Published in final edited form as:

Med Decis Making. 2021 January ; 41(1): 9–20. doi:10.1177/0272989X20973160.

USING CLINICAL TRIAL DATA TO ESTIMATE THE COSTS OF BEHAVIORAL INTERVENTIONS FOR POTENTIAL ADOPTERS: A GUIDE FOR TRIALISTS

Louise B Russell, PhD^{1,2}, Laurie A. Norton, MA¹, David Pagnotti, MS¹, Christianne Sevinc, MPH³, Sophia Anderson, BA¹, Darra Finnerty Bigelow, MPH¹, Lauren G. Iannotte, BA¹, Michael Josephs, BA¹, Ryan McGilloway, BA¹, Iwan Barankay, PhD⁴, Mary E Putt, PhD ScD⁵, Peter P Reese, MD MSCE^{2,5,6}, David A Asch, MD MBA^{1,2,7,8}, Lee R Goldberg, MD MPH^{1,7}, Shivan J Mehta, MD MBA MSHP^{1,2,3,7}, Monique S Tanna, MD⁷, Andrea B Troxel, ScD⁹, Kevin G Volpp, MD, PhD^{1,2,3,7,8}

¹Department of Medical Ethics and Health Policy, Perelman School of Medicine, and Center for Health Incentives and Behavioral Economics, University of Pennsylvania, Philadelphia, PA, USA.

²The Leonard Davis Institute of Health Economics (LDI), University of Pennsylvania, Philadelphia, PA, USA

³Penn Medicine Center for Health Care Innovation, University of Pennsylvania, Philadelphia, PA, USA.

⁴Department of Management and Department of Business Economics and Public Policy, The Wharton School, University of Pennsylvania, Philadelphia, PA, USA.

Corresponding author: Louise B. Russell, louisebrussell1983@gmail.com.

Authors no longer at Penn. Darra Finnerty Bigelow is Clinical Trials Manager at Teleflex, Inc., Ryan McGilloway is with PayPal's Business Financing Solutions team, Sophia Anderson is in the Graduate Biostatistics Program at Duke University, Durham NC, and Michael Josephs is in the graduate program for physicians' assistants at Thomas Jefferson University, Philadelphia PA.

Author contributions

Louise Russell designed the study; directed the development of time logs, computer platform estimates; drafted and revised the manuscript.

David Pagnotti, Ryan McGilloway, and Darra Finnerty Bigelow participated in developing the Habit Formation/Process v Outcomes time log and filled it out for the 11 days selected.

Darra Finnerty Bigelow recorded the time for creating lists of potential participants and mailing invitations.

Christianne Sevinc developed the cost estimates for the computer technology after conversations with Darra Finnerty Bigelow, Ryan McGilloway, and Louise Russell.

Laurie Norton, Lauren Iannotte, Sophia Anderson, and Michael Josephs participated in developing the EMPOWER time log and filled it out for four selected weeks.

Iwan Barankay, Mary Putt, Peter Reese, and Kevin Volpp, investigators on the Habit Formation and Process v Outcomes trials, participated in discussions of the costing of the interventions tested in those trials.

David Asch, Lee Goldberg, Shivan Mehta, Monique Tanna, Andrea Troxel, and Kevin Volpp, investigators on the EMPOWER trial, participated in discussions of the costing and cost-effectiveness analysis planned for that trial. Andrea Troxel advised on selecting the weeks for completing the EMPOWER time log.

All authors read the manuscript, provided critical feedback, and approved the final version.

Conflicts of interest. Drs Asch and Volpp are partners at VAL Health, and Dr Troxel is on VAL Health's Scientific Advisory Board. Dr. Volpp has served as a consultant for CVS Caremark and received research funding from CVS, Weight Watchers, Humana, Vitality/Discovery, Oscar, and Hawaii Medical Services Association. The other authors report no conflicts.

Trial registration numbers at ClinicalTrials.gov

EMPOWER: [NCT02708654](https://ClinicalTrials.gov/record/NCT02708654).

Habit Formation: [NCT01798784](https://ClinicalTrials.gov/record/NCT01798784).

Process v Outcomes: [NCT02246959](https://ClinicalTrials.gov/record/NCT02246959)

- ⁵Department of Biostatistics, Epidemiology & Informatics, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA, USA.
- ⁶Renal Division, Department of Medicine, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA, USA.
- ⁷Department of Medicine, University of Pennsylvania Perelman School of Medicine, Philadelphia, PA, USA.
- ⁸The Corporal Michael J Crescenz VA Medical Center, Philadelphia, PA, USA.
- ⁹Division of Biostatistics, Department of Population Health, New York University School of Medicine, New York, NY, USA.

Abstract

Behavioral interventions involving electronic devices, financial incentives, gamification, and specially trained staff to encourage healthy behaviors are becoming increasingly prevalent and important in health innovation and improvement efforts. While considerations of cost are key to their wider adoption, cost information is lacking because the resources required cannot be costed using standard administrative billing data. Pragmatic clinical trials that test behavioral interventions are potentially the best and often only source of cost information, but rarely incorporate costing studies.

