A Practical Guide to Using the Positive Deviance Method in Health Services Research

Adam J. Rose and Megan B. McCullough

Objective. To provide practical tips for health services researchers considering the use of positive deviance (PD) methods to help explain variations in quality of care or other meaningful parameters.

Data Sources. Published literature and personal experience.

Study Design. Narrative review.

Principal Findings. This review includes a discussion of possible applications of PD to health services research, some methodological choices applicable to PD, and some brief tips regarding publishing the results and incorporating them into future interventions.

Conclusions. It is hoped that this article will help health services researchers to use this valuable research method more effectively, especially those who have not done so before.

Key Words. Positive deviance, qualitative research, implementation science, variations in care, quality of health care

RATIONALE FOR THIS REVIEW

Over the past 6 years, we have increasingly turned to positive deviance (PD) research methods to help accomplish our health services research (HSR) goals, such as explaining the causes of variations in health care or informing the design of future interventions to address underuse of an evidence-based practice. Beginning with one study in 2010 (Rose et al. 2012b), we have progressed to leading or participating in two other studies which have been published (Goren et al. 2016; Razouki et al. In Press) and three whose research is still ongoing. These experiences have led to a lot of thinking and also empirical learning about what works well in PD studies and what does not. The aim of this article is to share these practical lessons that we have learned, to

facilitate the effective use of PD by other health services researchers. The reader is referred elsewhere for the basics of qualitative research methods (Patton 2002); our aim is to explain how these techniques can be applied to PD research.

THE POSITIVE DEVIANCE METHOD—APPLICATIONS TO HEALTH SERVICES RESEARCH

The term “positive deviance” (PD) originated in the 1970s, when it was noted that even in villages where most children were malnourished, some children were not (Wishik and Vynckt 1976). The researchers used the successful habits of these exceptional children to promote behaviors that were within the reach of all. While previous authors have suggested that PD can rely on a mix of quantitative and qualitative methods (Bradley et al. 2009), we think that PD is best construed as a primarily qualitative research endeavor. The great attraction of the PD approach is that by definition, it will identify solutions that are currently feasible because someone is already employing them. PD has, therefore, been advanced as a way to solve problems, including those that have been long-standing or even intractable. PD has gained considerable momentum over the past two decades (Positive Deviance Initiative 2016), especially through its applications to HSR (Bradley et al. 2009).

At least several important conditions must be present to use the PD approach to design an HSR project. There must be variation on a measure of interest across sites or providers, but also the measure must be quantifiable and evidence based (Bradley et al. 2009). Previous applications of PD to HSR have focused on explaining variations in quality of health care, such as risk-adjusted mortality rates in acute myocardial infarction (Curry et al. 2011) or risk-adjusted rates of surgical complications (Daley et al. 1997). However, applications of PD to HSR are not limited to explaining variation on a

Address correspondence to Adam J. Rose, M.D., M.Sc., F.A.C.P., Section of General Internal Medicine, 801 Massachusetts Avenue, 2nd Floor, Boston, MA 02118; e-mail: adamrose@bu.edu. Adam J. Rose, M.D., M.Sc., F.A.C.P., is also with the Center for Healthcare Organization and Implementation Research, Bedford VA Medical Center, Bedford, MA. Megan B. McCullough, Ph.D., is with the Center for Healthcare Organization and Implementation Research, Bedford VA Medical Center, Bedford, MA; Department of Health Policy and Management, Boston University School of Public Health, Boston, MA.

performance measure. There are at least four dimensions of care that have the potential to be explored using PD:

- Variations in quality of health care, measured using a (risk-adjusted) outcome of care
- Variations in the cost or the value (cost compared to quality) of care
- Variations in utilization or uptake of expensive new medications or technologies
- Variations in the rate of appropriate versus inappropriate care

We are currently involved with projects that are using PD to assess variations in all these dimensions of health care. The choice of key informants and the questions one asks will differ according to these aims, but the basic process of using PD to examine variations in care is similar.

It is important to remember that sites of care can excel in some areas while performing poorly in others. For example, a site might excel at managing acute stroke, but it may have poor rates of postoperative infections. Or a site may achieve excellent outcomes for patients with acute myocardial infarction, but at a higher price per patient than the industry average. When sites present these sorts of mixed results, researchers must make conscious decisions about what characteristic they will use to profile sites, and sometimes they may need to profile on more than one dimension.

