1. Protocol

Title: Performance Feedback in Health Care

Principal Investigator: Felipe Lobelo, MD PhD

Version Date: December 15th, 2019

Form Authors: Felipe Lobelo

2. Objectives

The Southeast Permanente Medical Group (TSPMG) provides performance feedback to its doctors. The performance feedback designs can vary and change over time in terms of targets, summary statistics, included measures, and frequency of delivery. We would seek to compare different performance feedback designs so that we could identify which are most effective at contributing to performance improvement. This would be non-disruptive in the sense that we would simply be providing performance feedback through existing avenues and with appearances that are not out of the ordinary.

The research team will randomly assign providers into the different performance feedback conditions, as specified in the protocol. Providers will receive performance feedback through the standard mechanism in which it is conveyed by their supervisor.

The objective is to investigate how to design performance feedback for doctors to best motivate and support them in improving performance along existing strategic priorities for care delivery. We would test alternative designs of performance feedback that vary on the following dimensions: 1) targets for comparison of one's own performance, 2) frequency of performance feedback provision, and 3) team vs individual performance measures. The intervention is technically feasible using existing performance feedback design and delivery software within the organization.

We will provide feedback on performance by measures that the organization already tracks internally and seeks to improve performance by. These include a doctor's utilization of opportunities to provide flu vaccinations, cervical cancer screening, and blood pressure management, as well as a doctor's efforts to reconcile medications (i.e., check on the appropriateness of existing prescriptions and make necessary updates). The study will randomly assign different designs of feedback to physicians. In order to understand which designs of feedback have the best effects on performance, we would test the following hypotheses:

1. Displaying the next-highest quartile will motivate improvement more than display of all quartiles.

- Theory: The next-highest quartile will serve as an injunctive norm, or suggested target, to repeatedly lift an individual's performance to the suggested level.[1]

2. The positive effect of displaying the next-highest quartile, relative to displaying all quartiles, will be most pronounced for initially low performers.
   - Theory: The next highest quartile averts upward social comparison to a much higher level of peer performance, which can be discouraging and so negatively affect performance.[2]
3. The positive effect of displaying the next-highest quartile, relative to displaying all quartiles, will diminish over time.
   - Theory: Individuals may become worn out as they see a target ratchet higher when their performance improves.[3]
4. Displaying team relative performance along with individual relative performance will be more effective than displaying either type of information alone.

3. Background

a. Scientific Background

Achievable targets benefit performance because they can motivate improvement through a series of "little victories." On the other hand, difficult targets come with the benefit of motivating people to perform better than they thought possible.[4] Either type, or a combination, may be most effective at motivating performance. Furthermore, it is unclear whether the best performance effect will result from information on one's own performance, team performance, or both. Team relative performance information can facilitate peer pressure and best practice sharing. Individual relative performance information can generate a desire to validate one's own ability by outperforming others.[5] Our study would test whether performance improves the most when feedback contains targets that are relatively achievable or difficult, and when individuals can view performance on individual performance, team performance, or both.

b. Preliminary Data

---

[1] Schultz, P. W., J. M. Nolan, R. B. Cialdini, N. J. Goldstein, and V. Griskevicius. "The Constructive, Destructive, and Reconstructive Power of Social Norms." *Psychological Science* 18(5) (2007): 429-434.

[2] Rogers, T., and A. Feller. "The Threat of Excellence: Exposure to Peers' Exemplary Work Undermines Performance and Success." Presented at the Society for Judgement and Decision Making 2015 36th Annual Conference (2015).

[3] Indjejikian, R. J., M. Matějka, and J. D. Schloetzer. "Target Ratcheting and Incentives: Theory, Evidence, and New Opportunities. *The Accounting Review*. 89(4) (2014): 1259-1267.

[4] Locke, Edwin A., and Gary P. Latham. "Building a practically useful theory of goal setting and task motivation: A 35-year odyssey." *American Psychologist* 57.9 (2002): 705.

