



Advancing the Science and Implementation of Patient Narratives: Summary of Presentations

October 7, 2021



Introduction and Overview.....	3
Meeting Background and Objectives	3
Plenary 1: The Power of Patient Stories.....	4
Plenary 2: Using Narratives for Quality Improvement	5
Concurrent Session 1: Presenting Narrative Information to Drive Improvement	6
Innovative Strategies for Internal Feedback Reporting	6
Use of Narratives in Experience-Based Co-Design.....	7
Concurrent Session 2: Extracting Insights from Narratives for Improvement	8
The Value of Patients’ Creative Ideas for Quality Improvement	8
Learning About Trust Through Narratives From People With Cystic Fibrosis	9
Plenary 3: Using Patient Narratives in Public Reports	11
Plenary 4: Introduction to the CAHPS Narrative Item Sets	14
Concurrent Session 1: CAHPS Clinician & Group Survey (CG-CAHPS) Narrative Item Set.....	14
Concurrent Session 2: CAHPS Health Plan Survey Narrative Item Set	16
Concurrent Session 3: CAHPS Child Hospital Survey (HCAHPS) Narrative Item Set.....	17
Plenary 5: Narrative Analysis Methods.....	18
Using Natural Language Processing (NLP) to Code Patient Narratives: Capabilities and Challenges	18
Using Public Source NLP Software to Code CG-CAHPS Comments in Massachusetts.....	19
Useful References.....	20

Introduction and Overview

On October 7, 2021, the Agency for Healthcare Research and Quality (AHRQ) convened a virtual invitation-only research meeting where approximately 120 participants discussed the use of patient narratives to assess and improve patients' experiences with care. During the five plenary and concurrent sessions, speakers and attendees explored some of the ways in which narratives are being used, the development and testing of open-ended items for specific surveys, and lessons learned about the use of natural language processing to analyze narratives. This meeting highlighted the extensive work done by the Consumer Assessment of Healthcare Providers and Systems (CAHPS) team over the past 10 years and expanded on themes explored in a similar research meeting five years earlier.¹

Meeting Background and Objectives

Dale Shaller, M.P.A., Principal, Shaller Consulting Group

There are multiple ways to listen to the “voice of the patient,” including both quantifiable metrics (e.g., standardized closed-ended surveys such as CAHPS) and qualitative feedback that captures patients' experiences in their own words (e.g., patient narratives, grievances and complaints). This meeting addressed the role that patient narratives can play in helping us understand and improve the patient's experience of care, with a particular focus on narratives collected through a structured series of open-ended questions added to the end of CAHPS and other patient experience surveys.

In response to the rapid growth in the reporting of patient comments about a decade ago, primarily through online rating sites and social media, the CAHPS Team initiated a research program for advancing the science and implementation of patient narratives. This program has four major components:

- 1) Narrative Elicitation:** The systematic prompting of patients to describe their experiences through a structured set of open-ended questions developed and tested with the same degree of scientific rigor used in the development of CAHPS closed-ended survey questions.
- 2) Narrative Analysis:** The development of analysis methods for carefully extracting key insights revealed by narratives through both human coding and machine learning techniques.

¹ See: [Advancing the Science and Implementation of Patient Narrative Elicitation and Reporting](#), September 8, 2016.

- 3) **Use of Narratives in Public Reporting:** The development of methods for reporting narrative information to consumers in ways that complement rather than obscure other performance metrics.
- 4) **Use of Narratives for Improvement:** The development of methods for presenting narrative feedback for use by providers and health care organizations to improve patient experience.

This research meeting was designed to shed light on each of these four components; the plenary and concurrent sessions featured CAHPS Team research findings as well as results from research and implementation projects sponsored by others.

Plenary 1: The Power of Patient Stories

Rachel Grob, Ph.D., Clinical Professor, Senior Scientist, and Director of National Initiatives, Center for Patient Partnerships, University of Wisconsin-Madison

Rigorously elicited and analyzed narratives provide insight about patients' experiences with care that cannot be gleaned via surveys alone. Stories collected using the CAHPS Narrative Item Sets provide concrete, actionable examples of aspects of care already being measured; emphasize elements of composite measures that are not being directly assessed; and highlight domains of care that are important to patients, but not yet included on the CAHPS survey to which narrative items are attached.

Narratives about COVID-19, collected within NewYork-Presbyterian's ambulatory care network between August 2020 and May 2021, illustrate all of these points. Analyses of narratives can also suggest ways to improve the structure and content of current surveys by capturing how patients' experiences are changing in real time. In over 400 comments analyzed for this presentation, patients shared perspectives on communication, access, office staff, and care coordination during the pandemic that were not captured by the closed-ended survey questions. They also told stories reflecting compassion with and empathy for providers; describing what it was like to seek and get care directly related to COVID, COVID testing, and COVID vaccines; detailing how changes to waiting areas affected them; and describing their perspective on COVID-related institutional policies.

Summary of the Discussion

Increase in patient comments during the COVID-19 pandemic: In response to a question asking if scores went up during the height of the pandemic because people were expressing their gratitude, Rachel highlighted the research team's interest in tracking the connection between the actionable information and details about patient experience provided in the comments and CAHPS scores. She added that although patients were asked about their experiences with outpatient care in general, and not specifically about their experiences with COVID, patients had a lot to say about their experiences with care during the pandemic. There were 407 comments related to care during the pandemic through May 2021, and the volume has not tapered off. The team is in the process of looking at discrete periods during the pandemic for further analysis.