This paper provides a guide for researchers to help them collect and analyze, during the trial and with little additional effort, the information needed to inform potential adopters of the costs of adopting a behavioral intervention. A key challenge in using trial data is the separation of implementation costs, the costs an adopter would incur, from research costs. Based on experience with three randomized clinical trials of behavioral interventions, this paper explains how to frame the costing problem, including how to think about costs associated with the control group, and describes methods for collecting data on individual costs; specifications for costing a technology platform that supports the specialized functions required; how to set up a time log to collect data on the time staff spend on implementation; and issues in getting data on device, overhead, and financial incentive costs.

Keywords

Behavioral intervention; clinical trial; costing; medication adherence; outcomes research; pragmatic trial; preventive services; translating research into practice

Behavioral interventions involving the use of electronic devices, financial incentives, gamification, and specially trained staff to encourage healthy behaviors are becoming increasingly prevalent and important in health innovation and improvement efforts. Systematic reviews have found them to be effective for some behaviors in some populations.¹⁻³ Potential adopters of these interventions – clinical practices, health plans, employers, and hospitals – need information not only about their effectiveness, but also about the costs of implementing them.

Behavioral interventions differ from many interventions evaluated in costing, cost-effectiveness, and budget impact analyses. Interventions such as screening for cancer, or surgery for heart disease, make use of established clinical services and tests and can be costed using sources such as the Healthcare Cost and Utilization Project (HCUP)⁴ and Medicare payment rates.⁵ In contrast, behavioral interventions often involve specialized staff and technology that are not part of standard clinical workflows or billing processes. As a result, the costs of the resources required, which can include electronic devices, integrated technology platforms, secure messaging, and staff to advise patients on their use, are not readily available from HCUP, Medicare, or health insurers' administrative data systems.

Pragmatic clinical trials that test behavioral interventions are a potential source for this cost information. A key challenge in using trial data, however, is the separation of implementation costs from research costs; research costs are often substantial, sometimes dwarfing the costs of the intervention itself, and must be separated out to give potential adopters an accurate idea of the costs they are likely to incur when using the intervention outside of a research setting. Additionally, because some of the resources required, such as secure messaging systems, are nonstandard, methods for costing them may be underdeveloped.⁶⁻⁷

This paper describes how to estimate a potential adopter's costs from the trial costs of a behavioral intervention. Although the fundamental costing principles are the same as for other interventions, behavioral interventions involve nonstandard costs, the need for more primary data collection than is typically the case, and careful work to separate research costs from costs that would be incurred by an organization that adopts the intervention. We use three randomized controlled trials (RCTs) as case studies to illustrate how behavioral interventions can be costed, one straightforward two-arm trial and two trials, each with several interventions, which were run by the same staff and so presented more complex costing issues.

THE TRIALS.

EMPOWER

The EMPOWER trial, which ended April 2020, enrolled 566 patients within 30 days of hospital discharge with a heart failure diagnosis in a pragmatic, randomized clinical trial to compare usual care with an "automated hovering" intervention that included patient incentives for daily weight monitoring and diuretic adherence. Electronic scales and pill bottles issued to each participant sent automated feedback into the clinical care pathway, enabling real-time response to concerning clinical symptoms.⁸ The intervention lasted 12 months. The primary trial outcome was time to hospital readmission with the primary aim to reduce readmissions. EMPOWER was a 2-arm trial with staff dedicated to that trial only.

Habit Formation and Process versus Outcomes

Habit Formation (HF) and Process versus Outcomes (PvO), completed in 2018 and 2019, respectively, were separate randomized controlled trials, each aimed at using financial incentives to improve adherence to statin medication. The goal was to reduce low-density

lipoprotein cholesterol (LDL-c), the primary outcome, in patients with elevated LDL-c. Each trial tested different financial incentive interventions against control. Because of their similar patient populations, goals, and resource requirements, the same staff worked on both trials.

Habit Formation (HF).—Habit Formation enrolled 805 patients in a six-month trial that offered financial incentives for statin adherence, monitored daily by means of an electronic pill bottle provided to each participant.⁹ Each day the pill bottle was opened, a signal was transmitted to the study team, and the participant was deemed adherent. Participants in all arms, including control, received daily reminders to take their statin. The intervention arms tested three alternative incentive designs to improve adherence. In all three, participants were eligible to win a daily sweepstakes, based on a personal 2-digit number selected at enrollment, if the subject was adherent the previous day. Arm 1 was a simple sweepstakes. Arm 2 (“Deadline Sweepstakes”) paid participants their full winnings only if they took their statin before receiving the daily reminder, half if they took it after the reminder. Arm 3 (“Sweepstakes plus Deposit”) also halved the daily winnings, compared to the simple sweepstakes, and put the other half in a monthly deposit, which was reduced each day the participant failed to take statins; the balance was paid at the end of the month. Expected winnings for a fully adherent participant averaged \$2.80 per day in all three interventions. LDL-c levels were measured at months 6 and 12.

Process versus Outcomes (PvO).—Process versus Outcomes enrolled 764 patients in a twelve-month trial of financial incentives. The Process arm rewarded statin adherence by making participants eligible for sweepstakes-based incentives each day they were adherent (as indicated by opening an electronic pill bottle). The Outcome arm paid participants if lab tests every 3 months showed that they had lowered their LDL-c by >10 mg/dl since the last test or remained stable below 100 mg/dL. The Process+Outcome arm combined daily sweepstakes incentives for adherence with 3-month incentives based on lab tests that confirmed declines in LDL-c or stability below 100 mg/dL. As in the HF trial, adherence in all arms, including control, was monitored through electronic pill bottles and the expected average payment for a fully adherent participant was the same in each arm.

