P hdvxuhvriVxffhvv=

Qhz P hwkrgviruFdofxodwqj FrvwVdylqjvlq d SrsxodwrqKhdowk Ip suryhp hqwSurjudp

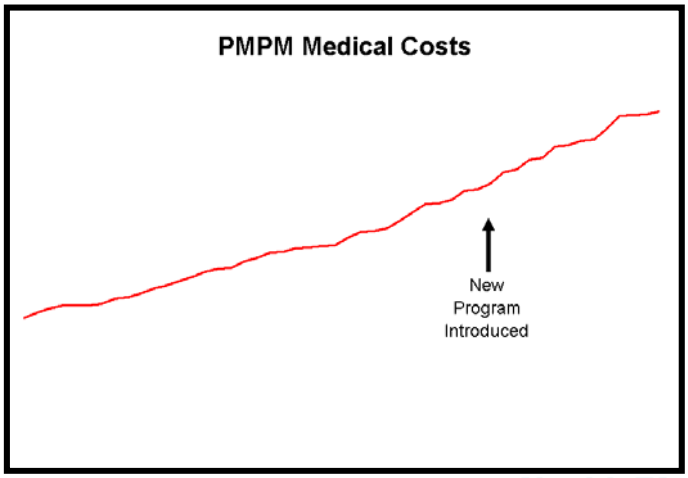
Health Dialog

Agenda

- The Challenge
- The Idea
- The Evidence
- Discussion

Health Dialog

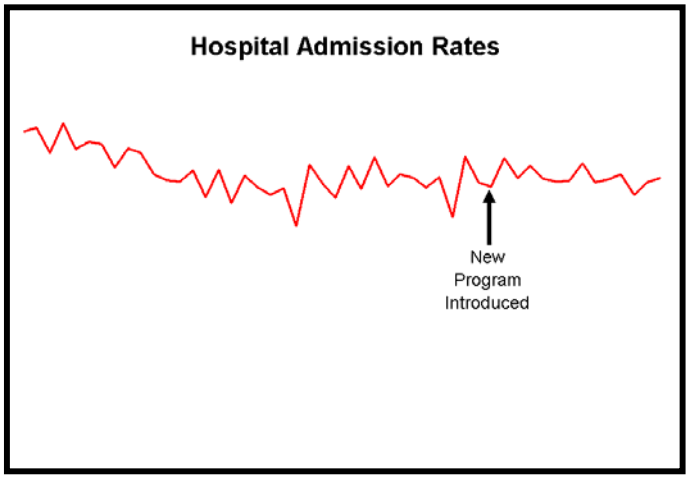
The Challenge
Did the program reduce costs?



Stylized for discussion only, not representative of actual program.

Health Dialog

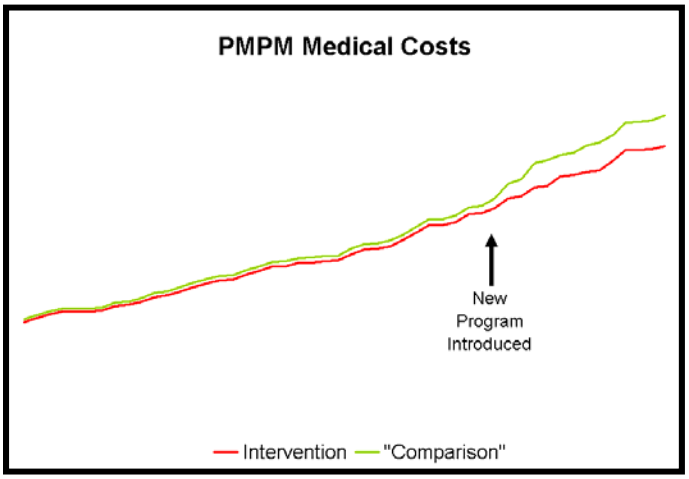
The Challenge
Did the program change utilization trends?



Stylized for discussion only, not representative of actual program.