Ensuring that comments from seriously ill patients are represented in the narratives: The narrative item set developed for outpatient settings was tested among seriously ill patients to ensure that their comments are seen in the data if such patients received care in outpatient settings. Based on data collected so far, there were narratives from patients who (1) talked about being very ill; (2) expressed how difficult it was to adjust to receiving care in a new location due to COVID; and (3) commented on how their care for a chronic condition was affected by COVID. The team is in the process of developing additional narrative item sets that target specific settings and types of patients (such as hospitalized children). Testing of narrative item sets is an important step in the development process that ensures that these questions gather rich data from a representative group of patients.

Providing a safe space for patients from racial and ethnic minority groups to be heard: The narrative item set was tested with a representative sample, including patients from racial and ethnic minority groups and people with lower literacy levels, and the items performed well among different sociodemographic populations. The CAHPS team is interested in improving the data collected from minority populations and in testing multiple modes of administering the narrative item sets to continue applying an equity lens.

Balance between closed-ended CAHPS survey questions and the narrative item set: The CAHPS team sees the value of the closed-ended and open-ended questions working together and how responses to each inform the other. The open-ended questions can illuminate the meaning of scores that are generated from the closed-ended questions. Research findings also affirm that patients appreciate being asked to comment in their own words through open-ended questions.

Plenary 2: Using Narratives for Quality Improvement

Ingrid Nembhard, Ph.D., Fishman Family President's Distinguished Professor, Associate Professor of Health Care Management, The Wharton School, University of Pennsylvania

Patient narratives are rich in insight. Given their power in that regard, a natural question is: how can we use these stories to improve the patient experience? This session presented quality improvement as a data-driven approach to improve care systematically and highlighted four requirements for successful improvement efforts: information, resources, workforce engagement, and creative ideas. The presentation then focused on the challenges to each requirement for improving patient experience and the ways in which narratives can address those challenges and facilitate improvement efforts. The conclusion is that patient narratives can serve as a useful tool in healthcare improvement.

Summary of the Discussion

Using patient narratives to engage the workforce: In the CAHPS team's research, healthcare staff were asked to share the frequency with which they had exposure to patient comments, as well as the usefulness of the comments they had received. The team found that frequent exposure of healthcare staff to patient comments alone was not sufficient to improve staff

experience and engagement; neither is the usefulness of the comments alone enough to improve staff experience and engagement. When healthcare staff were frequently exposed to useful patient comments, however, they reported being more inspired to engage in changes that would improve patient experience. Staff also reported feeling more confident in their knowledge to provide care. This highlights the importance of patient comments in staff's understanding of their role in the patient experience improvement process.

The importance of transparency when sharing patient experience data with stakeholders: It is important to have transparency as a key part of the data sharing process. Organizations often share only one of the following: (1) positive feedback; (2) negative feedback; or (3) negative feedback, but only with the specific people involved. Doing any of these is problematic because it provides an incomplete picture of what the patient experience really looks like in an organization. To build trust, it is important to share the raw data with stakeholders, so they do not receive only a filtered version of the data or an interpretation of the data. Sharing the full story builds trust in the data, and the narrative item set does a good job of eliciting full stories. The volume of the data also permits users to distinguish common experiences from outliers.

Presentation of both negative and positive patient stories with stakeholders: Regarding the ratio of positive and negative comments, Ingrid noted that behavioral researchers use a “sandwich approach” that involves sharing positive feedback, followed by negative feedback, and then more positive feedback to make it easier for people to digest the information. It is important to share data that are representative and then leave it to the organizations and staff to discern where they should focus their efforts. In many instances people are performing well, but there are also opportunities for improvement.

Sharing patient stories with patient advisory councils: Sharing data collected from patient stories with patient advisory councils is an area of interest for the CAHPS research team. The team's vision is that patient advisory councils with access to patient narrative data will help organizations come up with ideas on possible action steps to take.

CONCURRENT SESSION 1: PRESENTING NARRATIVE INFORMATION TO DRIVE IMPROVEMENT

Innovative Strategies for Internal Feedback Reporting

Rick Evans, Senior VP, Patient Services and Chief Experience Officer, NewYork-Presbyterian Hospital

By participating in a pilot project with the CAHPS team, NewYork-Presbyterian (NYP) Hospital's Ambulatory Care Network saw their staff culture quickly evolve to value narrative comments in addition to survey results. Using the Narrative Item Set for the CAHPS Clinician & Group Survey, they found that adding the narrative items resulted in the identification of important themes that could be reported to practice administrators in new ways to facilitate their use for improvement.

A previous version of NYP’s patient experience survey included three open-ended questions, with results shared with participating practices in an Excel spreadsheet. To test the use of enhanced feedback reporting methods, the team partnered with practice staff to design an interactive website with several innovative features. These features include:

- a visual depiction of trends in comment sentiment (coded as positive or negative),
- the number of comments received by date, which enables staff and clinicians to assess other patterns (e.g., the relationship to staffing issues, care continuity over holidays), and
- a range of filters to sort comments by key themes, key actors (e.g., comments about doctors, nurses, or office staff), and patient attributes (age, sex, and spoken language).

Practice managers found the website easier to use and more interesting than the Excel document. The ability to group comments thematically also allowed healthcare practices to identify and address specific issues. The pilot study demonstrated that enhanced narrative elicitation and feedback reporting can be successful in producing richer and more actionable information for improvement.