THE COSTING PROBLEM

What costs would potential adopters incur for one of these interventions and what do they need to know about those costs? To estimate the costs of an intervention that is not just a rearrangement of well-established tests and procedures, it is necessary to identify in detail the tasks involved and to cost the resources required to carry out those tasks. This process is known as *micro-costing* because it “starts with the detailed inventory and measurement of resources”.¹⁰ By contrast, *gross costing* focuses on “units of input and output that are large relative to the intervention being analyzed,” such as complete hospital stays.

Since the costs of a trial include both intervention and research, the first step is to identify all tasks and resources involved, then the subset required by the intervention. Trial staff and investigators, the people most knowledgeable about the trial’s day-to-day activities, should create a detailed list of tasks and resources, including any that do not require payment; for example, if volunteers help with the intervention, their time must be included to give an

accurate picture of resource requirements. Breaking tasks down in detail helps ensure that all tasks are identified, and that implementation tasks can be distinguished from research tasks. It also helps potential adopters understand why the resources are needed.

Table 1, based on the HF/PvO trials, lists all the tasks involved in those two trials in the first column, the subset of tasks required to implement the interventions in the second column, and the types of resources needed for each task in the third column. Thus it shows which tasks need to be included, and which excluded, to measure only those costs involved in the intervention.

Research and other tasks unrelated to the intervention, e.g., departmental staff meetings, will clearly be excluded, but deciding how to handle some research components can be complicated. An important example is the resources required by the control group, the group of people who are enrolled in and tracked by the trial, but who do not receive the intervention. If the control group is not contacted and does not receive anything from the trial after enrollment, there are no intervention costs for them (other than enrollment). More typically, the control group is contacted and surveyed during the trial in order to facilitate comparison with the intervention groups. In the HF/PvO trials, for example, the primary outcome, change in LDL-c, was measured by lab test, and the secondary outcome, adherence to statins, was measured by the frequency with which participants opened an electronic pill bottle. To determine the effects of the incentive payments, control as well as intervention participants had their LDL-c measured and received electronic pill bottles, as well as staff help in the use of the pill bottle and computer platform. Thus the trial incurred some of the same costs for control as for intervention participants, costs that would not be involved in implementing an intervention in a clinical practice or workplace setting. Removing those costs would require separating out and costing contacts with control subjects, a time-consuming and costly process. Instead, calculating costs *per trial participant*, which adjusts for the total number of people served, provides an approximate and inexpensive correction. The correction will be most accurate if control and intervention participants use resources similarly, e.g., with similar levels of staff assistance.

Once the subset of tasks involved in implementing the intervention(s) is identified, the quantities of resources required will be measured in natural units, such as types of staff and hours of time spent by each type. The following sections discuss how to do this for the resources required by the three example trials. The quantities of resources must then be valued in dollar terms; for example, staff time will be valued by wages plus fringe benefits. At this final stage, all component costs need to be adjusted to the price level of a specific year so that they can be summed. The Consumer Price Index can be used to adjust costs to the selected year.

To recap, the goal of the process is to estimate the total costs of the intervention, exclusive of research and other costs, and to present those costs in ways that are most useful to potential adopters. Expressing all costs in dollars of the same year allows potential adopters to adjust them to the price level of the year they expect to run the intervention. Expressing costs per participant lets them estimate costs for sites with more or fewer enrollees. Implementing an intervention also involves the initial costs needed for any new effort, such as meetings to

discuss and plan the implementation. These efforts involve additional costs, but the research costs of a clinical trial are not a useful guide for estimating them.

MICRO-COSTING EACH RESOURCE

To help guide the costing process Figure 1 provides an overview of the steps and methods involved in costing a behavioral intervention using trial data. The following sections describe the methods in detail.

Technology platform

Behavioral intervention trials conducted at the University of Pennsylvania use *Way to Health*, one example of a specially-designed technology platform that automates tasks required for interventions that involve remote monitoring. *Way to Health* supports implementation through automated communication with trial participants; data collection from Bluetooth-enabled devices such as pill bottles, scales, and blood pressure cuffs; application of complex rules that evaluate data from those devices to tailor communications, monitor and adjust patient care; and automated payments of financial incentives. To serve the research needs of trials, *Way to Health* includes additional capabilities, such as the ability to administer and summarize surveys at enrollment and during the trial, to randomize participants to trial arms, and to store and organize data on each trial participant securely for subsequent analysis.

Adopting organizations could develop their own technology platforms, adapt one they might already have developed, or purchase a subscription to a commercially available platform. Subscription prices may be per user, per feature or module, a flat rate per year, or some combination. Depending on how platform services are acquired, the cost can be variable (based on, say, the number of participants each month) or largely fixed (because of development costs or because the charge is a flat rate per year).

To estimate the cost of a platform with only the functions needed by potential adopters, we used information from the HF and PvO trials to specify how long the platform would be required, the likely number of enrollees, the design of the intervention, and the tasks the platform would be expected to automate; some, such as text messages, involve per-participant fees. In those two trials, an average of 49 people enrolled per month and required all the functions of the platform for the duration of the intervention. We used the average enrollment for both trials, rather than only one, because both trials drew from the same patient population, and, during the period when both were recruiting, eligible patients were divided between them. Another 73 people per month started to enroll but were determined through the enrollment surveys to be ineligible, thus incurring only the costs of enrollment; the cost of enrolling ineligible patients are relevant if adopters would follow a similar path to determine who is eligible for the intervention. Based on their clinical experience, the clinicians on the trial team estimated that a new program would be offered for at least three years before it was reviewed and perhaps discontinued. Thus, to allow the last enrollees to complete the intervention, the platform would be needed for 3.5 years for a six-month intervention, or four years for a twelve-month intervention. Over that period the total number of participants would be 1,764 (49 enrollees per month times 36 months).

