Title of Study or Project:	Exploratory analysis of preconditions for tailoring performance feedback to MPOG Anesthesiology Providers
Primary Institution:	University of Michigan Medical School, Dept of Learning Health Sciences
Principal Investigator:	Zach Landis-Lewis
Co-Investigators:	Nirav Shah
Statistician(s):	Zach Landis-Lewis, Colin Gross
Type of Study:	Exploratory
IRB Number/Status:	HUM00163880 - Determined not regulated as QI/QA
Hypotheses / Aims:	We aim to establish expectations for the number of providers in MPOG whose care quality data meet criteria for tailored feedback messages each month. This expectation will inform a power analysis for a grant application proposing a 3-arm trial comparing the effects of feedback messages that are a) tailored, b) a "one-size-fits-most", and c) an existing standard form.
Number of Patients/Participants:	~3,500 anesthesiology providers in MPOG
Power Analysis:	This exploratory proposal will be used to provide a power analysis for a subsequent full PCRC proposal
Proposed statistical tests/analyses:	Descriptive statistics for the monthly prevalence and magnitude of 4 preconditions (negative gap, positive gap, negative trend, positive trend) in provider performance data, across all 20 ASPIRE quality measures.
Resources (Brief summary of resources for data collection, personnel, financial):	DISPLAY lab team, Landis-Lewis NIH K01 award, ASPIRE team

Introduction

What is the significance of the clinical problem being addressed?

Healthcare organizations have tremendous opportunities to use their data to improve care quality and outcomes[1–3]. However, an epidemic of provider dissatisfaction with the electronic environment suggests that communication about clinical data is frequently inappropriate[4,5]. An increasingly common source of such communication is email-based feedback reports and notifications about quality dashboards (i.e. feedback interventions)[6,7]. Inappropriate feedback may lead to unintended consequences both in terms of low engagement (e.g. perceived spam, alert fatigue), and high engagement (e.g. gaming, tunnel vision)[8]. To reduce these unintended consequences we require new knowledge about appropriate electronic performance feedback.

What current gaps exist in the understanding of this problem?

Evidence about clinical feedback interventions shows a pattern of mixed effects over decades of trials (median 4.3% absolute improvement, IQR 0.5-16%)[9]. This effect variability reveals a significant lack of knowledge about appropriate feedback interventions. Knowledge about the mechanisms of action through which performance feedback appropriately influences clinical practice could significantly improve provider satisfaction with feedback interventions and the resulting quality of care[8,10]. A primary barrier to appropriate communication may be that feedback interventions lack functionality to prioritize feedback, thereby failing to accommodate providers' limited time and attention[8].

How will this project address this gap and advance clinical care and/or research knowledge?

Message tailoring has successfully changed health behaviors via theoretical mechanisms of influence on an individual's capability and motivation[11–13]. Causal pathway models[14] can represent the *mechanisms* through which feedback interventions influence provider capability (e.g. awareness of performance) and motivation (e.g. gain/loss orientation), and the *preconditions* for mechanism activation. *Preconditions* can be specified using information content of feedback, defined as *gaps* (a distance between current performance and a comparator) and *trends* (consistently directed change in performance). We developed message tailoring software that uses these models to analyze clinical quality performance data for prioritization of appropriate feedback interventions. For example, an anesthesiology provider participating in MPOG may have preconditions for some, but not all of the 20 routine quality measures that are delivered each month, due to changes in performance and the delivery of prior messages. A high-priority tailored message might emphasize a newly developed gap between a provider and a peer benchmark, rather than display data for all 20 measures.

What are the primary (and secondary if applicable) aim(s) / hypothes(es)?

Our primary objective is to establish expectations for the number of providers who could receive tailored feedback messages about care quality each month, based on preconditions of mechanisms for appropriate feedback. This analysis will inform a grant proposal for a 3-arm trial comparing the effects

of monthly emails containing: a) a tailored message, b) a “one-size-fits-most” message, or c) an existing standardized message.

Methods

Study Design

We propose to conduct an exploratory study to establish expectations for MPOG provider eligibility to receive tailored performance feedback messages. These expectations will inform the design of a subsequently proposed trial of tailored feedback messaging about the quality of care, to be submitted in a future PCRC proposal.