Use of Narratives in Experience-Based Co-Design

Bev Fitzsimons, Chief Executive, The Point of Care Foundation

Experience Based Co-Design (EBCD) is a participatory action research approach that combines a user-centered orientation (experienced-based) and a collaborative change process (co-design). This approach can be used to identify improvement priorities and devise and implement changes by changing mindsets, creating shared perspectives, and reconfiguring power within organizations. It is a cyclical process that generally starts by engaging staff and gathering their experiences, gathering patient experiences, conducting narrative interviews shared in a short film, and then bringing patients and staff together for co-design meeting(s). Together, patients and staff identify key themes and “touchpoints” – emotional or cognitive hotspots that stay with people for years afterwards. The co-design meetings focus on implementing changes by developing solutions to some of the challenges identified in the narrative interviews. Trust, long-term relationships, and high-quality communication are key to the success of this research approach.

Summary of the Discussion

Use of narratives to identify issues: Rick clarified that NewYork-Presbyterian has continued a longstanding process of reviewing comments to look for harm or reportable incidents, with comments scanned in real time and escalated as appropriate. With the expansion of narrative questions, they have not noticed a significant change in the number of reportable/harm events but have seen richer detail in the patients’ responses. It is difficult to compare new scores to previous ones as new patterns are emerging due in part to the ongoing effects of the COVID-19 pandemic. Responding to a question about concerns with patients naming individual providers and associated confidentiality or professionalism concerns, he noted that a filtering process is

used before comments are distributed internally, so that only leadership staff can see names. Practice administrators have guidelines about what to share, and this process conforms to longstanding guidelines about negative comments in support of a non-punitive safety culture.

Gathering and analyzing data: In response to a question regarding the frequency of data collection, Rick clarified that they conduct daily sampling, as there are not enough resources for census-based responses. The NewYork-Presbyterian team is currently working with the CAHPS team to analyze and report the data on a more frequent basis. The next step of this pilot project is to develop a plan for sustaining comment coding and narrative feedback reporting in the future. Seeing the results of the pilot convinced Rick's team of the value of the narrative data and reporting interface. The platform they are using to share results was built by the design firm Wowza and could be adapted by other organizations for similar purposes.

Encouraging participation in the design process: EBCD is a time-intensive process; it is important to prepare patients and build trust. In Bev's team's experience, people are generally happy to be asked to share their stories. To facilitate the success of the co-design workshops, they generally hold separate feedback events with staff and patients, then one larger co-design workshop that takes about an hour in-person. After that, smaller co-design teams meet to decide if they need to meet in-person or virtually, and the whole group is united together for a celebration at the end. An example of simple changes that can come from this co-design experience resulted from a suggestion by a new mother to improve the experience of giving birth in an operating theater by changing the way the gown is tied to facilitate skin-to-skin contact after birth. Bev stressed how critical it is to involve patients in co-design to support the overall success of improvement initiatives. If the end users of a process or system are not involved, it is quite possible that organizations will invest significant resources in unsuccessful improvement efforts that will fail to address the needs and preference of end users.

CONCURRENT SESSION 2: EXTRACTING INSIGHTS FROM NARRATIVES FOR IMPROVEMENT

The Value of Patients' Creative Ideas for Quality Improvement

Yuna Lee, Ph.D., M.P.H., Assistant Professor, Health Policy and Management, Columbia University

Healthcare leaders are calling for new approaches to patient experience improvement to address enduring quality gaps. Creating opportunities for patients to share their voice, specifically their creative ideas for health care improvement, may help close these gaps. To explore this idea, researchers developed and pilot tested an improvement-focused item that could be added to the CAHPS Clinician & Group Survey (CG-CAHPS) Narrative Item Set. They are also examining the relationship between patients' ideas and their CAHPS scores. The item development process involved developing and adapting items from the fields of service quality, quality improvement, design thinking, and user experience, and cognitively testing these items with patients and focus groups.

In the first step of the item testing process, patient experience officers, patients, and patient advocates rank-ordered item alternatives. The empirical test of the items involved a representative sample of patients from across New York State, oversampling for those with chronic diseases, who completed a selection of CG-CAHPS closed-ended questions, the five-question narrative item set, and four alternatives for the improvement item. The alternative that generated the most actionable items per idea and the highest average creativity score was *“If you could, what would you change about your care or interactions with your provider and their staff? Hearing from you is important to improving the care we deliver.”* This alternative represented an approach that induced creativity (generation of novel and useful ideas) and emphasized the system impact of the response (via the second sentence).

Compared to the original narrative item set, the item set with the improvement item doubles the number of actionable and creative responses. Improvement ideas span both CAHPS and non-CAHPS domains, and both idea types are negatively associated with CAHPS scores. Interpersonal-focused ideas are also negatively associated with CAHPS scores. Individuals who may systematically encounter poor care experiences (Black, female, low income) are more likely to share improvement ideas. This survey tool may offer a new channel to systematically access patients’ perspectives on improvement, elicit ideas from patients otherwise excluded or overlooked by patient engagement initiatives, and foster innovation.

Learning About Trust Through Narratives From People With Cystic Fibrosis

Stacy Van Gorp, Partner, See What I Mean, LLC

A research study funded by the Cystic Fibrosis Foundation examined trust in health care settings from narratives of people with cystic fibrosis and their families. The research team conducted secondary data analyses on patient narratives that were collected using the Patient and Family Experience of Care survey to examine whether patients and their families were talking about trust and what they said about trust. While there were no questions about trust specifically in the survey, 124 respondents used the word “trust” in six open-ended questions.