Because the details of the financial incentives had little effect on platform costs, we distinguished three types of interventions based on duration and types of messaging required.

1. All three HF interventions offered *daily sweepstakes incentives for 6 months* based on medication adherence and required one reminder and one feedback message each day (telling the participant whether and what s/he won) for a total of 362 messages over the course of the intervention.
2. The Process and Process+Outcome arms of PvO offered *daily sweepstakes incentives for 12 months* and required one feedback message each day based on medication adherence, plus messages to remind participants to get lab tests in the Process+Outcome arm, for a total of 360-385 messages over the course of the intervention.
3. The Outcome arm of PvO offered *payments based on LDL lab tests given every 3 months over the course of a twelve-month intervention* and required a minimum of 8 messages for participants who completed their labs after the first request (one to alert them that it is time for the lab and the second to let them know the results and any payments) and a maximum of 24 messages for participants who needed 4 reminders for each lab test.

Other messages served the research aims of the trials and were not included in the specifications used to estimate the costs of a potential adopter.

Based on these specifications, total platform costs for the period during which each intervention was active included implementation support, hosting and security in compliance with HIPAA requirements, ongoing support maintenance, integration with a payment service that issues debit cards, and per participant messaging fees. Estimated total costs came to \$156,920 for the six-month daily sweepstakes intervention, \$173,744 for the twelve-month daily sweepstakes intervention, and \$155,648 for the twelve-month lab-based incentives intervention (2019 dollars). For the specified 1,764 participants, total cost per participant is thus \$89, \$98, and \$88, respectively. The cost estimates assume that the adopter would not conduct research during or after the intervention, so would not need the additional secure data storage and staff required for such research, and that the intervention is offered by a health plan or employer with a patient population similar in size and composition to that of the clinical practices that participated in the trials. Because the trials are the sources of the cost data, it would be difficult to avoid this assumption.

Staff

Staff time is a key resource for implementation and can be a substantial expense. Potential adopters need to know what kind of personnel, with what kind of training, are needed to implement the intervention and how much time, on average, will be required per participant. The list of tasks developed to guide the costing (Table 1) helps them understand these requirements.

The starting point for estimating staff cost is a time log or diary, in which staff record the time spent on each task. Studies of the collection of time data indicate that the best way to

get accurate estimates is to use a log that covers an entire day – which means that it would cover both research and implementation tasks – and that is filled out during the day, not at some later point.¹¹ This design ensures that the time reported for individual tasks is controlled by the total number of hours in the workday. When tasks are identified in sufficient detail, it is straightforward to separate time spent on implementation from time spent on research. In trials with more than one intervention, like the HF/PvO trials, defining detailed tasks also allows tasks associated with different interventions to be distinguished so that each intervention can be costed accurately. It is essential to involve staff in identifying which tasks are part of implementation, for which intervention, and which are part of research or other work.

Time can be recorded daily for individual days scattered throughout the trial or for a week at a time, with several weeks of data collected during the trial. Collecting for an entire week is especially useful if tasks vary by day of the week. Data collection should be scheduled to take place once the trial is underway and the work has become routine, to avoid overestimating costs due to startup activities. To support estimates of staff time per intervention participant, the cost relevant to an adopter, the time log should include measures of activity, particularly the number of active trial participants that day or week, but also the number of devices prepared and sent, or the number of people who called for help with problems.

Table 2 shows a worksheet from the time log created for the EMPOWER trial. The EMPOWER trial's design made collecting staff time particularly straightforward. There was only one intervention arm, so no need to sort tasks by intervention, and the control arm was not contacted after enrollment, so there were no ongoing tasks associated with it (and control participants were not included in the count of active participants). The time log was set up to collect data for a full week at a time. It was completed five times – a total of five weeks – over the course of the trial. Four of those weeks fell during the period when participants were still being recruited, the fifth after recruitment was finished and the last participants were completing their year in the trial.

During recruitment, three full-time trial staff rotated tasks each week, focusing alternately on recruiting trial participants, monitoring electronic devices and alerts, or helping with either set of tasks as needed. After some experimentation, all three worksheets were set up the same way because even the recruiter and monitor sometimes helped with tasks that were not their primary responsibility that week.

In addition to the time reported by the three full-time staff, the project manager reported time spent managing the non-research tasks involved in the intervention, which averaged 4-5 hours per week, and the hours spent each week by the part-time screener, a medical student, who regularly reviewed selected medical records to identify eligible patients. To relate the screener's hours to the workload, the project manager recorded the number of charts the screener reviewed, the number of patients entered into the computer system for invitations, and the number enrolled that week, allowing screening time to be calculated per chart, per potentially eligible person, or per person enrolled.

