification factor, the clinical site at which the patient was treated. We did so because randomization was stratified by treating clinic. Using a log gamma model, these expected values and their differences can be estimated using mixed effects models that allow for correlations in longitudinal data.   For estimates of the differences between treatment groups of the change over time in asthma control, we will analyze the change in asthma control per our statistical analysis plan using mixed effects models with a random intercept and slope for patient, using an identity link models. If the observed distribution is skewed to the right, as we suspect a priori, then a log gamma generalized linear mixed effects model will be indicated and used. Time is modeled flexibly with splines.

Sensitivity analysis to irregular data collection times

Methods for sensitivity analysis for irregular outcome assessment times are under development (Lin; Pullenayegum, Van Ness, and Buzkova 2009, 2010).   In brief, recent research on sensitivity of estimate to informative dropout and related software assume regular outcome measurement times in longitudinal studies (Scharfstein 2018; 2017).   In our case, the issue is informative measurement times—an association of the timing of the measurement and the value of the measurement.   Current software does not yet meet the needs of a proper sensitivity analysis for possibly informative, irregular measurements.   It does cover regular follow up intervals.   Sensitivity analysis along the lines of the proposal and implementation of Scharfstein and colleagues remains for development.

References

1.   Buzkova P, Brown ER, John-Stewart GC. Longitudinal data analysis for generalized linear models under participant-driven informative follow-up: an application in maternal health epidemiology. American journal of epidemiology 2010;171:189-97.

2.   Buzkova P, Lumley T. Semiparametric modeling of repeated measurements under outcome-dependent follow-up. Statistics in medicine 2009;28:987-1003.

3.   Fitzmaurice GM, Laird NM, J.H. W. Applied Longitudinal Analysis, 2nd Edition. Hoboken, NJ: John Wiley & Sons; 2011.

4.   Howe LD, Tilling K, Matijasevich A, et al. Linear spline multilevel models for summarising childhood growth trajectories: A guide to their application using examples from five birth cohorts. Statistical methods in medical research 2016;25:1854-74.

5.   Lin H, Scharfstein DO, Rosenheck RA. Analysis of longitudinal data with irregular, outcome-dependent follow-up. . Journal of the Royal Statistical Society: Series B (Statistical Methodology) 2004;66:791-813.

6.   Pullenayegum EM, Lim LS. Longitudinal data subject to irregular observation: A review of methods with a focus on visit processes, assumptions, and study design. Statistical methods in medical research 2016;25:2992-3014.

7.   Scharfstein, D. , McDermott, A. , Díaz, I. , Carone, M. , Lunardon, N. and Turkoz, I. (2018), Global sensitivity analysis for repeated measures studies with informative drop-out: A semi-parametric approach. Biometrics.   2018; 74: 207-219. doi:10.1111/biom.12729

8.   Scharfstein D, McDermott A. samon. Sensitivity analysis for missing data.   2017. https://cran.r-project.org/package=samon

9.   Van Ness PH, Allore HG, Fried TR, Lin H. Inverse intensity weighting in generalized linear models as an option for analyzing longitudinal data with triggered observations. American journal of epidemiology 2010;171:105-12.

Appendix A8. Fitting mixed effects models to estimate individual improvement in asthma control and achievement of asthma control.

Background

The overall estimates of the average improvement in asthma control over time within the two intervention groups and the differences in the two improvements (see main text) do not reflect whether all patients improved, some improved and others did not, or some improvement markedly and others worsened.   To arrive at estimates of the number of patients who might have achieved asthma control over time and/or who might have improved individual levels of control over time, we had to implement model strategies that estimate individual levels of control.   We regard this exercise as a sensitivity analysis to our main results (see main report), or alternatively as a re-expression of those results in alternative terms that might have more clinical meaning.

For two reasons, we could not *and should not* use raw data to report individual asthma control.   First, as we report in the main text, patients were measured at times that departed from the pre-specified times. Most of these departures reflected delays in obtaining interviews.   Reasons for the delays we have already outlined: patients were difficult to locate and schedule, in part because of their many illnesses and their often poor living conditions and communication options.   For that reason, actual 12month measures of outcomes, e.g., asthma control, were routinely not observed. Rather, observed (measured) asthma control occurred at various times around 12 months from the date of randomization for the individual. Variation in times of the final measurement were substantial. We therefore had to develop a method of estimating measures of asthma control as of a common date – 12 months from randomization – to account for this variation in the timing of data collection.   Second, as with many health measures based on patient responses to an interview instrument, asthma control is measured with error. The observed asthma control is but one estimate of the patient's asthma symptoms on a particular day out of several days that represent actual, steady-state symptoms.   For both of these reasons, an observed measure will not accurately reflect the actual level of control.

