

The Population Health Information Technology Assessment (PHITA): Understanding the Ability of Primary Care Practices to Report Clinical Quality Metrics

Jeff Hummel¹, Ellen O'Meara², Laura-Mae Baldwin³, David Dorr⁴, Lyle Fagnan⁴, Ross Howell⁵, Leah Tuzzie², Kilian Kimbel², and Michael Parchman²

¹Comagine Health

²Kaiser Permanente Washington Health Research Institute

³University of Washington

⁴Oregon Health & Science University

⁵Fraser Health

May 5, 2020

Abstract

Rationale Most small-to-medium sized practices lack the software tools and analytic skills required for clinical quality reporting. We describe the development and initial testing of a measure to rapidly assess practices' clinical reporting readiness and guide technical assistance for population health. **Methods** Co-investigators developed the Population Health Information Technology Assessment (PHITA), a 5-point scale comprised of two 3-point sub-scales measuring Software Capability and HIT Skill Set. A practice's PHITA score was determined by interviewing practice facilitators (PF) who coached practices in a regional quality improvement (QI) study. Relative risk regression models were used to estimate the association between each practice's PHITA score and its ability to report two or more (of four) cardiovascular risk clinical quality measures (CQMs). A qualitative analysis of PFs' field notes on high and low PHITA scoring practices was used to describe differences in practices' HIT experiences. **Results** Each point increase in total PHITA score was associated with a 29% higher probability of reporting two or more CQMs. Only 21.4% of practices were found to have the highest score on both sub-scales. Independently owned sites had significantly lower PHITA scores than other ownership types. Qualitative analysis for low PHITA scoring practices revealed reporting challenges and mistrust of data but willingness to try improving quality. High PHITA scoring sites consistently expressed on-going need for assistance, a focus on data accuracy, and greater engagement in quality improvement. **Conclusion** The PHITA can help PFs quickly assess preparedness for clinical quality reporting in small-medium sized practices and guide coaching efforts.

Abbreviations:

ABCS: Aspirin, Blood Pressure, Cholesterol, Smoking AHRQ: Agency for Healthcare Research and Quality CQM: Clinical Quality Measure EHR: Electronic Health Record ESCALATES: Evaluating System Change to Advance Learning and Take Evidence to Scale FQHC: Federally Qualified Health Center H2N: Healthy Hearts Northwest HIT: Health Information Technology HITECH Act: The Health Information Technology for Economic and Clinical Health Act ICD: International Classification of Diseases ORPRN: Oregon Rural Practice-based Research Network PF: Practice Facilitator PHITA: Population Health Information Technology Assessment QI: Quality Improvement TA: Technical Assistance RUCA: Rural Urban Community Areas

INTRODUCTION

Clinical quality reporting burdens small practices with estimated costs of \$50,000 per provider per year 1. Practice burdens include data entry, interference with patient interactions, and lack of interoperability 2, 3. Basic EHR reporting functionality that meets requirements for Meaningful Use certification are often of limited utility in efforts to improve quality metrics in high risk populations 4. Moreover, it is not clear that smaller primary care practices have the ability to generate the types of quality performance reports that are essential to inform efforts to improve the quality of care they deliver 5, 6.

In 2015 the Agency for Healthcare Research and Quality (AHRQ) launched EvidenceNOW, a national project offering practice facilitation support to more than 1500 primary care practices with ten or fewer clinicians to improve the cardiovascular risk factors Aspirin, Blood Pressure, Cholesterol and Smoking (ABCS) 7. AHRQ funded seven EvidenceNOW cooperatives, including the Healthy Hearts Northwest (H2N) cooperative, to understand the technical assistance (TA) that small to medium sized primary care practices need to improve performance on the ABCS clinical quality measures (CQMs).