Health Dialog

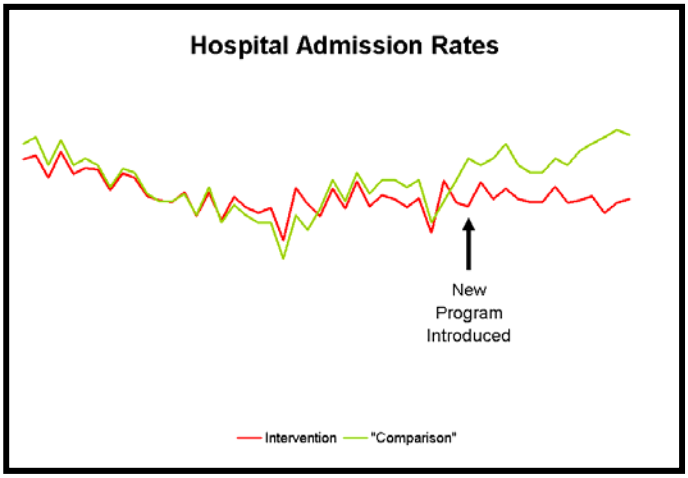
The Challenge
Did the program reduce costs?



Stylized for discussion only, not representative of actual program.

Health Dialog

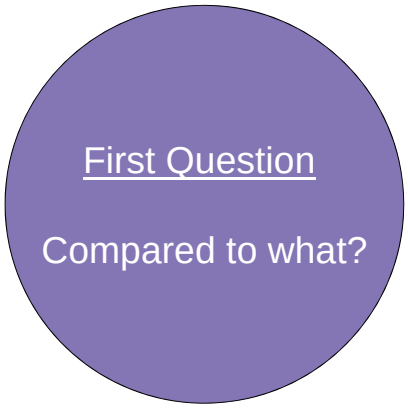
The Challenge
Did the program change utilization trends?



Stylized for discussion only, not representative of actual program.

Health Dialog

The Challenge
Did the program work?