[5] Song, H., A. L. Tucker, K. L. Murrell, and D. R. Vinson, "Closing the productivity gap: Improving worker productivity through public relative performance feedback and validation of best practices." *Management Science* 64(6) (2017): 2628-2649.; Tafkov, I. D. "Private and public relative performance information under different compensation contracts." The Accounting Review 88(1) (2012): 327-350.

N/A

4. Study Design

1. Common Procedures
   - Participants will receive an introductory email, sent from a standard administrative email used to provide such information to their work email address, with a link to the relative performance information, information describing the study.
   - Participants would be able to view their performance information using a link provided in a periodic (as delineated below) email announcing the performance information update.

2. Variable Procedures ("A" refers to one set of control and treatments and "B" refers to a second set of control and treatments. We would run these sets sequentially over the course of a year, randomly assigning the conditions in set A and observing results for three months, and then randomly assigning the conditions in set B and observing results for three months, or vice versa pending organizational discussion of which set—A or B—to implement first.)
   - Organization-Wide Target, Own Performance, Frequent (Control A.1a)
     - Doctors are shown their own performance relative to other doctors, and with the organizational target as a benchmark, sent weekly.
   - Organization-Wide Target, Own Performance, Infrequent (Control A.1b)
     - Doctors are shown their own performance relative to other doctors, and with the organizational target as a benchmark, sent bi-weekly.
   - Organization-Wide Target Plus Next-Highest Quartile, Own Performance, Frequent (Treatment A.1a)
     - Doctors are shown their own performance relative to other doctors, and with the organizational target and the next-highest quartile as benchmarks, sent weekly.
   - Organization-Wide Target Plus Next-Highest Quartile, Own Performance, Infrequent (Treatment A.1b)
     - Doctors are shown their own performance relative to other doctors, and with the organizational target and the next-highest quartile as benchmarks, sent bi-weekly.
   - Organization-Wide Target Plus All Quartiles, Own Performance (Treatment A.2a)
     - Doctors are shown their own performance relative to other doctors, and with the organizational target and all quartiles as benchmarks, sent weekly.
   - Organization-Wide Target Plus All Quartiles, Own Performance (Treatment A.2b)
     - Doctors are shown their own performance relative to other doctors, and with the organizational target and all quartiles as benchmarks, sent bi-weekly.
   - Organization-Wide Target Plus All Quartiles, Own Performance (Control B)
     - Doctors are shown their own performance relative to other doctors, and with the organizational target and all quartiles as benchmarks, sent bi-weekly.

- Organization-Wide Target Plus All Quartiles, Team Performance (Treatment B.1)
  - Doctors are shown the performance of their team (a "pod," or group of doctors who practice together) relative to other teams, and with the organizational target and all quartiles as benchmarks, sent bi-weekly.
- Organization-Wide Target Plus All Quartiles, Own Performance and Team Performance (Treatment B.2)
  - Doctors are shown their own performance and the performance of their team (a "pod," or group of doctors who practice together) relative to other teams, and with the organizational target and all quartiles as benchmarks, sent bi-weekly.

5. Study Population

a. Number of Subjects

The number of subjects (doctors) will be approximately 500. The number will include all of the physicians in the primary care and specialties for whom the performance feedback measures are relevant to their provision of health care. This number naturally fluctuates over time with employment status changes.

The subjects are doctors and identifiable data as well as the intervention will be restricted to their performance toward already existing organizational encouragement for their care of patients including providing flu shots, conducting pap screening, treating blood pressure, and reconciling medications.

b. Inclusion and Exclusion Criteria

The inclusion criterion is that the providers are:

- A physician practicing at The Southeast Permanente Medical Group in a specialty for which the performance measure being studied (flu, medication reconciliation, blood pressure management, or pap screening) is relevant.
- Physicians routinely receive similarly formatted performance feedback, and the formats and measures included regularly change every few months, and so the experimental performance feedback would in line with this precedent. Doctors can ignore the performance information be declining to open it both in ongoing instances of performance feedback delivery at The Southeast Permanente Medical Group and in the proposed study. Any such doctors' performance data will be retained by the organization and used in the study.

c. Vulnerable Populations

All of these groups will be excluded from the study.