In 2016 the EvidenceNOW evaluation team documented challenges participating clinicians encountered in their efforts to produce clinical quality reports for the project 5. They found that although the software performed well for billing data, functionality for population quality improvement activities was limited. Earlier observation by practice facilitators (PFs) 8 pointed to the importance of the ability of EHR software to extract, aggregate, and format data so clinicians can see care gaps 9. In addition, a practice must have personnel with analytic and reporting skills to use the software 5. Weakness in either factor has the potential to impair a clinicians' ability to see, understand, and act on quality metrics on which practice revenue increasingly depends 10.

A short and easy-to-use instrument to quickly assess a practice's ability to produce reliable practice-facing reports for quality improvement (QI) will benefit practice facilitators (PFs) who provide external support to a primary care practices. Here we describe a tool developed by the H2N study team for this purpose. We examine the relationships between a practice's PHITA score, its ability to report CQMs, and practice characteristics. Finally, we describe themes that emerged from an analysis of qualitative data derived from notes kept by PFs for practices with the lowest and highest PHITA scores.

METHODS

Study, Setting and Subjects :

H2N enrolled 259 practices in Washington, Oregon, and Idaho to test the effectiveness of external practice support to build QI capacity with a focus on controlling cardiovascular disease risk factors. Of the enrolled practices, 209 were randomized at the start of the intervention. Qualis Health and the Oregon Rural Practice-based Research Network (ORPRN) provided PFs for the effort directed by the MacColl Center for Health Care Innovation at Kaiser Permanente Washington Health Research Institute. Details about the study protocol and outcomes of the study have been previously reported 11, 12.

Data Collection and Measures

Quantitative

Study co-investigators from Qualis Health (JH, RH) created the Population Health Information Technology Assessment (PHITA) based on PF observations in prior work 13. The PHITA scale assesses 1) the technical capability of software, and 2) the available analytic skill set of the staff. Software Capability refers to data capture and analytic functionality in the EHR or affiliated software. Health Information Technology (HIT) Skill Set refers to the data management, query writing and analytic skills available to the practice. The categories were converted to 3-point scales in which 1 represented the lowest preparedness level and 3 the highest. See Table 1 for definitions. To assist the facilitator in prioritizing their work, the assessment describes three levels of preparedness (low, medium, and high) for each of these two sub-scales.

During the second half of H2N’s 15-month TA intervention (July - December 2016), two of the authors (JH, RH) administered the PHITA by interviewing all fifteen PFs so that each practice received sub-scale and total PHITA score.

We compared each practice’s PHITA score with its ability to submit data for ABCS CQMs. Ability to submit data was scored on a 5-point scale, from 0 to 4, defined by the number of ABCS CQMs submitted during the second quarter of 2017. H2N required a rolling 12-month lookback and excluded patients from the denominator with no office visit within 12 months. Criteria for each CQM are described in the primary outcomes paper 12. For the analyses we created a dichotomous outcome variable for whether a practice was able to report two or more (vs fewer than two) ABCS measures to improve interpretability of the findings.

To examine the relationship between PHITA scores and practice characteristics, we classified clinics geographically as rural or urban (RUCA codes 4-10 vs 1-3) 14, by practice size (solo clinician, 2-4 clinicians, or [?] 5 clinicians), and ownership (independent, part of a hospital or health system, Federally Qualified Health Center (FQHC)/migrant health clinic, or Tribal clinic).

Qualitative:

PFs documented all contacts with the 259 practices over the 15-month intervention and at 18 and 21 months after a baseline visit 12 using a password-protected online-database. Documentation included closed-ended and free-text entries after each in-person visit, phone call, or email encounter. In total, 4128 contacts were included in the database along with free-text documentation per contact from 1-5 pages in length. Qualitative data included PF notes about the practices’ QI and HIT progress and challenges, descriptions of rapid process improvement cycles, and general comments. This qualitative subanalysis only included practices that were scored as low or high PHITA scores.