The proposed analysis has been determined to have “Not Regulated” status via the University of Michigan IRB (HUM00163880). A letter of determination of “Not Regulated” status is included in Appendix 1. The letter includes the following statement:

Based on the information provided, the proposed study does not fit the definition of human subjects research requiring IRB approval (per 45 CFR 46, 21 CFR 56) because in this case, it is activities or procedures rather than human subjects that are the object of the study.

We have completed sections of the SQUIRE reporting guidelines that are relevant to the proposed analysis as an exploratory study that is preliminary to the initiation of a quality improvement intervention (Appendix 2).

Study Population

Participation in the MPOG provider feedback email program is determined at the institutional level. Each MPOG site has the ability to decide when to initiate (or stop) emails, which measures are included in the feedback emails, and who to send emails to (faculty, CRNAs, residents)

We propose to include all sites that participate in the MPOG provider feedback program in the analysis, and to include providers’ quality measure data recorded between January 1, 2015 and July 31st, 2019.

Inclusion Criteria

- Dates: Between January 1, 2015 and July 31st, 2019
- All providers receiving ASPIRE quality measure feedback email
- All measures included in the emails (Table 1)

Exclusion Criteria

- Measures with fewer than 10 observations per measurement period

The proposed object of study is the performance information of anesthesiology providers in MPOG that can be used to create tailored messages about the quality of care. The characteristics of the performance information that we will observe are the *preconditions* for activating theoretical

mechanisms which feedback messages can use to influence practice. Preconditions, together with mechanisms, moderators, and outcomes, form causal pathway models that message tailoring software can use to identify and prioritize appropriate performance feedback. We use causal pathway models to qualitatively represent knowledge about appropriate performance communication. These qualitative models are likely to form a foundation for quantitative predictive models that are beyond the scope of this analysis.

Primary outcome

The primary outcome of the analysis will be the average monthly sum of MPOG providers whose preconditions would enable them to receive a tailored message for any ASPIRE quality improvement measure.

Table 1: Performance Measures Studied

Anesthesiologist Performance Measure	Measure Specification with Inclusion/Exclusion Criteria
AKI 01: Acute Kidney Injury	https://mpog.org/files/quality/measures/AKI-01_spec.pdf
BP 01: Low MAP Prevention	https://mpog.org/files/quality/measures/BP-01_spec.pdf
BP 02: Avoiding Monitoring Gaps	https://mpog.org/files/quality/measures/BP-02_spec.pdf
CARD 02: Avoiding MI- Troponin I <0.6	https://mpog.org/files/quality/measures/card-02_spec.pdf
FLUID -01-C: Minimizing Colloid Use- Cardiac	https://mpog.org/files/quality/measures/FLUID-01-C_spec.pdf
FLUID 01-NC: Minimizing Colloid Use- Non-Cardiac	https://mpog.org/files/quality/measures/FLUID-01-NC_spec.pdf
GLU 01: High Glucose Treated	https://mpog.org/files/quality/measures/GLU-01_spec.pdf
GLU 02: Low Glucose Treated	https://mpog.org/files/quality/measures/GLU-02_spec.pdf
MED 01: Avoiding Medication Overdose	https://mpog.org/files/quality/measures/MED-01_spec.pdf
NMB 01: Train of Four Taken	https://mpog.org/files/quality/measures/NMB-01_spec.pdf
NMB 02: Reversal Administered	https://mpog.org/files/quality/measures/NMB-02_spec.pdf
PONV 01: Post Operative Nausea and Vomiting	https://mpog.org/files/quality/measures/ponv-01_spec.pdf
PONV 02: Post Operative Nausea and Vomiting- Pediatrics	https://mpog.org/files/quality/measures/ponv-02_spec.pdf
PUL 01: Low Tidal Volume- Less than 10 ml/kg	https://mpog.org/files/quality/measures/PUL-01_spec.pdf
PUL 02: Low Tidal Volume- Less than or equal to 8 ml/kg	https://mpog.org/files/quality/measures/PUL-02_spec.pdf
PUL 03: PEEP Utilization	https://mpog.org/files/quality/measures/PUL-03_spec.pdf
TEMP 01: Active Warming	https://mpog.org/files/quality/measures/TEMP-01_spec.pdf
TEMP 02: Core Temperature Measurement	https://mpog.org/files/quality/measures/TEMP-02_spec.pdf
TEMP 03 (MIPS 424): Perioperative Temperature Management	https://mpog.org/files/quality/measures/TEMP-03_spec.pdf
TOC 01: Intraoperative Transfer of Care	https://mpog.org/files/quality/measures/TOC-01_spec.pdf
TOC 02 (MIPS 426): Post-Anesthetic Transfer of Care	https://mpog.org/files/quality/measures/TOC-02_spec.pdf
TOC 03 (MIPS 427): Post-Anesthetic Transfer of Care	https://mpog.org/files/quality/measures/TOC-03_spec.pdf
TRAN 01: Transfusion Management Vigilance	https://mpog.org/files/quality/measures/TRAN-01_spec.pdf
TRAN 02: Post Transfusion Monitoring	https://mpog.org/files/quality/measures/TRAN-02_Spec.pdf
OPIOID: Opioid Equivalency	https://mpog.org/files/quality/measures/Opioid_Spec.pdf