Methods

Appendix A8 (continued). Fitting mixed effects models to estimate individual improvement in asthma control and achievement of asthma control.

Mixed effects models that include random effects to represent individual patient departures from average levels of outcomes, such as asthma control, have long been used to address these issues of measurement error.   The resulting "predictions" of individual levels of outcomes are in theory better, less biased, estimates of individual-level outcomes that are the raw observations. (Efron; Morris; Casella). In this application, we implemented longitudinal models with two different types of random effects.   Random intercepts reflect the individual departure of outcomes (asthma control) at baseline, in our case at randomization (month 0). Random slopes reflect the departure of individual level trajectories from the average trajectory of change in outcomes over time.   With these two types of random effects, we could predict each patient's level of outcome at two common times: the date of randomization and 12 months later.   We used these predictions, and their differences, of asthma control to estimate the number of patients who were in control (score <1.5) and who had achieved a clinically important reduction in score (decrease of 0.5), which reflects an improvement in asthma control.

We used restricted maximum likelihood (REML) as implemented in the program "mixed" in Stata version 15.1 to fit a linear mixed model.   Our attempts to fit generalized linear mixed models, such as a log gamma model with log link and gamma error, all failed to achieve convergence. Thus, this linear model did not fully reflect the skewed nature of the data but represented the best alternative.

In our application, defining and implementing random slopes became more complex because of our use of splines to estimate the changes over time in outcomes (See main text for description of the primary analysis).   The time line of observation as divided into segments, with each segment allowed to have a different slope (or trajectory) of the outcome over time. We began with a model that had the same set of 5 line segments for time as did the primary model.   We then added all of the baseline covariates (as in the primary analysis) in order to explain as much as possible the individual patient-level variation in outcome.   To this model, we added a random intercept for each patient and a single random slope to reflect overall

Appendix A8 (continued). Fitting mixed effects models to estimate individual improvement in asthma control and achievement of asthma control.

departure of the individual over time from the average trajectory not explained by these observed covariates.   The model converged. We then added a second random slope to represent an additional element of individual patient departure from average as the follow-up time progressed.   This model converged in 10 iterations.   Models with additional random slopes to reflect the additional time segments of the spline model would not converge.

The final model, included as fixed effects: Linear (marginal) spline with knots at 3,6, 9, and 12 months; clinic, Spanish as the primary language, age categorized (18-39, 40-49, and 50+), literacy levels (4 levels from latent class models), social support(3 levels from latent class models), smoking exposure(4 –levels: none, first hand smoke only, second hand smoke only, both first and second hand smoke), comorbidities (3-level latent class model: healthy, depressed, other chronic comorbidities), baseline asthma severity (low high), concern about using health data on the internet (5 levels), home crowding (3 levels: <1 per room, 1 up to 2 per room, 2+ per room).

Random effects included: a random intercept for subject and a random slope for the first two line segments constructed by the spline model.

This analysis can best be thought of as a sensitivity analysis to the main findings of the asthma control outcome.   Its limitations are that there were insufficient data to fit a log-link, gamma-family model with more than 2 random slopes.

References

1.   Efron B, Morris C. Data analysis using Stein's estimator and its generalizations. J. Amer. Statist. Assoc. 1975;70:311–319.

2.   Morris CN. Parametric empirical Bayes inference: Theory and applications. J. Amer. Statist. Assoc. 1983;78:47–65. with discussion.

3.   Casella G. An introduction to empirical Bayes data analysis. Amer. Statist. 1985;39:83–87.

Appendix A9: Sensitivity analysis for dropout—Using baseline covariates.

Two missing data problems might affect our results: (1) missing outcomes over time arising from dropout, and (2) missing baseline covariate values. This appendix does not cover missing covariate values. Details on how we used latent class models to revolve the small amount of data on baseline factor appears in the appendix on effect modification.   This appendix uses the latent classes and baseline factors as a sensitivity analysis of the effect of dropout on randomization. In addition, in this appendix we do not discuss the potential for dropout (or more appropriately delayed acquisition of outcome data) that might be not at random.   That subject we cover in Appendix A10, and a formal solution awaits methodological advances for irregular visit data.

We used marginal models adjusted for baseline covariates as combined using latent class models.   (See details in Appendix on effect modification).

Appendix A10. Estimating effect of home visitors adjusted for incomplete adherence – A sensitivity analysis.