One of the main themes from the findings is that narratives are a powerful way for people to express the vulnerability of living with a chronic illness. The team also found that factors that help patients and their family members to trust their care team include competence of the care team, reliability of the care team, feeling known by the care team, having a caring care team, having a care team that listens, and communicating with the care team. Patients also expressed wanting to feel trusted by their care team. Some patients and their family members also use distrust as a strategy to manage cystic fibrosis. Finally, patients and family members expressed the role of trust in developing an effective partnership between patients, family members and care teams.

Summary of the Discussion

Sharing feedback with participants: One attendee noted that her organization shared feedback from patients and family members with others during a journey mapping exercise. This helped

others feel freer to open up, even though some of them were initially reluctant to provide feedback due to fear of losing their coverage. Stacy noted that her team has also shared what they were learning from the open-ended responses with family members to validate their findings.

Using patient narratives for loop closures: Stacy noted that patients and family members find it very meaningful to learn that feedback they provided has informed changes in an organization. Yuna agreed that narratives provide a great opportunity for “loop closure.” In her study, patient narratives were collected from multiple health systems. Those health systems developed methods for using feedback from narratives and reported back to patients and other stakeholders the changes that were made related to their feedback. Ingrid added that narratives are a tool for quality improvement and the virtual cycle created through loop closure is important. However, it can be challenging for organizations to identify effective approaches for conducting loop closures.

Theoretical and operational definitions of trust: Stacy noted that their operationalization of “trust” is the willingness to be vulnerable when you do not have control, based on what you do know in a given instance that is both situational and relational. In this study, the team was interested in learning what people with cystic fibrosis and their families meant when they were talking about trust and how that showed up for them. The definition of trust would have been a good follow-up question if the team could go back to the study participants and ask them additional questions. Rachel added that in qualitative research, there is always a balance between mapping to a theoretical framework and freeing yourself for the inductive insight that comes through the process.

Providers’ reactions to patients and family members’ comments about trust: Stacy and her team found that they needed to present their data in a way that does not evoke defensiveness among providers. The shared decision-making model and language like complementary expertise is likely a better way into the conversation. The Cystic Fibrosis Foundation has been sharing some of their survey results with a group of providers who have already been working on their partnership skills for a number of years; these providers have already come to their own conclusion that a partnership between providers and care recipients needs trust to be at its center. It is also useful to start from earning trust and then moving on to extending trust.

Rachel added that the team found value in elevating patient comments about trust even though they did not use a dyadic approach that included providers’ perspectives, which providers would have preferred. Although studies that use this dyadic approach are also valuable, they are not always possible. It is therefore important to note when study findings represent patient experiences and perceptions, regardless of whether the providers agree with them. Stacy mentioned that some reports of mistrust reflected systems-related issues rather than interpersonal issues.

Avoiding pressure to suggest radical changes: Yuna noted that when we think of creativity, we often think of creativity that is radical (i.e., doing away with waiting rooms rather than improving wait times in waiting rooms). In quality improvement, we think of creativity as a

whole spectrum that involves starting with incremental changes and “low hanging fruit” that may not seem like a lot from the perspective of the provider and the organization, but that still matter to patients. The “innovation” question is asking for ideas that are new to the setting, but not radical or paradigm shifting. In Yuna’s study, the team was very careful about **not** asking patients to think about innovative approaches in that moment. Rather, their approach was to create a space that allows patients to describe the kind of care they may or may not have received.

Interpersonal elements of patient narratives: The observation that tied both presentations together for Ingrid was the role of interpersonal trust. Yuna’s presentation noted that interpersonal trust was linked to CAHPS scores more than organizational trust; Stacy’s presentation illuminated the role that interpersonal trust played in the patient care experience.

While there was a balance between organizational and interpersonal ideas in Yuna’s research, she noted that interpersonal encounters that were salient to patients were often infused with emotions. They highlighted how patients wanted to be treated or how they wished to be treated in the future. The hypothesis is that because the interpersonal encounters were front and center to their experience, the salience and impact of their experience is the reason why we see a direct quantitative relationship between patient narratives and the CAHPS scores. Stacy commented that the interpersonal encounters evoked a lot of feelings and emotions among patients. Rachel added that the COVID data from patient narratives that she presented earlier in the meeting also illustrated the role of emotions in interpersonal encounters.

Plenary 3: Using Patient Narratives in Public Reports

*Christian Dankers, M.D., Associate Chief Quality Officer, Mass General Brigham
Adrienne Boissy, M.D., Chief Experience Officer, Cleveland Clinic*

The past decade has seen an increasing number of online provider rating sites – such as Angi, Yelp and Healthgrades – that post comments volunteered by patients and consumers. Research has shown growing use of these sites by consumers, not only as a first step in finding a new doctor but also to research a physician after a referral. In response to the proliferation of these online rating sites, a growing number of health systems have begun posting their own physician ratings with patient comments, increasing from just a handful of health systems a few years ago to an estimated 150 in 2018. This session featured a panel discussion of both the challenges and opportunities associated with posting comments on health system websites by two prominent physician leaders from two large academic health centers.