Analysis :

Quantitative

We used Chi-square tests to determine whether scores for the PHITA subscales differed by practice characteristics. Spearman correlation coefficients were used to assess relationships between the two PHITA sub-scale scores. We used relative risk regression models to estimate the association between the PHITA score (by domain and the total score) and the ability of practice sites to report two or more CQMs for the 12-month period ending within the 2nd quarter of 2016. Preliminary analyses indicated that independently owned practice sites scored significantly lower than other ownership categories on both sub-scales. We therefore stratified the analysis of the relationship between a practice’s total PHITA score and its ability to produce two or more CQM reports by independent vs non-independent ownership.

Qualitative

Members of the study team with qualitative research training and expertise reviewed and coded all free text data for all contacts. An inductive approach was used to develop the code list 15 and an iterative process of coding, comparing, and discussing coding decisions was used to refine the code list and increase shared understanding of the data and codes. Each coder was assigned practices to code as sole coder. The coding team met regularly to discuss coding questions. All coded data were entered into Atlas.ti.16. To identify subthemes and synthesize the data, two qualitative analysts (LT, KK) reviewed the coded data and created a coding memo with a summary of the themes. The analysts met to discuss the themes and used the final iteration to structure the qualitative findings. Our methods for analyzing this rich qualitative set of data were consistent with the consolidated criteria for reporting qualitative research (COREQ) checklist for interview and focus groups 17.

KPWHRI’s Institutional Review Board approved this study.

Results

Quantitative

PFs completed the PHITA on all 201 practices actively participating during the 15-month period of coaching. There were no significant differences in PHITA scores between urban and rural sites, or by practice size (Table 2). However, independently owned practices had significantly lower scores on both sub-scales compared to other ownership types.

Most practices had PHITA scores in the lower range of possible scores. Only 66 practices (32.8%) had a high score for EHR analytic capability, while 60 practices (29.9%) had reliable and stable HIT analytic skills (Table 3); only 43 practices (21.4%) had both (Table 3). A higher total PHITA score was associated with increased ability to report ABCS CQMs. In all, 72.2% of practices with the highest PHITA total score were able to report two or more CQMs compared to only 20.0% of practices with the lowest score (Table 3). Each one-point increase in PHITA score was associated with a 29% higher probability of reporting two or more CQMs, (OR 1.29, 95% C.I 1.17 – 1.42) (Figure 1).

Both sub-scores were significantly associated with CQM reporting ability (Table 3) and with each other (correlation coefficient 0.62, $p < 0.005$). The association between the PHITA sub-scales and ability of the site to report CQM measures did not change when adjusted for practice ownership (data not shown).

Qualitative

Qualitative analysis of PF contact notes from practices scoring the lowest on both Software Capability and HIT Skill Set ($n = 40$) revealed three dominant themes: reporting challenges, a mistrust of their data, and a willingness to try to make improvements. (Table 4)

1. *Reporting challenges.* Limited staffing resources, access to data, and EHR capabilities presented challenges. Practices described inability to report, lack of HIT expertise or staff with only basic skills, and limited time to run reports. Other barriers included no interface between practice management and clinical data and difficulty mapping diagnostic codes to the problem list. Variations in data entry and data quality were often significant. For some practices a parent health system, network, or hospital controlled and limited access to the data. Limited reporting ability often reduced enthusiasm for engaging in QI.
2. *Mistrust of data :* Providers and staff seemed to mistrust or lack confidence in their data accuracy and EHR, which undermined motivation to use data for QI.
3. *Willingness to try :* Despite challenges, practices scoring low on the PHITA seemed eager to make improvements using rapid process improvement cycles and alternative data sources for QI. Some purchased registries, upgraded their EHRs, or transitioned to EHRs that promised better capabilities or a registry function.