Decisionally Impaired Adults

Decisionally impaired adults will not be included.

Other Populations Targeted for Recruitment

The performance feedback will be delivered physicians employed by The Southeast Permanente Medical Group. It would be provided in line with existing norms of the organization for providing performance feedback. It will not introduce risks in terms of altering contracts or pay for performance. It is meant to provide information that helps the doctors identify areas for improvement and track their progress toward those areas. This is a common practice at the organization.

d. Setting

The research would be conducted at The Southeast Permanente Medical Group in Atlanta.

e. Recruitment Methods

We will send an introductory email with a link to the performance feedback. This email will describe that performance feedback will vary among doctors as researchers study which is most effective. The Southeast Permanente Medical Group has an ongoing practice, which predates this study, of regularly providing doctors with emails that share performance feedback.

The provision of these feedback, with display designs that frequently change along the dimensions in our study, is the norm for doctors in their medical group. We thus expect that provision of performance feedback with varying designs should be expected by the doctors and would allow them to function in their roles with minimal if any disturbance to doctors' expected access to performance information.

f. Informed Consent Process N/A

Waiver of Informed Consent N/A

Waiver of Signed Informed Consent

We are requesting a waiver of signed informed consent. The research involves no more than minimal risk to the subjects and involves no procedures for which written consent is normally required outside of the research context. We would be following standard processes for providing performance feedback within the organization under the guidance of administrators. It is normal for performance feedback to contain all or a subset of the features that we are testing, and to vary over time in how these are combined. It is not normal for employees at the organization to be asked for their consent before being provided with performance feedback, and the performance feedback in this study is similar to what is often provided at the organization. The introductory email will contain the required consent elements. Clicking on the link to the performance feedback included in the email will be the study participant's consent to be in the study. See the introductory email located in the appendix.

Alteration of Informed Consent

N/A

Non-English-Speaking Subjects
N/A

Assent of Children and Parent Permission
N/A

Adults Unable to Consent/Decisionally Impaired
N/A

g. HIPAA Privacy Rule Authorization – if study will use or disclose Protected Health Information (PHI)

The study will not include PHI.

Waiver of HIPAA Privacy Rule Authorization
N/A

6. Study Procedures

The study will involve delivering performance feedback of the given type sent with the given frequency delineated in section 4 under "variable procedures". Organization administrators will review the data as part of standard internal control procedures during the implementation of the field experiment.

The results of the experiment will be provided in a way that does not include personally identifiable information to co-investigators to perform analysis.

The study will not involve the introduction or use of medical devices or treatments that are not already in practice and encouraged by the organization to be used as a means of preventative care (flu, cervical cancer screening, blood pressure management, and medication reconciliation).

The performance data will be recorded using existing practices and provided to the co-investigators. The survey data will be gathered using Qualtrics through a secure account that keeps respondents anonymous. The survey questions only address doctor opinions regarding the performance feedback, and do not involve any data related to personal health information.

A given doctor will be involved in the study for a maximum of 6 months. This will involve, at the most frequent, the delivery of weekly performance feedback.

The performance feedback will be delivered using standard approaches already in place in the organization and there will not be a period of seeking out participants for enrollment.

This study should be completed by 2020. Analysis of the data may need to continue thereafter depending on the number of rounds of review that the drafts undergo at journals.

Doctors can decline to open the performance feedback and the experiment-related survey. As described in section 5 of this protocol, the organization does not give doctors the option to request that performance feedback is not delivered. In compliance with these norms, doctors will be able to ignore and not open the performance feedback but will not be able to terminate its delivery.