It is well known that variation on these dimensions of care (quality, value, spread of new technologies, and appropriateness) is ubiquitous (Wennberg and Gittelsohn 1973); in fact, it would be difficult to imagine a situation in which researchers would look for variation and not find any. In the past, HSR has been criticized for often pointing out such variations in care but not doing as much to address them. Part of the promise of PD is its potential to leverage naturally occurring variation to produce both scientific insight into the causes of variation and also potential pathways to helping all sites or providers perform more like those that are currently performing best.

PLACE IN THE RESEARCH “PIPELINE”

For the purposes of this article, we will assume that every research effort intends to move all the way from describing variation, to explaining its causes, to using that information to design an effective intervention to improve care, and then deploying that intervention. The PD approach should be considered as a possible route anytime one has identified variation on a dimension of care

that matters to the researchers, funders, health care leaders, or society. Our own work demonstrates how PD can fit into a broader program of research and implementation. We had established that anticoagulation control is an important intermediate outcome of care (Rose et al. 2009), that improving anticoagulation control would improve patient outcomes and save money (Rose et al. 2011a), and that sites of care can be profiled on anticoagulation control (Rose et al. 2011c). We then investigated which structures and processes of care predicted better performance on anticoagulation control at the site level. We did find several processes of care that were highly linked with anticoagulation control (Rose et al. 2011b, 2012a, 2013), but our efforts to explore the structure and organization of care using quantitative data were less than revealing (Rose et al. 2011d). The lessons we drew from this, which have been confirmed by our later work, are that surveys and databases may not reveal enough details or context to fully explain how differences in the organization and management of care contribute to differences in outcomes.

We, therefore, undertook a PD study, wherein we visited three of the sites (within the VA system) with the best performance on anticoagulation control and three of the sites with the worst performance. This PD study led to insights about what differentiated the high and low sites (Rose et al. 2012b), which in turn informed the intervention we eventually developed to improve the organization and management of anticoagulation care (McCullough et al. 2015). The main results of this implementation study are forthcoming. In this way, our group has firsthand experience of how PD research can contribute to a more effective effort to improve care, and we similarly expect to use our more recent PD projects as a basis for improving care.

METHODOLOGICAL CHOICES AND OTHER PRACTICAL TIPS

Qualitative versus Quantitative Methods

One of the strengths of PD is its reliance upon qualitative methods, which has advantages in terms of understanding context, meaning, and details that may not be available in database-type studies (Giacomini and Cook 2000). Also, qualitative research methods have the potential to produce unanticipated findings, which is much more difficult with quantitative methods (Denzin and Lincoln 1994; Patton 2002; Bernard 2011; Charmaz 2014). PD is an especially apt choice when quantitative approaches to explaining variations in care, such

as surveys or database studies, have not worked out well. Database studies can be very quick and effort-efficient, so some researchers will naturally gravitate toward them first, but in our experience, numbers alone cannot fully explain the causes of variability in health care. We would suggest that most efforts to understand the causes of variability in health care should begin with the expectation of using (qualitative dominant) mixed methods, with PD as the organizing principle for the qualitative part of the approach, and the addition of quantitative studies using databases, surveys, or similar mechanisms as appropriate.

Choosing Sites

In our studies, we have always included sites with the highest and the lowest rankings we could find on our measure. To enhance the credibility of between-site comparisons, we have generally risk-adjusted the performance measure, which has sometimes required the extra step of deriving and validating a specific risk adjustment model for the measure (Rose et al. 2010). One logical goal for the PD study, in fact, may be to examine what factors seem to be important determinants of performance but were not included in the risk adjustment model, because of either an oversight or in many cases unavailability of those variables in the data. This can lead not only to an improved understanding of the determinants of performance but also to an improved risk adjustment model. It is important to remember, however, that unlike pay-for-performance schemes, reimbursement is not riding on how well the risk adjustment model has captured determinants of performance, so it may be sufficient to be “mostly” confident in the results.