Examination of PF contact notes for practices scoring highest on both sub-scales ($n = 43$) revealed three themes: an ongoing need for assistance, a focus on data accuracy, and engagement with QI activities. (Table 5)

1. *On-going need for assistance :* High performers had resources and skills to run reports, but faced remaining challenges and needed assistance from their PF, EHR vendor, or parent organization. Challenges were often related to non-familiarity with the complexities of quality reporting, competing priorities, and staff turnover. Specific barriers included limited date range reporting and inability to produce data at the provider or practice level.
2. *Focus on data accuracy :* High-scoring practices focused on improving data accuracy through better data capture and report validation. Improved data quality was viewed as a prerequisite to using it for quality improvement. Understanding information flow for reporting encouraged care teams to standardize data capture to improve reporting accuracy. Quality reporting methodology was new to some providers, including data definitions for clinical concepts and understanding population metrics.
3. *Engagement with QI :* High-scoring practices were engaged in understanding how reports were created and were willing to use them to guide patient care. They engaged their PF, their EHR vendor, and used other resources. These practices engaged all staff, including providers and leadership in using data for practice improvement.

DISCUSSION

The PHITA tool was designed to help PFs rapidly categorize practices into low, medium, or high levels of readiness for clinical reporting so that they may provide appropriate TA. Both technical capabilities and the presence of staff with the requisite knowledge and skills, separately and combined, are associated with the ability of a practice to report on clinical quality measures for cardiovascular risk factors. Further, qualitative data extracted from PFs' site visit documentation describing successes and challenges of practices with low and high PHITA scores add support for the validity of the scores. Least prepared sites struggled with data reporting challenges and a distrust of data that impeded their ability to do QI work, while best prepared sites put energy into improving data accuracy and engaging in QI activities. Both groups showed commitment to engaging in QI work and a need for on-going technical assistance.

A high PHITA score did not guarantee an ability to report CQMs. Practices with high scores on both sub-scales were generally those with centralized health IT resources, but these shops were often inundated with competing reporting priorities. In such settings, availability of tools and skills did not translate to producing additional reports unless the practice or health system leadership was willing to make reporting for H2N a strategic priority.

A few limitations deserve mention. The study was required to use the newly revised 2015 cholesterol guidelines for which canned reports were only available to a few participating sites at the time, limiting the ability of many practices to report on this measure. PHITA scores were assigned by interviewing the PFs rather than individuals in the practices. It is unknown how closely practice personnel would have agreed with the PF's assessment, or with each other. PHITA scores were assigned after PFs had 7 – 8 months of interaction to understand the health IT environment of each practice, whereas use of the PHITA as a field tool would entail an assessment based on more limited observation. Although few practices experienced improvements in software capability or skill set during that short time period, however, technical assistance provided by the PFs may have helped identify work-arounds that would improve the ability to produce two or more CQM reports. Error introduced by this would likely reduce the difference in reporting ability between levels of preparedness for both the sub-scales.

There is growing evidence for the importance of practice facilitation to support implementation of evidence into primary care practices 12, 18, 19. A significant portion of a PF's effort must be directed toward using available HIT tools effectively. The PHITA can help PFs set realistic expectations for data reporting and quickly identify strategies to meet reporting/analytic needs. For example, practices with limited reporting capability might create EHR patient lists and export them to spreadsheets where data can be manipulated to produce reports to identify patients with care gaps. In practices with limited reporting skills, a PF may be able to facilitate direct support from the vendor or help select a third-party registry.

Finally, the low levels of HIT capability and HIT analytic skills found among small-to-medium sized enrolled practices have serious implications for primary care infrastructure in the U.S. Barely over 20% of practices engaged in the H2N study had the necessary reporting and analytic skills for this work, and independent small practices were the least prepared. These findings are consistent with those of Cohen and colleagues 5 who reported challenges due to lack of functionality for generating reports, discordance between clinical guidelines and measures, questionable data quality, and unreceptive vendors. Our findings expand on prior work by documenting the lack of individuals with the software and analytic skills to write and validate reports. It is clear that clinicians need more than a meaningful use certified EHR to make significant progress in improving clinical outcomes as required for value-based reimbursement.