When the study is terminated, doctors will receive their last delivery of the experimental performance feedback but continue to be provided with performance feedback that The Southeast Permanente Group provides in varying iterations over time.

Exhibit A – Data Fields, Descriptions, and Examples

Field	Description	Example Values
PatientEncounterID	Unique numerical identifier (encrypted)	583, 373, 455
PatientID	ID code (encrypted)	372633, 42316
ProviderID	Medical staff office roster number (encrypted)	1206, 51, 707
SpecialtyNM	Physician medical specialty	Primary Care, Urology, Orthopedics
PatientEncounterDayNM	Day of the encounter relative to first day of the dataset	15, 24, 62
PatientAgeDSC	Age of patient, listed as 89 if older than 89	19, 58, 56
GenderDSC	Gender of patient	Male, Female, Unknown
EthnicGroupDSC	Ethnicity of patient	Hispanic/Latino, Non-Hispanic/Non-Latino, Unknown
ProviderTypeDSC	Provider role type	Physician, Anesthesiologist

cerv_scn_sched	Cervical screen scheduled during the encounter	Y, N
cerv_scn_admin	Cervical screen administered during the encounter	Y, N
Flu_scn_sched	Flu vaccination scheduled during the encounter	Y, N
Flu_scn_admin	Flu vaccination administered during the encounter	Y, N

a. Data Analysis

The statistical procedures would be t-tests that compare means among the groups of doctors. Doctors will be randomly assigned to receive a given performance feedback design included among those listed under the study design section.

In order to gain assurance that there will be sufficient statistical power, we have looked at prior studies of performance feedback to see their sample size. This includes a Journal of the American Medical Association study of performance feedback on doctors.[6] Our study would include more doctors per treatment and control group than in those studies and test similar interventions regarding performance feedback design; this provides confidence that there will be sufficient power to test our hypotheses.

The data are already collected using internally validated systems for measuring performance and are used for existing operational decision making. This helps to ensure the quality of the data.

b. Sharing of Results with Subjects

The results of the study would be intended for publication in a health care/policy and/or economics journal. This would be used by the organizations' administrators and could also be shared with the doctors, but this would be done at the discretion of the organization's administrators.

c. Data and/or Specimen Banking

[6] Kiefe CI, Allison JJ, Williams OD, Person SD, Weaver MT, Weissman NW. Improving Quality Improvement Using Achievable Benchmarks For Physician Feedback: A Randomized Controlled Trial. *JAMA.*2001;285(22):2871–2879. doi:10.1001/jama.285.22.2871; Tafkov, I. D. "Private and public relative performance information under different compensation contracts." The Accounting Review 88(1) (2012): 327-350.

There will not be data or specimen banking.

7. Privacy, Confidentiality and Data Security

Describe the plan for storage of data and/or specimens.

The data will be stored internally at the organization and also shared with the principle investigator and the co-investigators. The data will be kept on their encrypted computers. The data will have unidentifiable codes attached to subjects (doctors) and then the performance variables for those doctors as well as demographic information. There will be no PHI. The data will not be made available for release. The data will be retained internally at The Southeast Permanente Medical Group as is standard procedure for performance data but destroyed from the principal investigator and co-investigators' computers after the planned papers deriving from the study have been published.

Collection of data from subjects electronically

The study would involve a post-experiment survey to help in understanding whether doctors find the studied types of performance feedback informative and motivating. The survey would be distributed toward the end of the experiment in the last two weeks. The survey would be completed electronically and organization administrators would remove the email addresses of the respondents before providing the data to any non-Kaiser Permanente employed investigators on this study.

Description of data-sharing process with collaborators

The data for the project, which would be shared outside of Kaiser Permanente, is listed below. The data do not include PHI or identifiable information. The data sharing would occur through Kaiser Permanente's preferred secure file sharing service to the co-investigators who are not employed at Kaiser Permanente. **Exhibit A** highlights variables that will be shared with outside collaborators.