Another important issue relates to sample sizes and confidence in outlier status. We have generally focused on measures that occur frequently, such as continuous measures of performance or a process of care that happens with every patient. However, some outcomes are relatively rare, such as stroke or the formation of a pressure ulcer. Extra care should be taken when profiling sites on rare occurrences, and it may be necessary to apply advanced methods such as empirical Bayesian modeling to ensure that sites are truly outliers on performance (Berlowitz et al. 2002). It is also important to confirm that performance is consistent from year to year, as we have done in our past efforts, to ensure that one is observing actual variations in performance, rather than just statistical variation between years.

Some PD efforts have focused only on profiling the high outliers, but we have found it important to include low-outlier sites for contrast, and in fact, we

usually learn a tremendous amount from the low outliers about how *not* to organize care. We have found that it is important to maximize the *contrast* between these high and low sites. The idea of contrast is that some practices at a high-outlier site will look like “best practices” on the face of it, but some or all of the low-outlier sites will have the same practices, revealing these practices as “red herrings.” For example, in our anticoagulation study, we had expected to find that the staff at the best-performing sites would be distinguished by their willingness to go “above and beyond” for their patients, but in fact, we found that this was equally true at the high- and the low-outlier sites (Rose et al. 2012b), meaning that simply exhorting providers to go the extra mile for their patients would not be an effective approach. The point, then, is to find the practices that are reliably present at the high-outlier sites and reliably absent at the low-outlier sites—those are the high-value practices whose spread can help spur quality improvement (Rose et al. 2012b; Goren et al. 2016; Razouki et al. In Press). Our inclusion of low-outlier and high-outlier sites is an important methodological choice, and indeed, it makes the term “positive deviance” something of a misnomer because we are actually looking at both positive and negative deviance.

We have not included sites with intermediate performance in any of our studies, although grant reviewers have sometimes asked us to consider doing this. In our view, intermediate sites will not maximize contrast, but including them will increase the resources and time needed to perform the study. In addition, the goals of becoming more like the high-outlier sites, or less like the low-outlier sites, are not meaningfully advanced by also studying sites whose characteristics are presumably somewhere in between. However, a case could be made for specifically comparing the intermediate- and the high-performing sites, to focus on subtler differences in the organization of care that nevertheless contribute to excellent results. There may be limitations inherent in studying negative deviant sites, in that their flaws may be fairly obvious even to them, but may seem intractable because they simply have not been given the opportunity to address these obvious problems due to a lack of resources or other constraints. However, we have not found that the negative deviant sites have been aware of their poor performance or their deficiencies. On the contrary, informants from these sites have expressed a belief that their site is performing well above average, leading us to conclude that a lack of performance awareness (due to a lack of performance measurement) may contribute to low-outlier status (Rose et al. 2012b). While we continue to feel that comparing positive and negative deviant sites is the best and most efficient use of limited resources, it would

be useful in future studies to empirically examine whether the inclusion of intermediate sites adds meaningfully to what is learned.

Sometimes, the sites with the most extreme performance cannot participate or do not want to; one simply chooses the next site down on the list, and the erosion of contrast is minimal. Efforts should also be made to balance sites on other dimensions, such as urban/rural, large/small, or other relevant dimensions—as much as can be accommodated given the need to maximize contrast by focusing on the top and the bottom of the “rank list.” Extreme similarity within groups should be avoided as an obvious threat to validity. For example, a study where all the high-outlier sites are urban and all the low-outlier rural would not be acceptable.

How Many Sites to Sample

The question of how many sites to include is a challenging one, and one we have not yet fully solved. Ideally, a study would enroll sites prospectively until thematic saturation is achieved (Patton 2002), but often it is necessary to specify a number of sites in advance for a grant proposal and budget. Saturation is usually achieved within the first 12 interviews and often within the first six interviews (Guest, Bunce, and Johnson 2006). This provides a helpful guide to how many interviews would be necessary per site to learn about how the site organizes care, and indeed our experience (Rose et al. 2012b; Goren et al. 2016; Razouki et al. In Press) has been that 6–8 participants per site are usually sufficient to achieve saturation. However, for a PD study, one also needs to pursue saturation at the site level. According to the Guest study, it would be ideal to enroll 16 sites in each PD study—eight to fully explore the different ways to be a high-outlier site and eight to fully explore the different ways to be a low-outlier site. However, it is rarely practical to interview 6–8 participants at each of 16 sites, due to constraints of research funding, effort, and a need to produce results sooner rather than later. In our studies, we have tended to include something between three high and three low sites (“3 + 3”) and 5 + 5; although neither one seems to guarantee saturation, 5 + 5 at least seems likely to come close (Guest, Bunce, and Johnson 2006). That said, we have produced good results with as little as 3 + 3; although we may not have fully elucidated all possible pathways to high- or low-outlier status, we have certainly elucidated many of the most important ones (Rose et al. 2012b; Razouki et al. In Press). It is difficult to balance not only the substantive demands of the science but also the issues of how this appears to grant reviewers, as we have been variously told by reviewers that 3 + 3 is too few to be complete or that 5 + 5 is too