Methods

With less than full adherence to the scheduled number of home visits for all patients assigned to home visitors, the "as randomized" analysis does not answer the following question: What is the effect of home visitors among patients who would be responsive to scheduling for and meeting with a home visitor? The estimand to answer this question is sometimes called the "complier average causal effect" or CACE. Because actual adherence to the home visit protocol is not randomized, i.e., because patients receive or do not receive the full complement of home visits based on factors other than randomization, simplistic approaches can be subject to selection bias, from both observed and unobserved confounders. Those factors might influence both the patient's decision to meet the home visitor and the patient's observed outcome. In other words, the observed change in asthma control, for example, among those patients who completed all scheduled home visits does not account for the fact that this subgroup of patients has not been randomly assigned to full home visits versus no home visits. Nor does the simple approach consider that those who accept the home visits might also have been more likely to improve if they had been assigned to the patient portal arm alone. Whether the patient received the home visit might be related to the patient's outcome. (Sagarin 2014) To avoid the bias from less than complete adherence, special statistical methods are needed.(Stuart 2008).

One approach might be the use of models that weight data based on the probability of a patient's being adherent to the treatment at the specified discrete time that is the same for all patients. This approach could apply for example, in a study in which all patients are seen at the same time since randomization or when data collection is at set times (Toh 2010). These weighting approaches attempt to address the bias arising out of non-randomized adherence through modeling adherence based on baseline and post-baseline covariates. While this approach might apply in the case of regular data collection times, it is not feasible for studies, such as ours, in which data collection times are irregular. In our case, irregularity is the rule

Appendix A10 (continued). Estimating effect of home visitors adjusted for incomplete adherence – A sensitivity analysis.

rather than the exception. We elected an alternative approach to resolve this estimation problem.

Instrumental variable approach to adherence

To deal with this estimation problem in randomized designs, we adopted an instrumental variable approach, in which randomization serves as the instrument, and the study design reflects what can be called a "randomized encouragement design". By this approach, randomization to the home visitor group becomes an encouragement to the patient to respond to the staff's request to schedule a home visit, and then to meet with the visitor as scheduled. (Ten Have 2004). In our case, "adherence" is much more complex than just accepting a treatment. Rather, adherence is a combination of being able to remain in contact with the study staff, having a schedule that permits setting a home visit date, and allowing the home visitor to come to the patient's home and offer information.

The instrumental variable approach can support estimation of the effect of the home visits among those who accept them and receive them, and will control for confounding of unmeasured factors, at the expense of increased variance in the resulting estimates of the intervention effects. To be valid and unbiased, the instrumental variable analysis must satisfy several assumptions. (Baiocchi 2014; Stuart 2008). We realize that different authors sometimes describe these assumptions somewhat differently.

(1) The stable unit value treatment assumption requires two sub-assumptions.   (1a) The assumption of "no interference" -- that the outcome of one patient is not influenced by the treatment assignment of another patient.   This assumption was likely satisfied because home visits occurred patient by patient at individual homes rather than in groups. (1b) There is only one version of the home visitor intervention. The home visitors were trained, and their interviews were scripted with the help of an interview packet.   But as with any similar intervention, our study could not guarantee that the eight home visitors presented themselves consistently and according to the protocol, but that variation in home-visitor personalities gave

Appendix A10 (continued). Estimating effect of home visitors adjusted for incomplete adherence – A sensitivity analysis.

rise to somewhat different flavors of the intervention.   For that reason, part of the SUTVA assumption might not be satisfied.

(2) "Exclusion restriction". The effect of the assignment to the home visitor arm of the study must influence outcome only through the actual completion of the home visit.   A patient would have to react the same in terms of asthma outcome if he or she had been assigned to the home visitor treatment arm or the portal only arm, and had received no home visits.   In the context of this study, this key assumption would be violated if the mere assignment to the home visitor led to higher expectations of improved asthma outcome just by learning of the assignment, and diminished expectations upon learning of assignment to the patient portal arm.   Owing to the unavoidable fact that the patients could not be blinded to the treatment assignment, a violation of this assumption might be possible. But it might be that both the home-visitor patients and the portal-only patients could have similar expectations of some improvement merely by being recruited to the study and followed over time.   A variation on this assumption requires that the patient who is randomized to the home visitor group but who does not schedule or see the home visitor (lack of compliance with the protocol) will receive no benefit from just being assigned to the home visit group.   That assumption seems plausible in this setting.