Christian Dankers, M.D., Associate Chief Quality Officer at Mass General Brigham (MGB), began with an overview of the MGB transparency program initiated in 2016. Providers receive an individual scorecard every two weeks with comments to be posted, with an opportunity to appeal their publication. He noted several challenges to the program, including the need to balance the voice of the patient by posting more comments with the potential negative impact on providers. He also stressed the importance of providing appropriate support for providers,

such as provider coaching, as well as system improvement efforts to address issues outside the providers' domain.

Adrienne Boissy, M.D., Chief Experience Officer at the Cleveland Clinic, described an extensive market research effort that recommended the implementation of "star satisfaction ratings" along with patient comments on the Cleveland Clinic website. This research revealed that patients value comments because "they tell the full story behind the ratings" and provide insight into the personality of the physician. The study also concluded that negative comments must be included on the website since they make the ratings and comments more credible. She reviewed several statistics on the impact of the Clinic's reporting website, including a 17% increase in page views within one month of launch, a 145% increase in physician profile page views, a 133% increase in sessions that start on physician profile pages, and more demand for higher rated clinicians.

Summary of the Discussion

The impact of public reports on physician ratings: Christian noted that the Brigham and Women's Hospital, which has reported physician ratings for some time, has observed a slight uptrend in the ratings, though they have been relatively stable. He added that it is important to make use of the patient narrative data in a way that supports physicians who are struggling while celebrating and sharing the practices of those who are thriving. Patients also raise many concerns that are not under the direct control of the provider, such as issues with access, parking, etc. His team has been thinking a lot about how to identify and address issues at the system level that are big drivers of patient experience.

Adrienne noted that physician ratings have remained relatively stable. However, the organization stopped posting physician ratings during the pandemic. Instead, her office has been sending a "dose of gratitude" via email to clinicians as a way of encouraging physicians and thanking them for the work they are doing, which was a pleasant surprise for physicians.

Equity and discrimination in physician ratings: Christian noted evidence that suggests that non-white practitioners systematically get lower ratings, not as a result of their competence but because of discrimination. His team is still trying to figure out the best way to provide patients with guidance on how to interpret provider ratings and balance positive and negative comments, recognizing that some negative comments may be due to discrimination and also that some issues are beyond providers' control. It is important to make patients aware that they need to review comments with caution because health care systems do not verify these comments before posting them. Adrienne added that she has found that some patients provide positive ratings if they are told what they want to hear; in addition, some patients become very angry and file complaints against clinicians when they are told the truth. This highlights the need for and importance of multiple patient stories to show the true performance of the clinician.

Physician-level issues versus organizational-level issues: The Brigham and Women's Hospital has considered having a practice transparency site where organizational issues identified by

patients that are not attached to a particular physician would be publicly reported, but has been unable to do so to date. Because system-level issues may inform patient's choice of providers, they believe that it is important to share all that information to allow patients to decide whether to be seen by a provider in that system. This then places the responsibility on the organization to act on those comments and fix the organization-level issues.

Adrienne added that although some organization-level issues may be out of the physician's control, clinicians can control some of these issues directly. It is important for physicians to take action to influence organization-level changes because these issues affect patient experience and clinician ratings. Susan Edgman-Levitan noted that patient advisors at Massachusetts General Hospital really wanted to see comments about the overall experience of care, beyond just the providers, because that also influences the patient's decision to see a provider. Upon reviewing all the primary care comments at Massachusetts General, particularly those flagged as a problem, she noticed that several were not about the actual experience with a clinician, but rather about issues that the hospital has been aware of for years. Reporting these issues publicly may motivate providers to try and do what they can to have those issues addressed immediately.

Coaching for physicians: The Cleveland Clinic uses a centralized system where providers who want to appeal any negative comments have an opportunity to discuss their concerns, although this has only happened a few times in the last 8 years. They use a phased approach that begins by rolling out 6 months of blinded data, and then 6 months of unblinded data. When they see consistent negative information across multiple data sources, they bring it to the clinician's and leadership's attention, who jointly try to determine what is going on. They also have a chief experience officer who provides counseling for clinicians.

Improve provider well-being: Adrienne explained that many private practice providers want their ratings to be publicly reported because it is good for their business. A study conducted by her team found higher ratings on the Jefferson's scale of empathy, higher scores on the HCAHPS survey, and lower rates of clinician burnout after clinician trainings. In addition, bringing clinicians together to talk to each other about their shared challenges and celebrate positive comments serves to lift up these clinicians. Christian added that some clinicians can be really bothered about negative comments even if they receive several positive comments and only receive one negative one. It is therefore important to celebrate the positive ones and help people think through the negative comments.

Demand for highly rated clinicians: Adrienne commented that patients usually want to be seen by clinicians with high ratings but are willing to be seen by other clinicians too. Healthcare systems need to be innovative in how care is provided because the current model is not working for some patients. These innovations may involve going to patients' homes, doing video visits, doing immediate triage, etc. Systems like Kaiser Permanente and Mayo Clinic have been transforming access to care and moving away from the classic model of providing care. Adrienne further noted that the different settings in which clinicians provide care, including outpatient, inpatient, and virtual settings, may impact the experience of care. Christian commented that most of the clinicians at his organization are highly rated, so they have not

experienced a lack of highly rated clinicians. Susan also noted that among the 13,500 providers at Massachusetts General, only three have star ratings below four. Since there is not much variation, patients do not have to forego the 4.97-star clinician and settle for one with 2 stars.