The part-time screener reviewed electronic medical records that had been automatically selected by means of a “reporting tool” that pulled the records of patients aged 18-80 who had recently been discharged from hospital with a primary diagnosis of heart failure or a secondary diagnosis of heart failure and intravenous diuretics during the hospital stay. Reporting tools are an essential component of electronic record systems and many systems include built-in options that make it possible to set one up in a few hours. Reports can then be created daily, every few days, or weekly, depending on how quickly the practice wants to identify patients and start them in the intervention. In EMPOWER the part-time screener then reviewed the records pulled by the tool to determine more carefully whether the patient was a good candidate for the intervention, excluding, for example, those with terminal cancer.

Table 3 is a worksheet from the time log for the HF/PvO trials. These trials involved several complexities. Like intervention participants, control participants were sent electronic pill bottles, and, in the HF trial, daily reminders to take their statins, so staff spent time answering their questions. Additionally, there were a total of 8 trial arms: HF involved a control arm and three intervention arms, each intervention lasting 6 months; PvO involved a control arm and three intervention arms, each intervention lasting 12 months. The trials together had two full-time staff, a project manager, part-time help from a programmer who screened medical records to identify potentially eligible people, and several part-time student workers who assisted with research tasks.

The time log was set up to collect time for a full day and was completed on 11 separate days. The 11 days covered a representative sample of days and tasks, but experience with this trial suggested that, in general, collecting data for a week at a time might better reflect the time spent on implementation tasks, because tasks varied over the week and some were even allocated to specific days. Again, time for specific tasks was related, where possible, to a measure of that task – number of phone calls with participants, number of devices prepared and mailed, etc. – to allow potential adopters to understand staff needs and adjust staffing to their own circumstances.

For research purposes all participants in all trial arms submitted lab tests at specified intervals in order to allow evaluation of changes in LDL-c. Only two interventions, however, required lab tests for their *implementation* – the Outcomes and Process+Outcomes interventions in the PvO trial. In those two interventions, lab test results were the basis for the payments patients received. Thus, costs for lab tests would be necessary to implement those two interventions. The time log was set up to separate time spent dealing with lab tests from tasks that were common to all the interventions, to allow cost estimates specific to each intervention.

In addition, one-time estimates were made for the HF/PvO trials of the time spent screening medical records to identify potentially eligible patients and recruiting them to participate by letter or email, tasks that were handled by the project manager and a programmer, not by the full-time staff. Those tasks are not included in the time log worksheet shown in Table 3, but are shown in Table 4, the numerical example that pulls together all the cost information.

Once staff time has been measured, the hours are costed by applying an hourly wage plus fringe benefit rate. The appropriate hourly rate depends on the skills of the staff and will likely vary across potential adopters, depending on the type of people they ordinarily hire and whether existing staff are able to take on some of the work. The estimates should thus be presented in a way that allows potential adopters to use their own wage and fringe benefit rates. For purposes of illustration in this paper we used the national average hourly wage plus fringe benefits for private industry workers in March 2017, \$33.11.¹²

In all three trials, because of the exploratory nature of the research, the investigators chose not to automate some tasks that could have been handled by the technology platform. For example, the HF/PvO research team ordered lab tests and uploaded results manually. Similarly, EMPOWER staff reviewed weights recorded by the electronic scales before forwarding them to clinicians, rather than have the technology platform automatically filter out unlikely values and alert clinicians to the rest through the electronic health record. If adopters automate these tasks staff costs would be lower (and technology costs higher).

Overhead

Overhead includes the costs of the physical facilities in which and with which the staff works – computers, office space, office furniture, and supplies. It also includes the costs of functions that the organization performs for all staff and all projects, such as human resources, legal consultants, administration, etc. Potential adopters, and even private-sector trial partners such as employers, are often unwilling to share their overhead rates. Cost estimates should therefore include overhead as an explicit and separate element so that potential adopters can recalculate costs internally using their own overhead rate. We used Penn's NIH-approved overhead rate, 62%, for purposes of illustration.

Devices

Electronic devices that monitor and report health-related information are a rapidly growing industry and the devices available change frequently. Their cost consists of the device itself and the cost of its integration with the technology platform. The obvious choice is to use the price the trial paid for the device and its integration, but manufacturers frequently offer devices to researchers at lower prices than they would charge other buyers and the contracts may specify, as they did for these three trials, that the price cannot be published. The prices paid by adopters will also vary depending on volume discounts and other contract features. The only published price we could find for an electronic pill bottle, the Wisepill, was published in 2010 and is used here for illustration.¹³ Again, this element should be shown separately so that potential adopters can replace it with the price quoted by their supplier.

Incentive payments

In the HV/PvO trials, a key design feature, described earlier, was that participants were offered payments of various kinds to induce them to lower their LDL-c by taking their statins consistently. The payments earned by each participant were automatically calculated by the *Way to Health* platform, as they would be by an adopter's platform. In the Habit Formation trial, for example, payments based on reported adherence averaged \$2.35 per

participant-day across the three interventions, for a total of \$424.45 per participant over the course of the six-month trial.

PUTTING IT TOGETHER: AN EXAMPLE

What does a complete costing analysis done along these lines look like? How can it be presented to be most useful to potential adopters? Table 4 presents an example based on the six-month daily sweepstakes incentives intervention tested in the Habit Formation trial. The information is presented item by item so that potential adopters can replace items with their own estimates, experiment with their own wage and fringe rates, etc., to see how costs might differ in their circumstances.