Secondary outcome(s), if applicable

1. The average monthly prevalence and magnitude of performance gaps for each measure
2. The average monthly prevalence and magnitude of performance trends for each measure

Data source

MPOG Provider dashboard database

Exposure Variable

Not applicable for exploratory analysis; future PCRC proposal to study type of feedback (tailored message, “one-size fits most” message, and existing standard message)

Covariates

Preconditions are necessary (but not sufficient) characteristics of data for the purpose of tailoring performance feedback. Preconditions are composed of one or more performance gaps or trends:

1. **Performance gap**: Any difference between a provider’s performance score and a comparator, such as a peer average or goal. Performance gaps can be positive or negative.
2. **Performance trend:** Consistently increasing or decreasing performance values across the 3 most recent time intervals. Performance trends can be positive or negative.

We propose to analyze the following preconditions that each represent a type of opportunity to send a tailored message to a provider (with example supported messages in parentheses):

1. **Positive gap** (e.g. “You are a top performer”, “Your institution's performance is above the benchmark”)
2. **Negative gap** (e.g. “You are below average this month”, “You did not meet the standard”)
3. **Positive trend** (e.g. “Your performance is improving”, “Your institution improved by 5% this month”)
4. **Negative trend** (e.g. “Your institution's performance has dropped”, “Your performance decreased by 5% this month”)
5. **Negative gap elimination**: Positive trend, negative gap at prior time interval (t-1), and absent negative gap at the current time interval (t) (e.g. “You have reached your goal”, “Your institution’s performance has met the standard”)

ASPIRE uses two comparators in the emailed feedback report for failed case review: a 90% threshold and a peer average. We will identify preconditions for both comparators.
The future PCRC proposal will adjust for provider type, clinic setting, and institution as covariates.

Statistical analysis

We will use R to calculate descriptive statistics for all outcomes. We will calculate the mean and median sum of providers who are eligible for tailored messaging, and generate histograms to identify the distributions and assess changes in this sum and its distribution over time.

For the subsequent trial, we plan to compare the standardized effect size for each arm on provider performance across all ASPIRE measures, clustered by institution.

Sensitivity / Secondary subgroup / Secondary outcome analyses

For the subsequent trial, we intend to conduct sensitivity analyses for a) the specific measures that were included in the content of tailored messages and b) the measures with displayed gaps in one-size-fits-most and standard messages. We will conduct a secondary subgroup analysis for the providers who logged into their dashboard each month during their participation in each intervention arm.

Power analysis

This exploratory proposal will be used to provide a power analysis for a subsequent full PCRC proposal.

Handling of missing or invalid data

We will assess the percentage of missing or invalid data in the exploratory analysis to develop an expectation for data quality problems in the subsequent trial. The message tailoring process is based on an assumption that healthcare organizations are complex adaptive systems in which data is periodically missing, software is updated, and providers turnover, etc. The exploratory analysis will account for these routine changes to develop expectations for disruptions, to assess the risk of these disruptions impeding the message tailoring process, which we anticipate will be relatively low across all ASPIRE measures each month.