many to be feasible. More recently, we have used a hybrid approach, in which we briefly “screen” a larger number of sites (e.g., 5 + 5) by interviewing 1 or 2 key informants, and then select only some sites (e.g., 3 + 3) for further study, based on preliminary analysis of the initial interviews. Selection of sites for in-depth study should attempt to maximize thematic variation among sites, by not enrolling two sites for further study that appear similar in terms of how they achieved high or low results.

Choosing Key Informants

As in many qualitative studies, there is no single rule that can guide whom to use as key informants, but there are some general principles that can help in PD studies. PD studies often aim to uncover processes and structures of care, and therefore, the research subject selected in PD studies should be able to answer these types of questions. However, not every respondent will be able to answer all the questions: lower-level employees may not appreciate some of the higher-level policy decisions, while leadership figures may delegate some of the details to others. Therefore, it is important to include staff at various levels of the hierarchy (leaders, middle managers, frontline staff, and support staff). In many cases, the goal of constructing the most accurate view of how things work will also be advanced by including people from different disciplines (e.g., two primary care physicians, two specialist physicians, two nurses, and two pharmacists). As an example, in seeking to understand the differences between high- and low-performing anticoagulation clinics, we interviewed the Chief of Pharmacy (i.e., the supervisor), the front-line pharmacists who staff the clinic, the support staff who work in the clinic, and important collaborators and internal customers such as physicians who help run the clinic (Rose et al. 2012b). In contrast, for a study of sites with high and low rates of employee flu vaccination, we enrolled key figures from occupational health, infection control, employee union leaders, and hospital leaders (Razouki et al. In Press). It is generally important to develop several different versions of the interview guide, aimed at supervisors, front-line staff, support staff, and possibly other groups one expects to interview. We use what has been called a “snowball” sampling technique (Biernacki and Waldorf 1981; Patton 2002; Penrod et al. 2003), namely, that whenever a participant says that a certain person exerts a major influence on the process we are studying, we request an interview from that person. In this way, the study can remain open to

unexpected findings, including the discovery that an unexpected person has a major impact on the way the site provides care.

Telephone versus In-Person Site Visit

Travel for site visits can reveal much information that is of value, but it may not be necessary in every case. Travel is most important when there is a need to directly observe processes of care and to compare what people *say* they do with what they *actually* do when observed (Finkler et al. 1993; Angrosino 2007; Bernard 2011). Direct observation is labor intensive to conduct and to analyze, and time constraints may make this challenging to include. In many cases, we have found that it is not necessary to travel to sites to conduct PD research; telephone interviews have served well for many, although not all, of our projects. The methodological rigor of telephone interviews can be enhanced in PD by careful development and review of the interview guide, with iterative revisions and pilot testing (Burke and Miller 2001; Sturges and Hanrahan 2004). In any qualitative study, interim review of the first few interview transcripts should be performed to ensure that interviews are garnering the information that is sought. For PD studies, this would take the form of ensuring that enough detail is being collected about the structures and processes of care at the site, and how they may be contributing to outcomes. There is an additional concern that interviewees may be less forthcoming over the telephone, but we have not observed this to be the case. Phone interviewing also allows for flexibility in scheduling that would not be possible with site visits. PD studies in health care tend to explore topics that are not highly emotionally charged, so telephone interviews are likely to be sufficient. Travel for research has implications for research budgets and for one's personal life, and should only be done if it adds value.