CONCLUSION

The PHITA offers a way to quickly assess the ability of clinical practices to report CQMs. Substantial proportions of primary care practices are inadequately prepared for the reporting and analytic requirements of value-based reimbursement. The gap is greatest in independent practices, which are also those least likely to have resources to pay for technical assistance. If these practices are to successfully transition to value-based reimbursement, stable funding for TA will need to be an integral part of national health policy. TA can target

the elements of Health IT preparedness identified by the PHITA, and help small-to-medium sized practices identify and overcome the barriers they face meeting the requirements for value-based reimbursement.

References:

1. Casalino LP, Gans D, Weber R, et al. US Physician Practices Spend More Than \$15.4 Billion Annually To Report Quality Measures. *Health Affairs*. 2016;35(3):1-6.
2. Meigs SL, Solomon M. Electronic Health Record Use a Bitter Pill for Many Physicians. *Perspectives in Health Information Management*. 2016;(Winter):1-17.
3. Arndt B, Beasley JW, Watkinson MD, et al. Tethered to the EHR: Primary Care Physician Workload Assessment Using EHR Event Log Data and Time-Motion Observations. *Annals of Family Medicine*. 2017;15(5):419-426.
4. Cohen DJ, Dorr DA, Knierem K, Dubard CA, Hemler JR, Hall JD, Marino M, Solberg LI, McConnell KJ, Nichols LM, Nease DE, Edwards ST, Wu WY, Pham-Singer H, Ko AN, Phillips RL, Rasmussen LV, Duffy FD, Balasubramanian BA. Primary Care Practice's Abilities And Challenges In Using Electronic Health Record Data For Quality Improvement. *Health Affairs*. 2018;37(4):635-643.
5. Cohen DJ, Balasubramanian BA, Gordon L, et al. A national evaluation of a dissemination and implementation initiative to enhance primary care practice capacity and improve cardiovascular disease care: the ESCALATES study protocol. *Implementation Science*. 2016; 11(68): 1-13.
6. Balasubramanian BA, Marino M, Cohen DJ, et al. Use of Quality Improvement Strategies Among Small to Medium-Size US Primary Care Practices. *Ann Fam Med*. 2018;16:S35-43.
7. Buchholz L. EvidenceNOW Grants Focus on Heart Disease Care in Small Practices. *JAMA*. 2015;314(2):115. doi:10.1001/jama.2015.7036.
8. Qualis Oral Health Integration Implimentation Guide 2014. <http://www.qualishealth.org/sites/default/files/Guide-Oral-Health-Integration.pdf> viewed 10/27/2019.
9. Dorr DA, Cohen DJ, Alder-Milstein J. Data Driven Diffusion of Innovations: Successes and Challenges in 3 Large-Scale Innovative Delivery Models. *Health Aff*. 2018;37(2):257-265.
10. Friedberg MW, Chen PG, Simmons P, Sherry T, Mendel P, Raaen L, Ryan J, Orr P, Vargo C, Caralasare L, Botts C, Blake K. Effects of Health Care Payment Models on Physician Practice in the United States: Follow-Up Study. 2018; RAND Corporation, Santa Monica, California. https://www.rand.org/pubs/research_reports/RR2667.html (accessed 2019.09.22)
11. Parchman ML, Fagnan LJ, Dorr DA, et al. Study Protocol for "Healthy Hearts Northwest": A 2x2 randomized factorial trial to build quality improvement capacity in primary care. *Implementation Science*. 2016;11:138.
12. Parchman ML, Anderson ML, Dorr DA, Fagnan LJ, O'Meara ES, Tuzzio L, Penfold RB, Cook AJ, Hummel J5 Conway C, Cholan R, Baldwin LM. A randomized trial of external practice support to improve cardiovascular risk factors in primary care. *Ann Fam Med*. 2019;17(Suppl 1): S40-49.
13. Hummel J, Phillips KE. A Population Management Approach to Oral Health. *Journal of the California Dental Association*. 2016 Mar. 167-172.
14. Hart LG, Larson EH, Lishner DM. Rural definitions for health policy and research. *Am J Public Health*. 2005;95(7):1149-1155.
15. Crabtree BF, Miller WL. Chapter 10: Immersion/Crystallization. In: Crabtree BF, Miller WL, editors. *Doing qualitative research*. 2nd ed. Thousand Oaks, CA: Sage Publications; 1999. p 163-77.
16. Muhr T. ATLAS.ti. Berlin, Germany: ATLAS.ti Scientific Software Development GmbH; 1993.
17. Tong A, Sainsbury P, Craig J. Consolidated criteria for reporting qualitative research (COREQ): A 32-item checklist for interviews and focus groups. *Int J Qual Health Care* 2007 Dec;19(6):349-57. DOI: <https://doi.org/10.1093/intqhc/mzm042>
18. Nagykaldi Z, Mold JW, Robinson A, Niebauer L, Ford A. Practice facilitators and practice-based research networks. *J Am Board Fam Med*. 2006;19(5):506-10.
19. Berta W, Cranley L, Dearing JW, Doherty EJ, Squires JE, Estabrooks CA. Why (we think) facilitation works: insights from organizational learning theory. *Implement Sci*. 2015;10:141.