Preliminary Single Center Data

We will first analyze single center (UM) date, to understand limitations of dataset and to inform multicenter analysis

Areas for discussion/known limitations

In any given month, all ASPIRE measures will not be available for all providers due to data issues at participating sites. See section “Handling of missing or invalid data”.

1. What is the effect of “missing” emails due to data extraction issues at sites?

2. What is effect of sites having different measures on their emails?
3. What is the effect of measures with high performance/ low variability in performance?
4. What is the effect of changes in format over time (adding previous months performance, for example)?

References

1. Friedman C, Rubin J, Brown J, Buntin M, Corn M, Etheredge L, et al. Toward a science of learning systems: a research agenda for the high-functioning Learning Health System. J Am Med Inform Assoc JAMIA. 2015 Jan;22(1):43–50.
2. Adler-Milstein J, Nong P, Friedman CP. Preparing healthcare delivery organizations for managing computable knowledge. Learn Health Syst. 0(0):e10070.
3. Medicine I of, America C on the LHCS in. Best Care at Lower Cost: The Path to Continuously Learning Health Care in America. National Academies Press; 2013. 437 p.
4. Shanafelt TD, Dyrbye LN, Sinsky C, Hasan O, Satele D, Sloan J, et al. Relationship Between Clerical Burden and Characteristics of the Electronic Environment With Physician Burnout and Professional Satisfaction. Mayo Clin Proc. 2016 Jul 1;91(7):836–48.
5. West CP, Dyrbye LN, Shanafelt TD. Physician burnout: contributors, consequences and solutions. J Intern Med. 2018;283(6):516–29.
6. Tuti T, Nzinga J, Njoroge M, Brown B, Peek N, English M, et al. A systematic review of electronic audit and feedback: intervention effectiveness and use of behaviour change theory. Implement Sci. 2017;12:61.
7. Dowding D, Randell R, Gardner P, Fitzpatrick G, Dykes P, Favela J, et al. Dashboards for improving patient care: Review of the literature. Int J Med Inf. 2015 Feb 1;84(2):87–100.
8. Brown B, Gude WT, Blakeman T, van der Veer SN, Ivers N, Francis JJ, et al. Clinical Performance Feedback Intervention Theory (CP-FIT): a new theory for designing, implementing, and evaluating feedback in health care based on a systematic review and meta-synthesis of qualitative research. Implement Sci. 2019 Apr 26;14(1):40.
9. Ivers NM, Grimshaw JM, Jamtvedt G, Flottorp S, O'Brien MA, French SD, et al. Growing literature, stagnant science? Systematic review, meta-regression and cumulative analysis of audit and feedback interventions in health care. J Gen Intern Med. 2014 Nov;29(11):1534–41.
10. Ivers NM, Sales A, Colquhoun H, Michie S, Foy R, Francis JJ, et al. No more "business as usual" with audit and feedback interventions: towards an agenda for a reinvigorated intervention. Implement Sci. 2014 Jan 17;9(1):14.
11. Lustria MLA, Noar SM, Cortese J, Stee SKV, Glueckauf RL, Lee J. A Meta-Analysis of Web-Delivered Tailored Health Behavior Change Interventions. J Health Commun. 2013 Sep 1;18(9):1039–69.
12. Krebs P, Prochaska JO, Rossi JS. DEFINING WHAT WORKS IN TAILORING: A META-ANALYSIS OF COMPUTERTAILORED INTERVENTIONS FOR HEALTH BEHAVIOR CHANGE. Prev Med. 2010;51(3–4):214–21.
13. Hawkins RP, Kreuter M, Resnicow K, Fishbein M, Dijkstra A. Understanding tailoring in communicating about health. Health Educ Res. 2008 Jun;23(3):454–66.
14. Lewis CC, Klasnja P, Powell BJ, Lyon AR, Tuzzio L, Jones S, et al. From Classification to Causality: Advancing Understanding of Mechanisms of Change in Implementation Science. Front Public Health [Internet]. 2018 [cited 2018 Sep 8];6. Available from: https://www.frontiersin.org/articles/10.3389/fpubh.2018.00136/full