In this example, total implementation costs per trial participant came to \$903.54 in 2017-2018 dollars. Incentive payments, at \$424 per participant over the six months, were the largest component. Staff time including overhead, \$189 per participant, and the cost of the device, \$185 per device, were next and close in magnitude. The technology platform, although a large fixed up-front cost, averaged less per participant, \$89.

Where the data come from the trial, and the potential adopter has no other source of information, it can be helpful to provide information about the possible variation in some of the items. For example, Table 4 shows the variation in the hours of staff time required per participant: mean time was 3.524 hours per participant with a standard deviation of 0.752 hours over the 11 days for which time data were collected.

Participants may incur costs for participating in an intervention, e.g., time required to complete forms, time to get lab tests, and the cost of transportation to the lab. These costs may not be of concern to the potential adopter, a health plan or employer, but it is worth keeping them in mind since they will affect individuals' willingness to participate and their success in the intervention.

DISCUSSION

Behavioral interventions differ from the clinical interventions typically evaluated in costing, cost-effectiveness, and budget impact analyses. The resources required, which can include electronic devices, integrated technology platforms, secure messaging, and staff to advise patients on their use, are not readily costed using standard health utilization or billing databases. The best and often only sources of cost data are the clinical trials that test behavioral interventions for effectiveness.

This paper describes issues that researchers need to consider in order to provide potential adopters with estimates of the costs they will likely incur if they implement a behavioral intervention that has been tested in a clinical trial. The key issue is to separate implementation costs, the costs an adopter would incur, from the costs associated with the research purposes of the trial. We have described how to frame the costing problem, including how to think about costs associated with the control group, and methods for collecting data on individual costs: the information the technology staff needs to estimate the costs of a technology platform that supports the specialized functions of the intervention;

how to set up a time log to collect data on the time staff spend on various tasks and separate out the tasks involved in implementation; and issues in getting data on device, overhead, and financial incentives costs.

Clinical trials are a good source of information about implementing an intervention, but every implementation must be adapted to the circumstances of the adopting workplace, health plan, or clinical practice. We have described how to present the cost information from a trial so that potential adopters can explore how their circumstances may affect costs. In addition, there are costs for which clinical trials are a less useful source. For example, there are start-up costs to any new activity that involve reviewing information to decide whether the activity is likely to be a worthwhile investment, and, if so, how to introduce it successfully. Adopters may also want to evaluate the behavioral intervention after implementation to determine its acceptance by their patients or employees, and their satisfaction with it. Clinical trials, with their goal of conducting research to determine the effectiveness of an intervention, are not good guides for these costs as their start-up and evaluation costs address very different objectives. In addition to planning the intervention and its implementation, trial start-up costs include planning how to implement the sampling strategy, randomize participants, and measure and adjudicate outcomes.

It should also be noted that, although costs are generally estimated and reported as though they are constant per participant, this may not be true for technology platforms. If the platform is built in-house it will constitute a fixed cost, incurred at the beginning of the implementation process, and cost per participant will decline with more participants. If platform services are purchased, the cost may be fixed or vary with the number of participants, depending on the terms of the purchase. Other implementation costs may involve economies of scale, with potential adopters that have a small eligible population incurring higher costs for the same tasks than adopters with large eligible populations. Wages are different in different parts of the country, and in different countries, and staff costs will be higher or lower as a result; showing the estimates of staff time separately makes it easy for potential adopters to apply their wage and fringe rates. While staff might be asked to work overtime for a while, thus increasing costs initially, it seems unlikely that that solution would be used for an intervention that lasts several years. Costs may also vary with the demographic composition of the eligible population. Researchers can only provide estimates for populations reasonably close in size and composition to the ones from which the trial participants were drawn.

If, however, the intervention was offered through one of the growing number of companies that specializes in providing behavioral interventions, the size of the adopting organization, an employer or a health plan, and the expected number of participants, might not matter. The adopter would be only one of many customers and the specialist company would be able to spread costs over all of them.

Although validation is beyond the scope of this paper, a useful next step would be to study the costs incurred by organizations that adopt behavioral interventions to determine whether estimates based on trial data accurately predict adopters' costs.

Our primary purpose in this paper has been to describe how to estimate the costs that would be incurred by potential adopters of an intervention. The same costs and methods could provide some of the information needed for the cost-effectiveness analyses used by policy makers to decide whether to approve new medical services for coverage. Micro-costing is often used for costing interventions, so the detailed enumeration of resources and estimates of staff time would be particularly useful. The final cost estimates would, however, need to reflect the circumstances and perspective of the decision maker. For example, insurers would want to adjust for any cost-sharing by participants, while analyses from the perspective of the health care system as a whole, recommended by the Second Panel on Cost-Effectiveness in Health and Medicine,¹⁴ would need to include the full costs of the intervention, regardless of payer, including any additional costs incurred by participants not captured in the trial.

In summary, behavioral health interventions are increasingly prevalent but are not easily costed using standard administrative and billing data sources. Employers and health plans may be interested in offering behavioral interventions, but want to know more about the likely costs of implementing them. Clinical researchers can help prospective adopters understand these costs, increasing the potential for translation of their research into practice, by undertaking a few special efforts during the trial to cost implementation tasks and publishing their cost estimates.