The Issue of Blinding

Arguably, it would be ideal to blind interviewers and data analysts to a site's high- or low-outlier status, to avoid any sort of bias or selectivity in the way questions are asked or utterances analyzed. However, we have not blinded the researchers in any of our studies, and it does not seem to have compromised the research. In most cases, it may be hard to successfully blind the researchers. For example, in a previous study of 10 high- and 10 low-outlier sites on risk-adjusted surgical complication rates, blinded site visitors correctly predicted the high or low status of 17 of 20 sites ($p < .001$), calling into question whether attempts to

blind investigators are worth the effort (Daley et al. 1997). We would say that it is extremely important to be even-handed in data collection and analysis and to always explore what is good and bad (or high and low) at all of the sites. Even the best performers have weak points, and even the worst performers may have features worthy of emulation; it is important not to miss these because one is not looking for them (Rose et al. 2012b; Goren et al. 2016; Razouki et al. In Press). It is also important for interviewers to understand the underlying construct on which sites have been profiled and selected to participate, whether it is risk-adjusted mortality, cost per patient, or some other outcome. This will help ensure that interviewers will be able to ask appropriate follow-up questions to achieve a full understanding of how the organization and management of care at the site has contributed to the outcome of interest.

The other side of the issue of blinding is whether to reveal to the sites being studied that they are outliers on performance or some other metric. We have never shared this fact with sites, but have instead maintained that we are studying the range of how different sites manage a certain program, or use a certain medication, or the like. In this way, we are telling 99 percent of the truth, but omitting one small fact. We have explicitly received approval for this strategy from both grant reviewers and ethics reviews (i.e., the institutional review board). The alternative, telling sites that they are high or low outliers, seems likely to impair participants' ability to answer questions forthrightly and in a nondefensive manner.

HOW TO COLLECT AND ANALYZE DATA IN POSITIVE DEVIANCE STUDIES

In PD studies, one should try to triangulate different sources of data to gain a complete and nuanced understanding of how things work at a site (Kimchi, Polivka, and Stevenson 1991; Creswell and Clark 2011). While triangulation is common to all qualitative studies, in PD studies it serves a somewhat unique purpose, as it enables researchers to create a full portrait of the organization of care at each site, drawing on all available data sources. Key informant interviews are always the main data source, but they should be supplemented by review of important documents and direct observation, when applicable. It is our practice to always ask participants to send us a copy of any document they mention as important, such as a training manual, handouts given to patients, or note templates.

We use the semistructured interview approach, which is by far the most commonly used in HSR (Patton 2002). As with any qualitative study, it is important to ask open-ended questions, to allow the participant to speak without interrupting, and to follow up on statements that seem ambiguous or unclear because those follow-up questions often lead to the greatest insights. The interviews in PD research tend to be fairly unambiguous, in that participants say what they mean in a fairly straightforward way. It, therefore, makes sense to choose a fairly straightforward, unfussy approach to data analysis, to match the relatively circumscribed nature of the data and the project aims.

There are three main approaches to incorporating key documents alongside interview transcripts, field notes, and observation data. It is possible to code them in a similar manner to the interview transcripts, or one can analyze them through writing summaries of them and memos about them, using the memo feature of NVivo or similar functions in other software. In most cases, we have favored the memo approach (Miles, Huberman, and Saldana 2014). One can also use a matrix approach, which allows one to order processes and do a sequence analysis (Miles, Huberman, and Saldana 2014). In this way, the research team can organize process data into events or units and elucidate what the steps in a process are, when they happen, and how the different steps relate to each other.

We have found that it is important to combine top-down and bottom-up approaches to data analysis. For example, a top-down (theory-driven) approach might start with the Rogers Theory of Diffusion of Innovations to explain why some sites have adopted a new practice faster than others (Rogers 2003). Codes would be developed to encompass the key constructs of this organizing theory. The top-down analysis must be complemented by a bottom-up (emergent) analysis, whose purpose is to include concepts that may not be contained within the chosen theory. While the importance of the theory-based approach should not be minimized, in our experience it has been the emergent ideas, which are not necessarily contained within the preselected conceptual framework, that provide the most valuable insights.

PD studies are amenable to traditional sorts of in-depth qualitative analysis (e.g., thematic analysis, grounded theory approaches, and case study analysis) as discussed above. However, PD also can be an extremely practical choice for time-pressured circumstances when rapid qualitative or rapid ethnographic analysis is required (Manderson and Aaby 1992; Vlassoff and Tanner 1992; Utarini, Winkvist, and Peltó 2001).