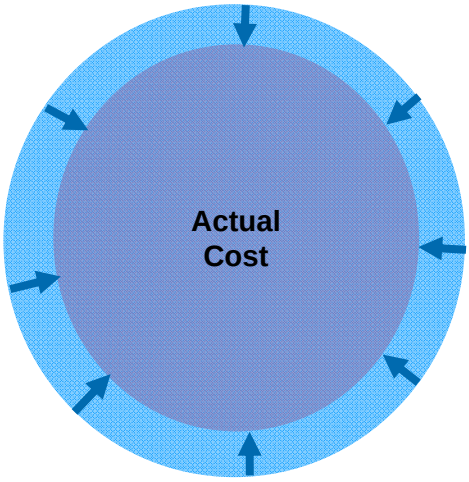


Health Dialog

Financial Impact: Basic Approach

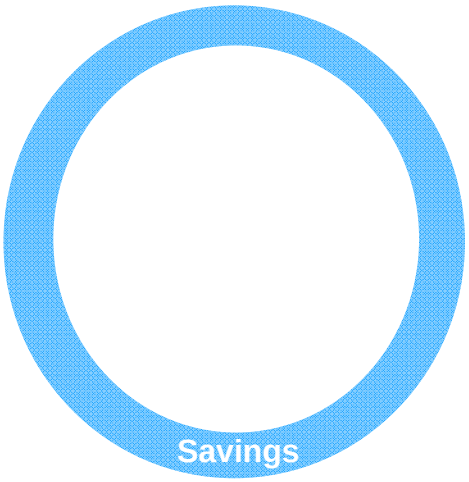
Health Dialog

Financial Impact: Basic Approach



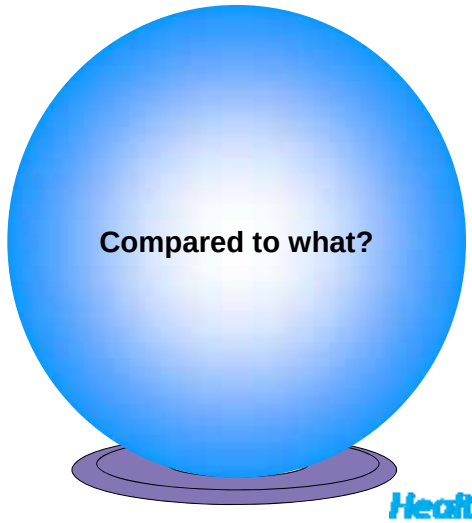
Health Dialog

Financial Impact: Basic Approach



Health Dialog

The Core Challenge



Health Dialog

The Core Challenge

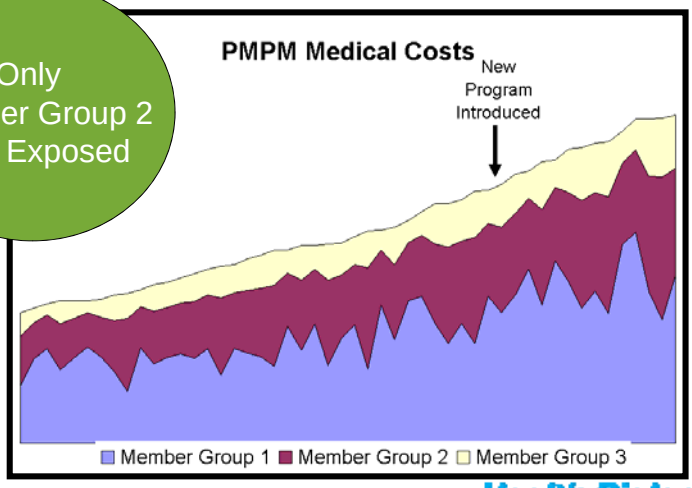
Common comparison methods

- **Randomized Controlled Trials**
- **Non-Random External Comparison**
- **Non-Random Internal Comparison**
- **Generalized Experience**

Health Dialog

The Challenge
Did the program reduce costs?

Only Member Group 2 Was Exposed



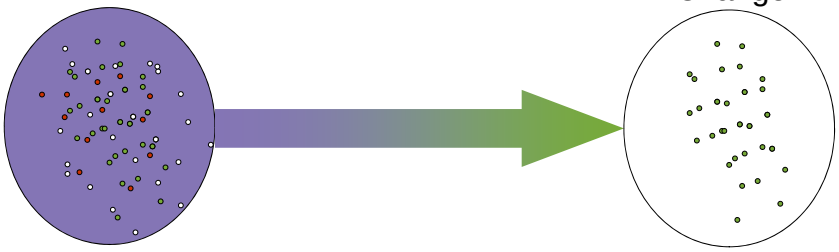
Stylized for discussion only, not representative of actual program.

Health Dialog

The Challenge
Did the program reduce costs?

Total Population

Population Expected to Drive Change

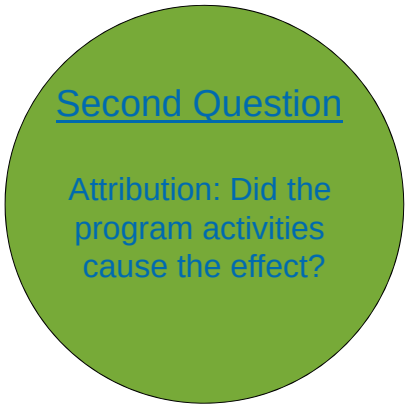


- Sent mailers
- Used web resources
- Received automated reminders
- Spoke with a Health Coach

Stylized for discussion only, not representative of actual program.

Health Dialog

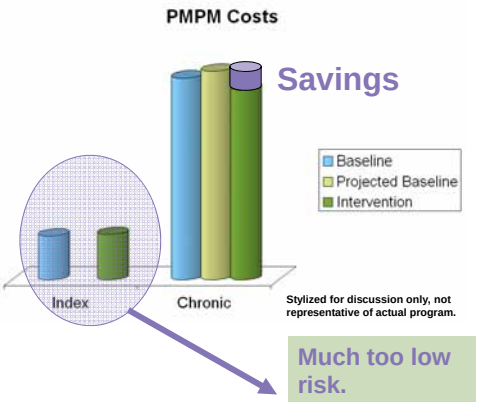
The Challenge
Did the program work?



Health Dialog

The Last Challenge
How to address the two key questions when a program is much broader than a typical disease management program?

- There are well-established methods for measuring typical chronic condition management programs.
- Typically, *Total Population Programs* target members across a broad range of conditions and risk levels, cutting deeply into the typical index population.
- The index population has risk levels that are too low to serve as a valuable reference population.



Health Dialog

The Opportunity
Ideal Test Environment

- A new whole person population health program was explicitly designed to widely replicate a program tested in a large randomized trial.
- The randomized controlled trial (RCT) provided a unique resource to test new savings methods because we had a "gold standard" measurement of actual savings.

Population Health Management Trial (RCT)

Tested standard program vs. enhanced program.*

Involved over 170,000 individuals from employers served by different health plans in different parts of the country.

Demonstrated that the enhanced program saved **$X.XX** PMPM** more than the standard program after 12 months.

*The enhanced program extended outreach well beyond the standard program, both in terms of scope (numbers of conditions) and depth (risk levels).