Does this study involve the disclosure of PHI to a collaborator?

No, this study does not involve the disclosure of PHI to a collaborator.

8. Provisions to Monitor Data to Ensure the Safety of Subjects

The study will not involve more than minimal risk to subjects. They will be receiving performance feedback that is similar in nature to the performance feedback that they already receive. The study does not introduce financial payoffs or contractual changes.

9. Risks and Benefits

a. Risks to Subjects

The study will provide performance feedback on dimensions of performance that the field site already reports to doctors, displayed in formats that vary on dimensions that the field site already frequently varies across all doctors among periods of time. Thus, our intervention is not likely to introduce risks that are not already inherent to the type of performance feedback sharing that the field site conducts. The main difference is that we will be varying the dimensions of performance feedback among doctors at the same time. This introduces a risk in that doctors may feel displeased with the difference between their access to performance feedback of a certain variant and other doctors' performance feedback. We discussed this risk with TSPMG executives, who themselves are doctors, who said that they expect little-to-no sharing among physicians of performance feedback reports given that doctors rarely if at all discuss personal reports with each other. In order to minimize any displeasure among doctors in the case that they notice variation in performance report formatting, we will explain in our introductory email that the performance feedback report display designs will vary among doctors as researchers study which is most effective. This will provide doctors with a reason for the variation so that they understand that there is not a doctor-specific reason why a certain doctor receives a certain type of feedback. Each doctor will receive a temporary, anonymous, subject ID during their session. All study data, outside of that which would be kept on the field site's servers as part of its regular operations, will only be tied to that ID, rather than to any identifying information. The fact that a subject participated in the study will be recorded by the field site for record keeping and auditing purposes.

b. Potential Benefits to Subjects

The study's goal is to facilitate and encourage improvement in the delivery of medical care. The performance feedback can provide an immediate benefit to the doctors receiving it by allowing them to track their progress toward key organizational objectives. After the study, the results can guide the organization's administrators in designing subsequent performance feedback, after the study is complete, and so could continue to help the organization and its employees' be successful in their efforts to improve the delivery of medical care.

10. Economic Burden to Subjects

There would not be an economic burden for subjects–the performance feedback is free and does not introduce penalties for doctors who have low performance or payoffs for doctors who have high performance.

11. Compensation to Participants

In line with prior research, including studies of performance feedback, we would offer compensation for subjects to open the experimental material, in recognition of time inconvenience as described above.[7] We would give each doctor $10 if he or she opens a

[7] Gill, D., Kissova, Z., Lee, J., & Prowse, V. First-place loving and last-place loathing: How rank in the distribution of performance affects effort provision. Management Science. 2018; Warner AS, Shah N, Morse A, et al. Patient

performance feedback report that we send as part of the experiment, which would be added to an Amazon gift card that the organization would distribute to each doctor at the end of the experiment.

12. Resources Available

The research team has prior experience and existing logistical capabilities within the organization (i.e., for performance feedback generation and delivery), but no other special resources will be required. The team includes advisors who are administrators at The Southeast Permanente Medical Group and who understand performance in the delivery of health care. The administrators are able to design and distribute the performance feedback as laid out in this protocol. The researchers have conducted performance feedback experiments, including in the context of health care, producing papers that are either under review or accepted at leading academic journals.

13. Prior Approvals

The AMD at The Southeast Permanente Medical Group who initially contacted the researchers about this experiment have held multiple meetings, including a full day on-site visit with the full research team, to plan the experiment. In emails, the administrators have indicated approval of the experiment as proposed herein.

14. Drugs or Devices

N/A

15. Multi-Site Research

N/A

16. Community-Based Participatory

Research N/A

---

and Physician Attitudes Toward Low-Value Diagnostic Tests. JAMA Intern Med. 2016;176(8):1219–1221. doi:10.1001/jamainternmed.2016.2936

Appendix.