(3) The final assumption, "monotonicity", requires that patients assigned to the portal group did not obtain home visits.   Our design satisfied this assumption because only patients assigned to the home visit arm of the study were able to receive home visits. Although this assumption was very likely satisfied, meeting this assumption limits the inference of the effect of the home visitors to the patients with the characteristics of those who actually had the full set of home visits.   To be able to generalize the effect of home visits to all patients would require yet an additional assumption – that the effect of treatment is the same across all patient types.   We feel that this assumption could not be satisfied because home visits, almost by definition, are not going to produce the same reaction and outcome across a diverse group of patient who

Appendix A10 (continued). Estimating effect of home visitors adjusted for incomplete adherence – A sensitivity analysis.

have sometimes markedly different comorbidities, ages, educations, poverty levels, and home environment.

Additional assumptions implied by this approach are (4) random assignment and (5) that at least some patients complied with the protocol and met with the home visitors having been randomized to that arm. Both implied assumptions were satisfied in our study.

Based on the assumptions, we determined that the instrumental variable approach might help to estimate the effect of home visits among those who complied with the home visit protocol.

Measuring adherence

To implement the instrumental variable approach, we measured adherence as follows. The data collection times at the protocol were 0 12 24 36 and 48 weeks. The scheduled home visit times were baseline, 3, 5, 8, and 25 week. But many patients could not or did not follow either plan. For reasons of difficulty in scheduling both the home visit and the data collection, patients' actual schedules slipped, sometimes by more than a few months, from the schedule specified by the protocol.

By comparing the dates of the actual home visits with the dates of the actual outcome data collection, we could determine for each data collection the number of home visits that the patient should have completed as of the data collection date and compare that number to the number of home visits actually received at that time. For example, if the patient had a first post-baseline data collection as of day 90 and by that time the patient should have received three home visits, the expected number of visits = 3. If the patient had received only 2 home visits as of that time, then the ratio of actual to expected number of visits = 2/3= 0.67. We then concluded that falling behind in the home visit schedule could attenuate the effect of the home visitor, and for that reason, we categorized each patient at each visit as adherent if this ratio was greater than or equal to 1.0 and non-adherent otherwise.

Estimating the effect of home visits under complete adherence

Appendix A10 (continued). Estimating effect of home visitors adjusted for incomplete adherence – A sensitivity analysis.

To estimate the effect of the home visitors among those who would adhere to the visit schedule, we implemented the two-stage residual inclusion (2SRI) approach to instrumental variable analysis in longitudinal settings. (Small 2006). By this approach, in the first stage we modeled the probability of a patient in the home visit group being adherent at a given time as a function of baseline covariates. At this stage we also computed the predicted probabilities of adherence based on this model for each patient, and then the residuals (Wij= actual adherence – predicted adherence, for patient=i at time=j), where the residuals =0 for patients assigned to receive portal training only.

In the second stage, we then fit a random effects log linear model (to account for skewed distribution of asthma control as an outcome) of asthma control as a function of: time (T), baseline covariates (X), the residuals (Wit), and the indicator for being adherence at a given time (Ait). This model we fit for all 301 patients. To this model we then added random effects for patient, in the form of a random intercept for each patient and a random slope for time. Because of the added complexity of irregular data collection times, we fit the same spline model for time as with the primary outcome model for asthma control (without an adjustment for less than complete adherence). The coefficient for the term for Aij represents the estimate of the effect of the home visitor at time=j.

For implementing the analysis of the effect of home visits among the compliers, we modified the two-stage residual inclusion (2SRI) approach as follows to arrive at a similar difference of expected values as we estimated in the "as randomized" analysis.

Step #1: Estimating the residuals

The initial model is a longitudinal logit model with Ait as the outcome (adherence as of the data collection at time = t). Then

$E(Aij) = alpha0 + Xi*ALPHA + BETA*Tij$, where X is a vector of baseline covariates, and Tij is a vector of elements of a spline for patient=i at time=j.

$Wij = Aij - E(Aij)$ represents the residuals from this first stage equation.

Step #2: Estimating the response model

Appendix A10 (continued). Estimating effect of home visitors adjusted for incomplete adherence – A sensitivity analysis.

Using a log link model (for asthma control), the response model was:

Log(Yij)=Xi*DELTA1 + GAMMA*Wij + BETA*Tij + *PSI*Aij, + Z*b, where this model is log gamma to account for skewed data, Xi is a vector of baseline covariates including an intercept, Wij is the vector of residuals from step 1, Tij is a vector of times in spline format, and Aij is the vector of adherence over time.   Z*b represents random effects at the patient level to allow for variation of individual patients from the average.   For the patients in the portal only group, Wij=0 and Aij=0 as these patients did not have access to the home visitor.   (The upper-case represents a vector of coefficients corresponding to the vector of factors.   For example, GAMMA is a vector of coefficients for the residual (from Step #1) at each data collection time. PSI then represents the vector of coefficients over time).