Query Specification

https://docs.google.com/spreadsheets/d/12TIMeFXtu8tu7iwzfmfeixVdtxh87SfzWuRdEBZcsow/edit?usp=sharing

Appendix 1

To: Zachary Landis-Lewis

Cc:

Zachary Landis-Lewis

Astrid Fishstrom

Colin Gross

Nirav Shah

Subject: eResearch System-Generated Notice of “Not Regulated” Status for HUM00163880

SUBMISSION INFORMATION

Title: Exploratory analysis for tailored feedback messaging to MPOG Anesthesiology Providers

Full Study Title (if applicable): Exploratory analysis of preconditions for tailoring performance feedback to MPOG Anesthesiology Providers

Study eResearch ID: HUM00163880

Date of this System-Generated Notice: 5/15/2019

IRB "NOT REGULATED" STATUS:

Based on the information provided, the proposed study does not fit the definition of human subjects research requiring IRB approval (per 45 CFR 46, 21 CFR 56) because in this case, it is activities or procedures rather than human subjects that are the object of the study.

Appendix 2

Revised **Standards for QUality Improvement Reporting Excellence** (**SQUIRE 2.0**) publication guidelines

Notes to authors
▸ The SQUIRE guidelines provide a framework for reporting new knowledge about how to improve healthcare. ▸ The SQUIRE guidelines are intended for reports that describe system level work to improve the quality, safety and value of healthcare, and used methods to establish that observed outcomes were due to the intervention(s). ▸ A range of approaches exists for improving healthcare. SQUIRE may be adapted for reporting any of these. ▸ Authors should consider every SQUIRE item, but it may be inappropriate or unnecessary to include every SQUIRE element in a particular manuscript. ▸ The SQUIRE glossary contains definitions of many of the key words in SQUIRE. ▸ The explanation and elaboration document provides specific examples of well-written SQUIRE items and an in-depth explanation of each item. ▸ Please cite SQUIRE when it is used to write a manuscript.

Text section and item name	**Page/line no(s). info is located**
Title and abstract	
1. **Title**	
Indicate that the manuscript concerns an initiative to improve healthcare (broadly defined to include the quality, safety, effectiveness, patient-centredness, timeliness, cost, efficiency and equity of healthcare).	N/A
2. **Abstract**	
a. Provide adequate information to aid in searching and indexing.	(N/A for PCRC)
b. Summarise all key information from various sections of the text using the abstract format of the intended publication or a structured summary such as: background, local problem, methods, interventions, results, conclusions.	(N/A for PCRC)
Introduction: Why did you start?	
3. **Problem description** - Nature and significance of the local problem.	2
4. **Available knowledge** - Summary of what is currently known about the problem, including relevant previous studies.	2
5. **Rationale** - Informal or formal frameworks, models, concepts and/or theories used to explain the problem, any reasons or assumptions that were used to develop the intervention(s) and reasons why the intervention(s) was expected to work	2
6. **Specific aims** - Purpose of the project and of this report.	2
Methods: What did you do?	

7. **Context** - Contextual elements considered important at the outset of introducing the intervention(s).	N/A
8. **Intervention(s)**	
a. Description of the intervention(s) in sufficient detail that others could reproduce it.	N/A
b. Specifics of the team involved in the work.	N/A
9. **Study of the intervention(s)**	
a. Approach chosen for assessing the impact of the intervention(s).	N/A
b. Approach used to establish whether the observed outcomes were due to the intervention(s).	N/A
10. **Measures**	
a. Measures chosen for studying processes and outcomes of the intervention(s), including rationale for choosing them, their operational definitions and their validity and reliability.	3-4
b. Description of the approach to the ongoing assessment of contextual elements that contributed to the success, failure, efficiency and cost.	N/A
c. Methods employed for assessing completeness and accuracy of data.	N/A
11. **Analysis**	
a. Qualitative and quantitative methods used to draw inferences from the data.	4
b. Methods for understanding variation within the data, including the effects of time as a variable.	4
12. **Ethical considerations** - Ethical aspects of implementing and studying the intervention(s) and how they were addressed, including, but not limited to, formal ethics review and potential conflict(s) of interest.	3
Results: What did you find?	
13. **Results**	
a. Initial steps of the intervention(s) and their evolution over time (eg, time-line diagram, flow chart or table), including modifications made to the intervention during the project.	(N/A for PCRC)
b. Details of the process measures and outcomes.	(N/A for PCRC)
c. Contextual elements that interacted with the intervention(s).	(N/A for PCRC)
d. Observed associations between outcomes, interventions and relevant contextual elements.	(N/A for PCRC)
e. Unintended consequences such as unexpected benefits, problems, failures or costs associated with the intervention(s).	(N/A for PCRC)
f. Details about missing data.	(N/A for PCRC)
Discussion: What does it mean?	
14. **Summary**	
a. Key findings, including relevance to the rationale and specific aims.	(N/A for PCRC)
b. Particular strengths of the project.	(N/A for PCRC)
15. **Interpretation**	
a. Nature of the association between the intervention(s) and the outcomes.	(N/A for PCRC)
b. Comparison of results with findings from other publications.	(N/A for PCRC)
c. Impact of the project on people and systems.	(N/A for PCRC)