Acknowledgments

Funding. Financial support for this study was provided in part by grants for three randomized trials from the National Institutes for Health and CVS Caremark: EMPOWER trial, grant R01 HL128465 from National Heart, Lung, and Blood Institute (Drs Asch and Volpp, PIs); Habit Formation trial, grant R01 AG043844 from the National Institute on Aging (Drs. Volpp, Barankay, and Reese, PIs) and CVS Caremark (Grant/Award Number: 30024841), and; Process versus Outcomes trial, grant 1R01HL118195-01A1 from the National Heart, Lung, and Blood Institute (Drs. Reese, Volpp, and Barankay, PIs). The funding agreement ensured the authors' independence in designing the study, interpreting the data, writing, and publishing the report.

References

1. Notley C, Gentry S, Livingstone-Banks J, Bauld L, Perera R, Hartmann-Boyce J. Incentives for smoking cessation. *Cochrane Database Syst Rev*. Version published: 17 7 2019, 10.1002/14651858.CD004307.pub6
2. Thirumurthy H, Asch DA, Volpp KG. The Uncertain Effect of Financial Incentives to Improve Health Behaviors. *JAMA* 4 16, 2019 Volume 321, Number 15 1451–2. [PubMed: 30907936]
3. van Driel ML, Morledge MD, Ulep R, Shaffer JP, Davies P, Deichmann R. Interventions to improve adherence to lipid-lowering medication. *Cochrane Database of Systematic Reviews* 2016, Issue 12 Art. No.: CD004371. DOI: 10.1002/14651858.CD004371.pub4.
4. Agency for Healthcare Research and Quality (AHRQ). Healthcare Cost and Utilization Project (HCUP). <https://www.hcup-us.ahrq.gov/>. Accessed 10 October 2019.
5. Centers for Medicare and Medicaid Services (CMS). Fee Schedules – General Information. <https://www.cms.gov/Medicare/Medicare-Fee-for-Service-Payment/FeeScheduleGenInfo/index.html>. Accessed 10 October 2019.
6. Wagner TH, Willard-Grace R, Chen E, Bodenheimer T, Thom DH. Costs for a health coaching intervention for chronic care management. *Am J Manag Care*. 2016 4 1;22(4):e141–6. [PubMed: 27143350]
7. Wagner T. Estimating the Cost of an Intervention. Veterans Administration Health Services Research and Development, Health Economics Resource Centers webinar, 31 1 2018.

8. Mehta SJ, Volpp KG, Asch DA, Goldberg LR, Russell LB, Norton LA, Iannotte LG, Troxel AB. Rationale and Design of EMPOWER, a Pragmatic Randomized Trial of Automated Hovering in Patients with Congestive Heart Failure. *Circulation: Cardiovascular Quality and Outcomes*. 2019 4;12(4).
9. Putt ME, Reese PP, Volpp KG, Russell LB, Pagnotti D, McGilloway R, Finnerty D, Hoffer K, Chada S, Barankay I. The Habit Formation Trial of Behavioral Economic Interventions to Improve Statin Use and Reduce the Risk of Cardiovascular Disease: Rationale, Design & Methodologies. *Clinical Trials: Journal of the Society for Clinical Trials*, forthcoming.
10. Luce BR, Manning WG, Siegel JE, Lipscomb J. Estimating costs in cost-effectiveness analysis, chapter 6 in Gold MR, Siegel JE, Russell LB, Weinstein MC, eds, *Cost-Effectiveness in Health and Medicine*. Oxford, 1996.
11. Juster FT, Stafford FP. The allocation of time: empirical findings, behavioral models, and problems of measurement. *J Econ Lit*. 1991;29:471–522.
12. Bureau of Labor Statistics, U.S. Department of Labor. News release: Employer costs for employee compensation – March 2017. USDL-17-0770 released 6 9, 2017.
13. Haberer JE, Kahane J, Kigozi I, Emenyonu N, Hunt P, Martin J, Bangsberg DR. Real-time adherence monitoring for HIV antiretroviral therapy. *AIDS Behav*. 2010 12; 14(6): 1340–6. doi: 10.1007/s10461-010-9799-4. [PubMed: 20809380]
14. Sanders GD, Neumann PJ, Basu A, Brock DW, Feeny D, Krahn M, et al. Recommendations for conduct, methodological practices, and reporting of cost-effectiveness analyses in health and medicine: the Second Panel on Cost-Effectiveness in Health and Medicine. *JAMA*, 2016 9 13; 316(10):1093–103. [PubMed: 27623463]

Figure 1.
Flow chart of steps and methods.

Table 1.
Identify Tasks and Resources required

The purpose of this table is to identify the tasks involved in a trial (column 1), the subset of those tasks that a potential adopter would need to do (column 2), and the resources needed for the task. The table is an example based on the HF/PvO trials. Tasks are listed in chronological order and research tasks are shown in italics.