Advice for health systems considering publicly reporting patient comments on physician rating websites: Christian Dankers noted the importance of being clear about the purpose of reporting physician ratings publicly, as well as implications of this decision for providers. Specifically, health systems should communicate with their providers, obtain buy-in from providers, create the structures that will enable the organization to achieve its goal, and support providers who need help. The goal should be to use physician ratings as a means for quality improvement and better patient care. Adrienne added that publicly reporting physician ratings is a powerful tool to help patients make a choice to use a given facility. She reiterated the need to be prepared to manage clinicians' reactions to negative patient comments and to celebrate positive comments that clinicians receive.

Plenary 4: Introduction to the CAHPS Narrative Item Sets

Mark Schlesinger, Ph.D., Professor, Yale School of Public Health

Over the past two years, the joint Yale-RAND Narratives Team has been developing two additional narrative item sets (NIS) to complement the existing NIS that supplements CG-CAHPS. The first new NIS is designed to link with the CAHPS Health Plan Survey, while the second is designed to supplement the Child HCAHPS Survey. These new narrative item sets follow the same process of rigorous development and testing that was pioneered for the CG-CAHPS NIS, which was designed to assure that elicited narratives provide the most complete, balanced, and representative portrait of patient experiences possible.

The sequence of open-ended questions in each NIS collects actionable information (that is, with enough concrete detail for QI staff to leverage for improving performance) in 40-70% of all collected narratives. Between 35 and 50 percent of the narrative content extends beyond the scope of the closed-ended questions in the matching CAHPS survey. As the CAHPS Team has developed these new narrative item sets, it has also been attentive to eliciting patient perspectives regarding changing health care practices, to collecting feedback that can foster creative improvements in healthcare, and to ensuring that the narratives fully reflect experiences of ethnic/racial minority groups that have previously had restricted voice in many healthcare settings.

CONCURRENT SESSION 1: CAHPS CLINICIAN & GROUP SURVEY (CG-CAHPS) NARRATIVE ITEM SET

Mark Schlesinger, Ph.D., Professor, Yale School of Public Health

Working with its partner, NewYork-Presbyterian, to implement the CG-CAHPS Narrative Item Set (NIS) in field trials, the CAHPS team has continued to refine the sequence of open-ended questions, resulting in several enhancements. First, the team has been experimenting with

revised question wording designed to elicit creative ideas from patients more effectively: these changes significantly increased the number of narratives (currently about half of all of those with actionable feedback) that propose concrete solutions to problems. Second, the team has developed new wording for an open-ended question on telehealth, to complement both visit-based and time period-based versions of the CG-CAHPS Survey. Adding this new question more than doubled the narrative feedback related to telehealth interactions, compared to the previous set of CG-CAHPS narrative items.

Finally, the team has been increasingly focused on understanding differences based on patients' race and ethnicity in terms of engagement with narrative items as well as the content of elicited narratives. Black and Hispanic respondents are as engaged with the narrative questions as their non-Hispanic white counterparts, though engagement is lower among other minority groups. Although narratives elicited from all minority respondents are generally more negative than those from non-Hispanic whites, the domains of negative experience differ:

- Black patients are reporting more negative experiences related to communication,
- narratives from Hispanic respondents are more negative regarding access and perceived quality, and
- narratives elicited from patients in other minority groups are more negative regarding access, communication, and interactions with office staff.

Summary of the Discussion

Eliciting comments about how the provider treated the patient as a person: While developing the original narrative item set for the CG-CAHPS Survey, the team discovered a need for a question specifically about patient-clinician relationships, which was added as the fifth question in the sequence: *“Please describe your interactions with this provider and how you get along.”* The narratives provided in response to that question included comments on whether the provider:

- saw the patient as a person,
- engaged with the patient outside of the clinic,
- asked the patient about their family,
- remembered who the patient was, and
- understood the context of the patient's community.

With this sequence, the team has collected a very rich set of data about patient-clinician relationships. This type of information was found in about 15-20% of the narratives. The 2019 Milbank article, *“What Words Convey: The Potential for Patient Narratives to Inform Quality Improvement”* provides the exact prevalence of these kinds of responses.

Length and frequency of administering the patient narrative set: Depending on their utilization of healthcare services, some people at NewYork-Presbyterian receive the survey more than once a year; they occasionally report having already shared information earlier in the year. The team is also aware of concerns about the potential burden of adding questions and the possibility that people will not complete the survey because of the number of

questions. The evidence to date indicates that the additional items do not pose a hardship and that people enjoy answering the open-ended questions. Among the people who start answering the narrative items, 95% complete all the questions.

Narrative Item Set that would elicit diagnosis-related concerns: The research team has recently been funded by the Moore Foundation to develop a special diagnosis-related narrative. The development of a new set of questions will take a couple of years to be completed.

CONCURRENT SESSION 2: CAHPS HEALTH PLAN SURVEY NARRATIVE ITEM SET

Lise Rybowski, M.B.A., President, The Sevryn Group

Blake Hodges, Senior Consultant, National Market Research, Kaiser Permanente

CAHPS researchers are currently developing and testing a series of open-ended questions that organizations could add to the CAHPS Health Plan Survey to learn more about enrollees' experiences with their health care and health plan. The team worked with Kaiser Permanente as a key stakeholder and testing partner to pilot test these draft items and then conduct a much larger validity test of a revised set of items.