**Because a publication is under review on this study, actual finding is embargoed for written dissemination. The finding will be shared in person at the presentation.

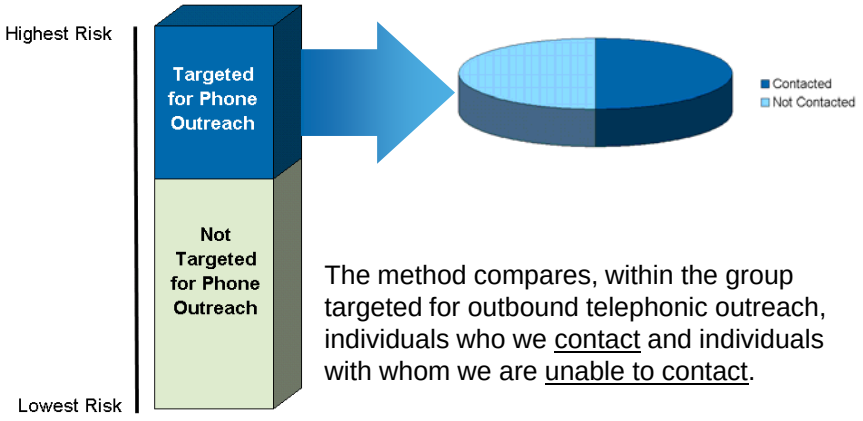
Health Dialog

Tested and Failed

- Predicting costs
- Predicting hospitalization trends
- Historical control methods
- Using external reference populations

Health Dialog

Tested and Worked!



Stylized for discussion only, not representative of actual program.

The method compares, within the group targeted for outbound telephonic outreach, individuals who we contact and individuals with whom we are unable to contact.

Members contacted have spoken with a Health Coach and not refused contact or have listened to an IVR script.

Health Dialog

The Evidence
Comparing RCT Result to Result from New Method

Test

1) Apply new savings method separately to control and intervention groups from RCT.
2) Subtract savings in the control group from savings in the intervention group to assess marginal savings level.
3) Compare to RCT savings result.

Result

Savings from RCT: *$X.XX* PMPM*
Savings from New Method: *$X.XX* PMPM*

The result using the new method was within 10 percent of the RCT result.

*Because a publication is under review on this study, actual finding is embargoed for written dissemination. The finding will be shared in person at the presentation.

Health Dialog

Why did it work?

The First Question

Compared to what?

Contacted and Not Contacted Are Very Similar

- Similar risk profiles (and baseline utilization/costs),
- Similar geographic profiles,
- Similar disease/condition prevalence levels,
- Similar demographic profiles,
- Similar benefit structures,
- Similar provider arrangements, and
- Drawn from the same time period.

Health Dialog

Big Question about Comparison

Is there participation bias?

"Contact" here means little more than answering the phone and not hanging up.

So, there is bias, but not what you'd expect…

- Being available to be contacted requires these individuals to be home.
- We try **much harder** to contact people who continue to be high risk.
- Therefore, contacted individuals are sicker on average.

Health Dialog

How do we manage this?

- First, we line up baseline risks as well as possible by:
  - Stratifying the analysis by reason for outreach and
  - Making reference periods based on months since initial outreach date.
- Then, we perform very straightforward clinical risk adjustment in the period after outreach to adjust for real clinical differences that occur after outreach.
  - Adjustments are for HCC, age, and gender.

Health Dialog

Why did it work?

Second Question

Attribution: Did the program activities cause the effect?

The measurement is focused on people we could have affected

The measurement is based on people targeted with whom some contact occurred compared to those with no contact.

While other program benefit may have occurred (mailings, web access, etc.), the core driver of impact is all that is being measured.

Health Dialog

Virtues of this approach

- It is concurrent (there is no need to account for secular forces).
- It includes all eligible members (no members are dropped from the analysis).
- It is transparent, transferable, and replicable.
- It focuses on member-specific **actual intervention** periods and it therefore reduces noise from exogenous factors unrelated to the intervention.
- Because the comparison group has similar risks to the intervention group at the point of intervention, regression to the mean affects the comparison group in very similar ways as the intervention group.

Health Dialog

Considerations

- This method requires a fairly large sample (~30K) to secure a comparison group of sufficient power
- Validation and testing of the model has been performed with a Commercial population. While our work shows that this approach should work with Medicare population measurement, the testing and validation has not reached the same level of maturity as for Commercial.

Health Dialog