Trial tasks	Implementation tasks	Resources
Identify potentially eligible people	Identify potentially eligible people	Staff
Letters of invitation	Letters of invitation	Staff and computer
Email followup	Email followup	Computer platform
Confirm eligibility	Confirm eligibility	Computer platform
Enroll participants	Enroll participants	Computer platform
<i>Randomize participants to trial arms</i>		<i>Computer platform</i>
<i>Administer surveys</i>		<i>Computer platform</i>
Prepare and send electronic device	Prepare and send electronic device	Staff; device
Messages: automatic reminders, feedback on progress, notification of results	Messages: automatic reminders, feedback on progress, notification of results	Computer platform
Collect and store data for incentives, calculate incentives	Collect and store data for incentives, calculate incentives	Computer platform
Pay incentives	Pay incentives	Computer platform
<i>Lab tests needed for research: reminders, lab orders, pay lab and participants</i>		<i>Computer platform</i>
Lab tests needed for incentives: reminders, lab orders, pay lab	Lab tests needed for incentives: reminders, lab orders, pay lab	Computer platform
Help enrollees enroll and use the technology platform and electronic device; approve payments; resolve problems with labs or payments	Help enrollees enroll and use the technology platform and electronic device; approve payments; resolve problems with labs or payments	Staff
<i>Collect and store data for research</i>		<i>Computer platform</i>
<i>Meet with research team to monitor trial progress</i>		<i>Staff, research team</i>
<i>Conduct statistical analyses of trial outcomes</i>		<i>Staff, research team, computer</i>
<i>Prepare administrative and budget reports for funder</i>		<i>Staff</i>
<i>Prepare reports for safety monitoring board and institutional review board</i>		<i>Staff, research team</i>
<i>Prepare reports of trial results for journals</i>		<i>Research team, staff</i>

Table 2.
EMPOWER time log

	Week beginning (enter date)				
Day	Mon	Tue	Wed	Thu	Fri
Time on research tasks					
Team meetings devoted to research issues					
Reviewing screener's work to find more eligible patients					
Calculating adverse events					
Other research-related tasks and activities unrelated to the project (e.g. monthly staff meetings)					
Time on non-research intervention tasks					
Time screening					
Time preparing for calls, calling, and followup					
Number of phone numbers dialed					
Time on in-person recruitment visits (includes travel)					
Number of visits					
Time preparing, sending, setting up devices					
Number of devices set up					
Number of device kits sent (scale and pill bottle)					
Time troubleshooting devices					
Time reviewing/responding to alerts and nonadherence					
Time on other monitoring tasks					
Number of active enrolled participants this week					
Time on any non-study tasks					
Total intervention time (exclude research and non-study tasks)					

Table 3.**HF/PvO time log, two days**

Two full-time staff members each filled out worksheets like the one below, which were then combined. Only time spent on implementation tasks is shown. Tasks highlighted in gray were involved in only some of the interventions; see text. Early in the trial (June 15, 2015) participant communication and payment reconciliation took more time. Later (October 12, 2016) ordering, following up, and checking lab results took more time.

Activities to Track	Date: June 15, 2015		Date: October 12, 2016	
	Hours	Workload measure	Hours	Workload measure
Active trial participants on this date	--	150 people	--	135 people
Communication with participants (email/phone)	3	25 people	1	6 people
Payment reconciliation	0.5	2 people	0	0 people
Platform maintenance (create/resolve tickets)	0	0 tickets	0	0 tickets
Staff training	0	-	0	-
Order lab tests	0.75	5 lab orders	2.5	19 lab orders
Mail lab slips	0.75	5 lab slips	0.75	19 lab slips
Enter lab results	0.5	2 lab results (checked 10 people for results received)	0.75	4 lab results (checked 64 people for results received)
Lab follow-ups (phone, email, text)	-	-	2.75	40 people followed up
Ship electronic pill bottles	0.75	4 bottles	0.25	3 bottles
Configure electronic pill bottles	1	10 bottles	0.1	3 bottles
Troubleshoot electronic pill bottles	1	7 people/bottles	0	0 people/bottles
Replace electronic pill bottles	0	0 bottles	0	0 bottles
Re-run lotteries	0.5	5 people	0	0 people
total hours	8	-	8.1	-
Active trial participants on this date	-	150	-	135
Total trial participants	-	360	-	805

Table 4.
Costs for a six-month daily incentive intervention, 2017-2018 dollars

Intervention: daily incentives for six months (181 days)

Sponsor is a health plan similar in size to clinical practices at Penn Medicine

Intervention is offered for three years (156 weeks), ends after 3.5 years (182 weeks)

49 participants enroll per month, 1764 in total over three years

Type of cost/Includes	Unit cost	Cost/participant	Range	Source (see text for details)
Computer platform	\$156,920*	\$88.96*	none	Based on <i>Way to Health</i>
Recruitment of participants (ends at three years)				
Staff time to screen, set up letters	527.8 hours/3 years	\$9.91		Time recorded by staff for one week's recruitment
Invitation letters	\$0.62/letter	\$6.20	none	Cost of a letter, divided by the recruitment rate 0.1 (1 of 10 who received letters enrolled)
Participation in intervention				
Electronic device and maintenance	\$185	\$185	\$20-250	Price paid device supplier (Haberer 2010)
Incentive payments	\$2.12-2.57 per participant-day	\$424.45	\$383.72-465.17	Recorded by computer platform: \$424.45 is the average payment/participant/day times 181 days.
Staff time to help participants enroll, prepare and ship devices, help participants use computer platform and device, check and correct incentive payments, answer questions, etc.	3.524 hours/participant	\$116.68	\$91.78-\$141.58 (95% confidence interval using the standard deviation of hours, 0.752)	Hours from time log (Table 3, excluding tasks highlighted in gray), costed at the national average hourly wage plus fringe benefits for private industry workers in March 2017, \$33.11.
Overhead rate applied to staff cost	62%	\$72.34	none	NIH-approved overhead rate, University of Pennsylvania
Total cost/participant		\$903.54		

The cost of the computer platform is a fixed cost and cost per participant would vary with the number of participants.