Research Communication Example:

PERMANENTE MEDICINE®
The Southeast
Permanente Medical Group

Variable target setting to improve flu vaccination performance in quality of care at TSPMG

The quality and patient safety department is conducting research that affects providers in adult primary care and many specialties. The objective of the research is to study if variable target setting changes behavior that leads to improvement by comparing the effects of providing different designs for flu vaccination reports. There is a small incentive for participation and improving quality metrics. The research will involve providers who are PCP's in Adult Primary Care as well as providers from most specialties. The projects is fully funded by TSPMG.

The specific question we are trying to answer is whether customizing a target (blinded) closer to the current performance for a provider or a team affects improvement in the metric. For example, if PCP A has a 10% performance and PCP B has a 50% performance, we would set the target for PCP A to 20% and 60% for PCP B.

This research is not related to P4P. There is an expectation that there is not open discussion about the individual targets amongst providers unless they are part of an identified team for research purpose.

We would like your participation in this important research as there is little data on provider behavior for target setting. This would help shape more appropriate and individualized goals for providers and teams of providers that will eventually improve the health of our entire membership @KPGA.

———————————————————————————————————

DISCUSSION

There has been a slow but steady improvement in many quality metrics over the last several years, however, we have not been able to reach our goals in a few measures. While many physicians have exceeded the targets set, there are more that have not achieved threshold. There is little more that the top performers can contribute to our overall success. This demonstrates that there is a great deal of variation in physician level performance for most quality measures. How do we get to the next level to improve the quality of care to all of our patients? There is little known research on standard goal-setting techniques applied to physician's performance in actual clinical practice and even less on setting goals that are individualized and variable, based on actual performance. An overwhelming amount of data available also presents its challenges to the providers.

We are excited to inform you of new research that will involve Adult Primary Care providers and other specialties. Our key research investigates if individualized target setting based on current physician's performance can motivate improvement in quality metrics with financial incentives. The hypothesis is the following: if the target appears to be attainable, there will be concerted effort by the provider to achieve this target rather than ignoring a target that seems elusive, for which no effort will be made to achieve.

We hope that this research stimulates thinking and action to advance our collective ability improve the performance of our clinical quality measures and the care we provide as The Southeast Permanente Medical Group.

FAQ's

Why are you doing this research?

There is little research in provider behavior as it relates to motivation for improvement of quality metrics. We would like to test if there is behavior change to increase performance for chosen metrics when receiving data with variable targets, with or without incentives. There is presumption that if a target for a metric is attainable, that the provider will be motivated to improve compared to a provider who has a target that is very far from performance.

What do I need to do?

First, continue to work on increasing your rate of successful closure of flu vaccination when there is an opportunity during a face to face encounter, regardless of reason for visit. View your performance by reviewing your Successful Opportunities Report for Flu Vaccination that will be sent biweekly via email in Outlook. All you need to do is:

- Open the email and click on the link to view your individual report.

This aligns with the organizational goal to increase flu vaccinations, and the expectation of closing the care gap during any visit.

When will the research start and how long will the project last?

The first phase is anticipated to start December 15th and last for 10 weeks.

Depending on the initial results, we may continue the project later into 2020 with different metrics and/or different teams.

What compensation can I possibly receive?

The incentive is small and will be further explained when the research project begins. A maximum of 5 payments of $10 each up to total $50. If you receive an incentive, TSPMG will cover any tax burden from this incentive. Future phases may or may not have an incentive.

Who is funding this research?

This research project is fully-funded internally by TSPMG. Kaiser Health plan is not involved and that there is no external or government funding for this project.

Does this affect my P4P?

No, the incentive will provide extra compensation, outside of P4P. This is separate from any MOU incentive. If you receive an incentive, it is your responsibility to declare the amount received on your individual tax return.

How will I receive my incentive?

At the end of the research period, you will receive a cumulative incentive for each time you clicked on the link and reviewed your performance for any given 2-week period.