In the validity test, 94% of those who responded to Kaiser's survey online or by phone answered the narrative items. About two-thirds of those who responded offered a significant amount of substantive information (i.e., more than half of the questions had a codable response). Moreover, the level of drop-off was low: about 95% of respondents who answered the first question also answered the last one. The CAHPS team also compared the elicited narratives to the content of structured, hour-long interviews with a subset of respondents. This analysis assessed the extent to which the content from both sources matched, as well as the balance of positive and negative passages in the narratives and interviews. The team also looked at the extent to which the narrative respondents were representative of different subgroups in the larger population of survey respondents. Finally, an analysis of narratives indicated that nearly half of the narratives had at least one actionable passage, with some having four or more.

Kaiser Permanente regards open-ended questions as an extremely valuable addition to the Health Plan Survey. Narrative items that measure domains covered in the CAHPS survey reveal opportunities to improve survey scores, while narratives discussing domains not addressed in the survey give a more complete picture of members' experiences with their health care and health plan.

Summary of the Discussion

Ideal number of narrative items: When asked about the use by some organizations of a single open-ended question, Blake and Lise agreed that one narrative item is not sufficient. It is difficult to specify the ideal number of narrative items, but in their experience, one broad question is not adequate to elicit complete, balanced, and actionable feedback on a complicated topic like patient experience with healthcare. While the Health Plan Narrative Item

Set is not yet final, organizations were encouraged to consider implementing the full set of questions to be able to collect comprehensive and usable feedback.

CONCURRENT SESSION 3: CAHPS CHILD HOSPITAL SURVEY (HCAHPS) NARRATIVE ITEM SET

Kerry Reynolds, Ph.D., Senior Behavioral/Social Scientist, RAND
Kelli Carroll, PFAC Leader, UCLA Mattel Children's Hospital

This session focused on a narrative item set currently under development that is intended for use with the Child HCAHPS Survey, which is administered to parents of children who have been recently discharged from a hospital stay. The draft Child HCAHPS Narrative Item Set includes six questions that ask parents to provide information on the most important aspects of the hospital stay; positive experiences; negative experiences; parent interactions with providers, nurses and others; child interactions with providers, nurses and others; and care coordination.

Formal evaluation efforts comparing narrative item responses to structured interviews with parents are currently underway. Initial findings suggest that the draft Narrative Item Set elicits descriptions of important positive and negative experiences in a variety of domains that include relationships with providers/nurses, communication, thoroughness, efficiency, hospital environment, emotional experiences, advocacy, hospital processes, and more. A consultation with parent/family advisory committees suggested that parents value the opportunity to provide open-ended responses to questions about their child's hospital stay and highlighted issues related to implementation (e.g., mode, timing, confidentiality of responses). Once evaluation efforts are completed, the team will publish the results of the evaluation and finalize items and administration guidance.

Summary of the Discussion

One participant remarked that being admitted to the hospital can be very traumatic and wondered how narratives might address trauma. Kerry noted that trauma as a concept can apply to both the child who is hospitalized and the parent. One of the things that may be learned from narratives is the way in which hospitals can better support parents who are dealing with these situations. Previous research has shown that many medical diagnoses can be very difficult for parents to handle. Having some compelling examples of exactly how emotional factors are affecting the experiences of parents and children in the hospital may support quality improvement efforts and help physicians communicate with parents more effectively given the traumatic nature of the event.

Plenary 5: Narrative Analysis Methods

Using Natural Language Processing (NLP) to Code Patient Narratives: Capabilities and Challenges

Steven Martino, Ph.D., Senior Behavioral Scientist, RAND Corporation

Osonde Osoba, Ph.D., Senior Information Scientist, RAND Corporation

Patient narratives about experiences with health care contain a wealth of information about what is important to patients. These narratives are valuable for both identifying strengths and weaknesses in healthcare and developing strategies for improvement. However, rigorous qualitative analysis of the extensive data contained in these narratives is a resource-intensive process, and one that can exceed the capabilities of human analysts. One potential solution to these challenges is the application of natural language processing (NLP), which uses computer algorithms to extract structured meaning from unstructured natural language.

Because NLP is a relatively new undertaking in the field of healthcare, Martino, Osoba, and their colleagues at RAND set out to demonstrate the feasibility of NLP for organizing and classifying these data in a way that can generate actionable information. In doing so, the team focused on two steps that must be performed by a machine learning system designed to classify narratives into codes similar to those typically applied by human coders (e.g., positive or negative statements regarding care coordination). These steps are (1) numerically representing the text data (in this case, entire narratives) and (2) classifying the data by codes based on that numerical representation. The team also compared four related approaches to deploying machine learning algorithms, identified potential pitfalls in the processing of data, and showed how NLP can be used to supplement and support human coding. This presentation described the results of this study and what they suggest for health systems and other organizations considering using NLP to analyze patient narratives, either alone or in concert with coding by human analysts.

The presenters answered a clarifying question, explaining what they meant by the term “balanced accuracy.” Rather than simply looking at the overall accuracy—the number of cases the tool correctly coded (i.e., given a code similar to what a human coder chose) divided by the number of samples—they disaggregated the data to look at the proportion of positive cases (which were more common in the dataset) that were correctly coded and the proportion of negative cases (which were more uncommon) that were coded correctly. They used the analogy that if a dataset contained zeros and ones but was highly unbalanced, with almost all ones, an algorithm could be highly accurate simply by always predicting a value of one. Using balanced accuracy would, in effect, penalize the algorithm for incorrect guesses so it takes that dimension into account when making predictions. Using balanced accuracy as a marker of success forces the algorithm to correct for cases it codes incorrectly, rather than leaning on cases it codes correctly.