Tables

Table 1: Definition of each level of preparedness for the PHITA Sub-scales.

HIT Technical Capability: Given the analytic capability of the EHR as deployed within the clinical delivery system thus far
Available HIT skills: Given the staffing of the clinical site or delivery system thus far in the project, if the CEO/CMO were

Table 2: Practice Characteristics and PHITA Scores

	All Practices	PHITA Capability Score	PHITA Capability Score	PHITA Capability Score	PHITA Capability Score	PHITA Skill Set Score	PHITA Skill Set Score	PHITA Skill Set Score	PHITA Skill Set Score
Characteristic	N (%)	1 Row %	2 Row %	3 Row %	p*	1 Row %	2 Row %	3 Row %	p*
Overall	201 (100.0)	39.8	27.4	32.8		21.9	48.3	29.9	
Location	112 (55.7)	42.9	36.0	28.6	25.8	28.6	38.2	0.35	
Urban	91 (44.3)	42.9	36.0	28.6	25.8	28.6	38.2	0.35	
Rural									
Size (# providers)	38 (18.9)	42.1	44.1	28.9	30.1	28.9	25.8	0.16	
1 (solo)	93 (46.3)	32.9		22.9		44.3			
2	70 (34.8)								
to 4									
5 or more									
Ownership	91 (45.3)	53.8	29.5	34.1	23.1	12.1	47.4	<0.001**	
Independent	78 (38.8)	18.2	40.0	13.6	30.0	68.2	30.0		
Health System	22 (10.9)								
FQHC	10 (5.0)								
IHS/Tribal									

* P-value from chi-square test

** P-value from chi-square test of independent vs other ownership

Table 3. Relationship between PHITA Scores and Number of CQMs Reported with 12-month Lookback

		Number of CQMs Re-reported with 12-month Lookback	Number of CQMs Re-reported with 12-month Lookback	Number of CQMs Re-reported with 12-month Lookback	Number of CQMs Re-reported with 12-month Lookback	Number of CQMs Re-reported with 12-month Lookback	Number of CQMs Re-reported with 12-month Lookback	
PHITA Element Capability	PHITA Score	0 %	1 %	2 %	3 %	4 %	Total N (%)	p*
	1	62.5	6.2	10.0	13.8	7.5	80 (100.0)	<0.001

		Number of CQMs Re- ported with 12- month Lookback	Number of CQMs Re- ported with 12- month Lookback	Number of CQMs Re- ported with 12- month Lookback	Number of CQMs Re- ported with 12- month Lookback	Number of CQMs Re- ported with 12- month Lookback	Number of CQMs Re- ported with 12- month Lookback	
	2	36.4	5.5	14.5	32.7	10.9	55 (100.0)	
	3	15.2	12.1	9.1	43.9	19.7	66 (100.0)	
Skill set	1	65.9	11.4	9.1	11.4	2.3	44 (100.0)	0.001
	2	39.2	4.1	12.4	26.8	17.5	97 (100.0)	
	3	21.7	11.7	10.0	45.0	11.7	60 (100.0)	
Total N (%)		80 (39.8)	16 (8.0)	22 (10.9)	58 (28.9)	25 (12.4)	201 (100.0)	

* P-value from relative risk regression model of PHITA score and number of CQMs reported using 12-month lookback (2-4 vs 0-1).