PUBLICATION OF RESULTS

It is certainly important to publish the results of PD studies, both to advance our shared understanding of the underlying causes of variation in health care and to build credibility and support for interventions that will ultimately be based on the study findings. Fortunately, many journals have been willing to publish these sorts of studies, even journals that do not usually publish qualitative studies. In part, this may be because the aims and methods for PD studies are fairly intuitive. The choice of journal may be guided somewhat by whether one's major focus is to advance HSR (in which case a methods-oriented journal may be preferred) or to inform clinicians and leaders in the field under study (in which case a clinical journal may be preferred). We have seen examples of PD research published in both kinds of journals.

USING RESULTS TO INFORM INTERVENTIONS

As stated above, the main purpose of PD investigations like those described here is to inform effective interventions to improve care. The findings of PD research have many obvious advantages as a basis for interventions, foremost among them that the strong practices described will always be feasible, because they are already being done somewhere. The practice of designing and implementing effective interventions to improve care is beyond the scope of this article; the reader is referred to a large and growing literature in implementation science to learn more about what has been shown to work (Damschroder et al. 2009; Stetler et al. 2011). The reader should not take this challenge lightly, because many more implementation efforts fail to achieve their aims than succeed. The reader is advised to partner with a recognized expert in implementation science to help design and implement an intervention that will truly achieve its goals.

APPLICATION OF PD OUTSIDE THE VA SYSTEM

Our PD efforts have all occurred in a particular setting, namely the VA national health care system. VA is the nation's largest integrated health care system, although it is far from the only one. As an integrated health care system, there may be a degree of uniformity and connection among sites that may enhance PD efforts. For example, soliciting participation from sites and

providers may be somewhat easier within the context of a shared health care system. However, there are ways that a PD study could be even more powerful outside of an integrated health care system, because one would have the potential to find an even wider variation in how care is organized and delivered. Indeed, important PD studies have succeeded outside the VA as well (Curry et al. 2011).

SUMMARY

Positive deviance methods can be a useful bridge for health services researchers between observing performance variation, explaining its causes, and formulating effective solutions to address it. This article provides practical tips for how to conduct PD studies based on our past experience. In the absence of a formal literature about how to apply qualitative methods to doing PD studies in HSR, such wisdom has likely been acquired only by individuals and not shared widely. Here, we have shared what we learned so that others may benefit.

ACKNOWLEDGMENTS

Joint Acknowledgment/Disclosure Statement: Dr. McCullough receives general support (space, computer, etc.) from the Bedford VA Medical Center. The sponsor did not have any role in the research or in the drafting or approval of this manuscript. The views expressed in this article do not necessarily reflect the official policies of the U.S. Department of Veterans Affairs.

Disclosures: None.

Disclaimers: None.

REFERENCES

- Angrosino, M. V. 2007. *Naturalistic Observation*. Walnut Creek, CA: Left Coast Press.
- Berlowitz, D. R., C. L. Christiansen, G. H. Brandeis, A. S. Ash, B. Kader, J. N. Morris, and M. A. Moskowitz. 2002. "Profiling Nursing Homes Using Bayesian Hierarchical Modeling." *Journal of the American Geriatrics Society* 50 (6): 1126–30.
- Bernard, H. R. 2011. *Research Methods in Anthropology: Qualitative and Quantitative Approaches*. Lanham, MD: Altamira Press.