Using Public Source NLP Software to Code CG-CAHPS Comments in Massachusetts

Amy Stern, Director of Operations and Commercial Surveys

Raji Rajan, Senior Programming Manager, Massachusetts Health Quality Partners

Amy Stern and Raji Rajan explained the process Massachusetts Health Quality Partners (MHQP) used to develop a natural language processing tool to efficiently extract the meaning of the information provided by the five-item CG-CAHPS Narrative Item Set. MHQP has been involved in piloting early versions of the narrative item set since 2015. With the marked increase in the volume of responses—from less than 2,000 in the 2015 pilot to almost 35,000 in 2019—they realized they would need to leverage technology in order to be able to report patient narrative information back to provider organizations in a way that was easy, effective, and actionable for improvement purposes.

After testing different commercial options that did not provide meaningful classifications of comments, they chose Amazon Comprehend. They trained the tool to identify key themes in a given narrative by using a training file that included 300 manually coded comments for both the adult and child versions of the survey. They iterated the training process about four or five times to improve the tool for both versions of the survey, then used the 2019 data as the test file. Based on their results, they believe that in the future they will be able to use the tool to produce analyses of data specific to participating organizations, enabling those organizations to take action on the feedback their patients provide. The presenters emphasized that this is an iterative process; the more they train the file, the better it will be at predicting, and the more accuracy will improve, even as new questions are added.

Summary of the Discussion

Need for continual training: Asked about how much continual training is needed, particularly with new questions or emerging issues, Amy responded that they believe they have done most of the labor-intensive portion already. They are currently in the process of manually coding a telehealth classification but did not have enough comment volume to do as much training as they had hoped. They anticipate that the higher volume of comments expected with the addition of a telehealth question in the 2021 survey will allow for more robust use of their tool.

Accounting for systemic bias: A participant asked how systemic bias (race, gender, age) in coding is being addressed and accounted for in the NLP models, and how oversight by patients, bioethicists, and others can be incorporated into the interpretation of the models. The RAND team is aware of this important question but has not yet had the chance to examine their model.

Researchers are working to identify areas where bias can be introduced in these tools in order to be able to apply corrective mechanisms. The most direct way to address bias is to try to measure it in the system and observe it in the processes. Doing this successfully requires information on the demographic characteristics of the people supplying the unstructured text the algorithm is analyzing. In an ideal world, there would have to be information about the race, gender, age, etc., of the patient attached to each sample, but the collection of this


information would raise privacy issues. If the samples are coded with that information, however, users can do a desegregated evaluation to see how the accuracy of the coding model breaks down across different demographic groups. Although this approach will never be perfect, if analysis indicates that some demographic groups tend to have higher rate of errors of these classifications, the algorithm could be modified to address that balance. The primary issue is developing a secure, safe way to incorporate markers or demographic status in the output of these models to allow for further evaluation of bias. The question of oversight is important and should be further considered.

Practicality for other practices: A participant noted that these two presentations showed two different approaches to building an NLP tool, with the RAND group taking a bottom-up build-it-yourself approach, whereas MHQP modified an off-the-shelf tool. The presenters were asked if they thought this level of analysis would be available to most practices. Osonde shared the RAND team's view that the medical context affects the use of language, which differs from the free, unstructured language utilized for training purposes with most commercial models. They suggest careful evaluation to see if the commercially available models are able to account for the specific use of language and medical context and if it changes the results in meaningful ways. If it does, practices will need to modify the off-the-shelf tools to account for this issue. If not much modification is required, the cost of using an off-the-shelf tool would likely be lower than developing an algorithm in-house.

MHQP's team explored a number of the off-the-shelf tools; Raji noted that the level of hands-on technical support was sometimes less than they had expected, although the tools were cost-effective. He suggested that if practices have someone on staff with data science or other technical skills available to navigate some of these issues and handle training of the algorithm, they would likely be able to use a tool like Amazon Comprehend.

Useful References

The following publications were authored by researchers funded by AHRQ's CAHPS program:

- Grob R, Schlesinger M, Barre LR, et al. What Words Convey: The Potential for Patient Narratives to Inform Quality Improvement *Milbank Q.* 2019 Mar, 97(1): 176-227. <https://pubmed.ncbi.nlm.nih.gov/30883954>
- Grob R, Schlesinger M, Parker AM, et al. Breaking Narrative Ground: Innovative Methods for Rigorously Eliciting and Assessing Patient Narratives. *Health Serv Res* 2016 Jun;51 Suppl 2:1248–72. <https://pubmed.ncbi.nlm.nih.gov/27126144>
- Ish D, Parker AM, Osoba OA, et al. Using Natural Language Processing to Code Patient Experience Narratives: Capabilities and Challenges. 2020 Jan https://www.rand.org/pubs/research_reports/RRA628-1.html 
- Kanouse DE, Schlesinger, M, Shaller D, et al. How Patient Comments Affect Consumers' Use of Physician Performance Measures. *Med Care* 2016 Jan;54(1):24-31. <https://pubmed.ncbi.nlm.nih.gov/26551765>

- Schlesinger M, Grob R, Shaller D, et al. Taking Patients' Narratives about Clinicians from Anecdote to Science *N Engl J Med* 2015 Aug;373(7):675–679.
<https://pubmed.ncbi.nlm.nih.gov/26267629>
- Learn more about [CAHPS Patient Narrative Item Sets page on AHRQ's CAHPS website](#)



AHRQ Publication No. 22-0018
January 2022
www.ahrq.gov