d. Reasons for any differences between observed and anticipated outcomes, including the influence of context.	(N/A for PCRC)
e. Costs and strategic trade-offs, including opportunity costs.	(N/A for PCRC)
16. **Limitations**	
a. Limits to the generalisability of the work.	N/A
b. Factors that might have limited internal validity such as confounding, bias or imprecision in the design, methods, measurement or analysis.	N/A
c. Efforts made to minimise and adjust for limitations.	N/A
Conclusions	
a. Usefulness of the work.	(N/A for PCRC)
b. Sustainability.	(N/A for PCRC)
c. Potential for spread to other contexts.	(N/A for PCRC)
d. Implications for practice and for further study in the field.	(N/A for PCRC)
e. Suggested next steps.	(N/A for PCRC)
Other information	
18. **Funding** - Sources of funding that supported this work. Role, if any, of the funding organisation in the design, implementation, interpretation and reporting.	N/A

Variable	Comment
Institution ID	anonymized institution
AIMS Staff ID	Transform
Staff Type	CRNA, Resident, Attending
Measure Name	Included Measures
Month	Month/Year
Number of Cases	Integer
Passed Count	Integer
Failed Count	Integer
Excluded Count	Integer
Pass Percentage	Pass/ Pass + Failed
Threshold	90-95% depending on measure
Same role average performance	Calculated on a per insitituion per role per month basis

Measures	Threshold (%)	Measure Type
AKI 01: Acute Kidney Injury	90	Outcome
BP 01: Low MAP Prevention	90	Process
BP 02: Avoiding Monitoring Gaps	90	Process
CARD 02: Avoiding MI- Troponin I <0.6	95	Outcome
FLUID -01-C: Minimizing Colloid Use- Cardiac	n/a	Process
FLUID 01-NC: Minimizing Colloid Use- Non-Cardiac	n/a	Process
GLU 01: High Glucose Treated	90	process
GLU 02: Low Glucose Treated	90	process
MED 01: Avoiding Medication Overdose	90	Outcome
NMB 01: Train of Four Taken	90	process
NMB 02: Reversal Administered	90	process
PONV 01: Post Operative Nausea and Vomiting	90	process
PONV 02: Post Operative Nausea and Vomiting- Pediatrics	90	process
PUL 01: Low Tidal Volume- Less than 10 ml/kg	90	process
PUL 02: Low Tidal Volume- Less than or equal to 8 ml/kg	90	process
PUL 03: PEEP Utilization	n/a	process
TEMP 01: Active Warming	90	process
TEMP 02: Core Temperature Measurement	90	process
TEMP 03 (MIPS 424): Perioperative Temperature Management	90	Outcome
TOC 01: Intraoperative Transfer of Care	90	process
TOC 02 (MIPS 426): Post-Anesthetic Transfer of Care	90	process
TOC 03 (MIPS 427): Post-Anesthetic Transfer of Care	90	process
TRAN 01: Transfusion Management Vigilance	90	outcome
TRAN 02: Post Transfusion Monitoring	90	outcome
OPIOID: Opioid Equivalency	n/a	process