- Biernacki, P., and D. Waldorf. 1981. "Snowball Sampling: Problems and Techniques of Chain Referral Sampling." *Sociological Methods & Research* 10 (2): 141–63.
- Bradley, E. H., L. A. Curry, S. Ramanadhan, L. Rowe, I. M. Nembhard, and H. M. Krumholz. 2009. "Research in Action: Using Positive Deviance to Improve Quality of Health Care." *Implementation Science* 4: 25.
- Burke, L. A., and M. K. Miller. 2001. "Phone Interviewing as a Means of Data Collection: Lessons Learned and Practical Recommendations." *Forum: Qualitative Social Research*, 2(2): 7.
- Charmaz, K. 2014. *Constructing Grounded Theory*. Thousand Oaks, CA: Sage.
- Creswell, J. W., and V. L. P. Clark. 2011. *Designing and Conducting Mixed Methods Research*. Thousand Oaks, CA: Sage.
- Curry, L. A., E. Spatz, E. Cherlin, J. W. Thompson, D. Berg, H. H. Ting, C. Decker, H. M. Krumholz, and E. H. Bradley. 2011. "What Distinguishes Top-Performing Hospitals in Acute Myocardial Infarction Mortality Rates? A Qualitative Study." *Annals of Internal Medicine* 154 (6): 384–90.
- Daley, J., M. G. Forbes, G. J. Young, M. P. Charns, J. O. Gibbs, K. Hur, W. Henderson, and S. F. Khuri. 1997. "Validating Risk-Adjusted Surgical Outcomes: Site Visit Assessment of Process and Structure. National VA Surgical Risk Study." *Journal of the American College of Surgeons* 185 (4): 341–51.
- Damschroder, L. J., D. C. Aron, R. E. Keith, S. R. Kirsh, J. A. Alexander, and J. C. Lowery. 2009. "Fostering Implementation of Health Services Research Findings Into Practice: A Consolidated Framework for Advancing Implementation Science." *Implementation Science* 4: 50.
- Denzin, N. K., and Y. S. Lincoln. 1994. *Handbook of Qualitative Research*. Thousand Oaks, CA: Sage Publications Inc.
- Finkler, S. A., J. R. Knickman, G. Hendrickson, M. Lipkin Jr, and W. G. Thompson. 1993. "A Comparison of Work-Sampling and Time-and-Motion Techniques for Studies in Health Services Research." *Health Services Research* 28 (5): 577.
- Giacomini, M. K., and D. J. Cook. 2000. "Users' Guides to the Medical Literature: XXIII. Qualitative Research in Health Care B. What Are the Results and How Do They Help Me Care for My Patients? Evidence-Based Medicine Working Group." *Journal of the American Medical Association* 284 (4): 478–82.
- Goren, J. L., A. J. Rose, R. L. Engle, E. G. Smith, M. L. Christopher, N. M. Rickles, T. P. Semla, and M. B. McCullough. 2016. "Organizational Characteristics of High- and Low-Clozapine Utilization." *Psychiatric Services* 67 (11): 1189–96.
- Guest, G., A. Bunce, and L. Johnson. 2006. "How Many Interviews Are Enough? An Experiment with Data Saturation and Variability." *Field Methods* 18: 59.
- Kimchi, J., B. Polivka, and J. S. Stevenson. 1991. "Triangulation: Operational Definitions." *Nursing Research* 40 (6): 364–6.
- Manderson, L., and P. Aaby. 1992. "An Epidemic in the Field? Rapid Assessment Procedures and Health Research." *Social Science & Medicine* 35 (7): 839–50.
- McCullough, M. B., A. F. Chou, J. L. Solomon, B. A. Petrakis, B. Kim, A. M. Park, A. J. Benedict, A. B. Hamilton, and A. J. Rose. 2015. "The Interplay of Contextual Elements in Implementation: An Ethnographic Case Study." *BMC Health Services Research* 15 (1): 62.