Our interest was not in the estimates for coefficients PSI but in the expected value of outcome (asthma control, for example, as of two time points, month=0 and month=12). For that reason, we estimated the solution to the question of the effect of the home visitor among those who complied by augmenting the original dataset (n=301) with four additional datasets and then using the original dataset to estimate predictions for each patient and then applying the predictions to the augmented datasets.   The predicted outcome accounted for each patient's baseline characteristic as well as the patient's random variation from the average. The four augmented datasets had the following characteristics.

(1) For the patients who complied/adhered, assume that they did not adhere at all (Aij=0) and that time = month 0, i.e., that the vector T of spline values for time are set to the equivalent of month =0.

(2) For the patients who adhered, assume that they did not adhere at all (Aij=0 and that time = 12 months.

(3) For the patients who adhered, assume that they did adhere throughout (Aij=1) and that time= 0 months.

(4) For the patients who adhered, assume that they did adhere throughout (Aij=1) and that time = 12 months.

Appendix A10 (continued). Estimating effect of home visitors adjusted for incomplete adherence – A sensitivity analysis.

These augmented datasets were confined to the patients who adhered, because we could not justify the assumption that the effect of home visitors would be the same across all patients. By estimating predictions for the each of the patients in these four datasets, we were able to estimate individual patient outcomes assuming two different adherence regimes and at two times. Thus, each patient had four predicted values.

The last step was to calculate the mean over the patients in each of the four augmented datasets. These four means then represented the expected values of the outcome among the patients who adhered at two time points (month=0 and month=12), assuming that they did adhere and alternatively that they did not adhere at all.   The estimator of interest was then: $[E(Y_{0,12}) - E(Y_{0,0})] - [E(Y_{1,12}) - E(Y_{1,0})]$, where $E(Y_{A,B})$ is the expected value (mean), A=0 represents the assumption of no adherence, A=1 represents the assumption of full adherence, and B represents times 0 and 12 months.

Confidence intervals

We estimated 95% confidence bounds for the resulting estimate by bootstrap resampling using 999 iterations and percentile bounds, with replacement stratified by treatment arm (home visitor vs portal training only).

Implementation of the two-stage residual inclusion approach

Both for the initial stage of this method (estimating residuals) and for the response model, we used same set of covariates as with the primary asthma control model (with baseline covariate adjustment) for the effect modification analyses.

References

1.   Baiocchi M, Cheng J, Small DS. Instrumental variable methods for causal inference. Stat Med. 2014; 33(13):2297-340. [PMID: 24599889]

2.   Nagelkerke N, Fidler V, Bernsen R, Borgdorff M. Estimating treatment effects in randomized clinical trials in the presence of non-compliance. Stat Med. 2000; 19(14):1849-64. [PMID: 10867675]

3.   Sagarin BJ, West SG, Ratnikov A, Homan WK, Ritchie TD, Hansen EJ. Treatment noncompliance in randomized experiments: statistical approaches and design issues. Psychol Methods. 2014; 19(3):317-33. [PMID: 24773358]

4.   Small DS, Ten Have TR, Joffe MM, Cheng J. Random effects logistic models for analysing efficacy of a longitudinal randomized treatment with non-adherence. Stat Med. 2006; 25(12):1981-2007. [PMID: 16220487]

5.   Stuart EA, Perry DF, Le H, Ialongo NS. Estimating Intervention Effects of Prevention Programs: Accounting for Noncompliance. Prevention Science. 2008; 9(4):288-98.

6.   Ten Have T,R., Elliott MR, Joffe M, Zanutto E, Datto C. Causal Models for Randomized Physician Encouragement Trials in Treating Primary Care Depression. Journal of the American Statistical Association. 2004; 99(465):16-25.

7.   Ten Have TR, Normand SL, Marcus SM, Brown CH, Lavori P, Duan N. Intent-to-Treat vs. Non-Intent-to-Treat Analyses under Treatment Non-Adherence in Mental Health Randomized Trials. Psychiatr Ann. 2008; 38(12):772-83. [PMID: 20717484]

8.   Toh S, Hernandez-Diaz S, Logan R, Robins JM, Hernan MA. Estimating absolute risks in the presence of nonadherence: an application to a follow-up study with baseline randomization. Epidemiology. 2010; 21(4):528-39. [PMID: 20526200]