Table 4 Examples of PF notes from practices scoring lowest on both sub-scales.

Dominant Themes	Practice Facilitator notes
On-going need for assistance	<i>Their IT department will be able to provide a good portion of the resources they want, but there will be some issues with more complex requests. . . it has been traditionally difficult to get support from [vendor] for these changes. – Practice #21203 They are hosted by [vendor], so don't have free reign on their reports, but have to request from [vendor], so what is available ad hoc is limited. – Practice #32288</i>
Focus on data accuracy	<i>It was great to see the team so interested in getting to the root of a measure. – Practice #21223 They seem to be getting held up with thinking their data must be perfect to move forward. – Practice #21132 They run their reports at the provider level and add up the patients reported in the denominator and numerator in order sum up at the practice level aggregate. But they were uncertain how the EHR defined their patient population either by visit or by the assigned primary care provider. - Practice #21203</i>

Dominant Themes	Practice Facilitator notes
Engagement with QI	<i>The doctor requested more information, he felt there was a black box around the measure and would like to know more on what the measure is generated from. – Practice #21223</i> <i>Focused improvement efforts led to positive data changes. – Practice #21004</i>

Table 5: Examples of PF notes from practices scoring highest on both sub-scales.

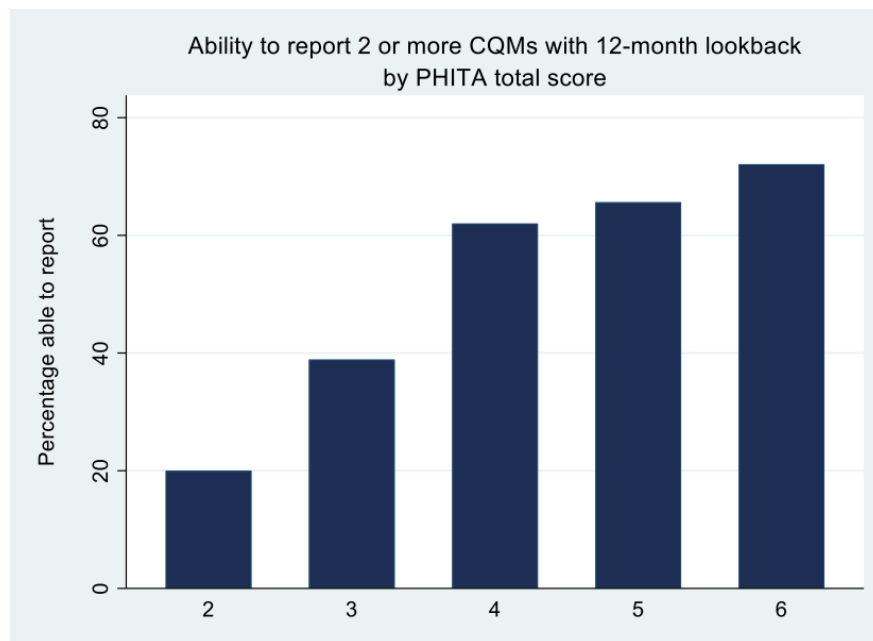


Figure 1. Ability to Report [?]2 CQMs with 12-month Lookback by Total PHITA Score

Hosted file

PHITA Figure.docx available at <https://authorea.com/users/287193/articles/411821-the-population-health-information-technology-assessment-phita-understanding-the-ability-of-primary-care-practices-to-report-clinical-quality-metrics>