- Miles, M. B., A. M. Huberman, and J. Saldana. 2014. *Qualitative Data Analysis: A Methods Sourcebook*. Thousand Oaks, CA: Sage.
- Patton, M. Q. 2002. *Qualitative Research and Evaluation Methods*. Thousand Oaks, CA: Sage.
- Penrod, J., D. B. Preston, R. E. Cain, and M. T. Starks. 2003. "A Discussion of Chain Referral as a Method of Sampling Hard-to-Reach Populations." *Journal of Transcultural Nursing* 14 (2): 100–7.
- Positive Deviance Initiative. 2016. "Positive Deviance Initiative: Projects" [accessed on March 11, 2016]. Available at www.positivedeviance.org/projects
- Razouki, Z. A., T. Knighton, R. A. Martinello, P. A. Hirsch, K. M. McPhaul, A. J. Rose, and M. B. McCullough. In Press. "Organizational Factors Associated with Health Care Provider Influenza Campaigns in the Veterans Health Care System: A Qualitative Study." *BMC Health Services Research*.
- Rogers, E. M. 2003. *Diffusion of Innovations*. New York: Free Press.
- Rose, A. J., D. R. Berlowitz, S. M. Frayne, and E. M. Hylek. 2009. "Measuring Quality of Oral Anticoagulation Care: Extending Quality Measurement to a New Field." *Joint Commission Journal on Quality and Patient Safety* 35 (3): 146–55.
- Rose, A. J., E. M. Hylek, A. Ozonoff, A. S. Ash, J. I. Reisman, and D. R. Berlowitz. 2010. "Patient Characteristics Associated with Oral Anticoagulation Control: Results of the Veterans Affairs Study to Improve Anticoagulation (VARIA)." *Journal of Thrombosis and Haemostasis* 8 (10): 2182–91.
- Rose, A. J., D. R. Berlowitz, A. S. Ash, A. Ozonoff, E. M. Hylek, and J. D. Goldhaber-Fiebert. 2011a. "The Business Case for Quality Improvement: Oral Anticoagulation for Atrial Fibrillation." *Circulation: Cardiovascular Quality and Outcomes* 4 (4): 416–24.
- Rose, A. J., E. M. Hylek, D. R. Berlowitz, A. S. Ash, J. I. Reisman, and A. Ozonoff. 2011b. "Prompt Repeat Testing after Out-of-Range INR Values: A Quality Indicator for Anticoagulation Care." *Circulation: Cardiovascular Quality and Outcomes* 4 (3): 276–82.
- Rose, A. J., E. M. Hylek, A. Ozonoff, A. S. Ash, J. I. Reisman, and D. R. Berlowitz. 2011c. "Risk-Adjusted Percent Time in Therapeutic Range as a Quality Indicator for Outpatient Oral Anticoagulation: Results of the Veterans Affairs Study to Improve Anticoagulation (VARIA)." *Circulation: Cardiovascular Quality and Outcomes* 4 (1): 22–9.
- Rose, A. J., E. M. Hylek, A. Ozonoff, A. S. Ash, J. I. Reisman, P. P. Callahan, M. M. Gordon, and D. R. Berlowitz. 2011d. "Relevance of Current Guidelines for Organizing an Anticoagulation Clinic." *American Journal of Managed Care* 17 (4): 284–9.
- Rose, A. J., D. R. Berlowitz, D. R. Miller, E. M. Hylek, A. Ozonoff, S. Zhao, J. I. Reisman, and A. S. Ash. 2012a. "INR Targets and Site-Level Anticoagulation Control: Results from the Veterans Affairs Study to Improve Anticoagulation (VARIA)." *Journal of Thrombosis and Haemostasis* 10 (4): 590–5.
- Rose, A. J., B. A. Petrakis, P. Callahan, S. Mambourg, D. Patel, E. M. Hylek, and B. G. Bokhour. 2012b. "Organizational Characteristics of High- and Low-Performing Anticoagulation Clinics in the Veterans Health Administration." *Health Services Research* 47 (4): 1541–60.

- Rose, A. J., D. R. Miller, A. Ozonoff, D. R. Berlowitz, A. S. Ash, S. Zhao, J. I. Reisman, and E. M. Hylek. 2013. "Gaps in Monitoring during Oral Anticoagulation: Insights into Care Transitions, Monitoring Barriers, and Medication Nonadherence." *Chest* 143 (3): 751–7.
- Stetler, C. B., L. J. Damschroder, C. D. Helfrich, and H. J. Hagedorn. 2011. "A Guide for Applying a Revised Version of the PARIHS Framework for Implementation." *Implementation Science* 6: 99.
- Sturges, J. E., and K. J. Hanrahan. 2004. "Comparing Telephone and Face-to-Face Qualitative Interviewing: A Research Note." *Qualitative Research* 4 (1): 107–18.
- Utarini, A., A. Winkvist, and G. H. Peltó. 2001. "Appraising Studies in Health Using Rapid Assessment Procedures (RAP): Eleven Critical Criteria." *Human Organization* 60 (4): 390–400.
- Vlassoff, C., and M. Tanner. 1992. "The Relevance of Rapid Assessment to Health Research and Interventions." *Health Policy and Planning* 7 (1): 1–9.
- Wennberg, J., and A. Gittelsohn. 1973. "Small Area Variations in Health Care Delivery." *Science* 182 (4117): 1102–8.
- Wishik, S. M., and S. Vynckt. 1976. "The Use of Nutritional 'Positive Deviants' to Identify Approaches for Modification of Dietary Practices." *American Journal of Public Health* 66 (1): 38–42.

SUPPORTING INFORMATION

Additional supporting information may be found in the online version of this article:

Appendix SA1: Author Matrix.