BMJ Open

Identifying important and feasible primary care structures and processes in the US healthcare system: a modified Delphi study

Stephanie L Albert ,[1] Lorraine Kwok,[1] Donna R Shelley,[2] Maggie M Paul,[3] Saul B Blecker,[1] Ann M Nguyen,[4] Daphna Harel,[5,6] Charles M Cleland,[1] Bryan J Weiner,[7] Deborah J Cohen,[8] Laura Damschroder,[9] Carolyn A Berry[1]

To cite: Albert SL, Kwok L, Shelley DR, *et al*. Identifying important and feasible primary care structures and processes in the US healthcare system: a modified Delphi study. *BMJ Open* 2024;**14**:e082989. doi:10.1136/bmjopen-2023-082989

► Prepublication history and additional supplemental material for this paper are available online. To view these files, please visit the journal online (https://doi.org/10.1136/bmjopen-2023-082989).

Received 08 December 2023
Accepted 15 September 2024

For numbered affiliations see end of article.

Correspondence to
Dr Stephanie L Albert; stephanie.albert@nyulangone.org

ABSTRACT

Objective To identify primary care structures and processes that have the highest and lowest impact on chronic disease management and screening and prevention outcomes as well as to assess the feasibility of implementing these structures and processes into practice.

Design A two-round Delphi study was conducted to establish consensus on the impact and feasibility of 258 primary care structures and processes.

Participants 29 primary care providers, health system leaders and health services researchers in the USA.

Outcomes Primary outcomes were (1) consensus on the impact of each structure and process on chronic disease management and screening and prevention outcomes, separately and (2) consensus on feasibility of implementation by primary care practices.

Results Consensus on high impact and feasibility of implementation was reached on four items for chronic disease management: 'Providers use motivational interviewing to help patients set goals', 'Practice has designated staff to manage patient panel', 'Practice has onsite providers or staff that speak the most dominant, non-English language spoken by patients' and 'Practice includes mental health providers and/or behavioural health specialists in care team' and seven items for screening and prevention: 'Practice utilizes standing protocols and orders', 'Practice generates reports to alert clinicians to missed targets and to identify gaps in care, such as overdue visits, needed vaccinations, screenings or other preventive services', 'Practice has designated staff to manage patient panel', 'Practice sets performance goals and uses benchmarking to track quality of care', 'Practice uses performance feedback to identify practice-specific areas of improvement', 'Practice builds quality improvement activities into practice operations' and 'Previsit planning data are reviewed during daily huddles'. Only 'Practice has designated staff to manage patient panel' appeared on both lists.

Conclusion Findings suggest that practices need to focus on implementing mostly distinct, rather than common, structures and processes to optimise chronic disease and preventive care.

STRENGTHS AND LIMITATIONS OF THIS STUDY

⇒ The list of Delphi items reviewed was comprehensive and developed through a rigorous review of the literature.
⇒ The sample of Delphi experts included in this study were majority active practising primary care providers and health systems leaders.
⇒ We used a novel adaptation of the Delphi process in which members of the panel reviewed a subset of the 258 total items to reduce fatigue and burden.
⇒ Results from the Delphi are sensitive to the specific sample of respondents.

INTRODUCTION

Primary care is a critical point of access to the healthcare system because it often serves as patients' first point of contact for their healthcare needs, especially for preventive and chronic care.[1] In 2018, about half of all physician office visits (over 440 million) in the USA were to a primary care provider (PCP).[2] Of those primary care visits, 44.7% were for preventive care.[2] Although the receipt of primary care is associated with improved adherence to guideline-recommended preventive health screenings[3] and chronic condition care,[4] there remains a lag in the translation of these guidelines into practice in primary care settings.[5] Thus, it is important to identify which primary care structures (ie, the context in which care is delivered, such as care team composition, clinic characteristics and use of electronic health records (EHRs)) and processes (ie, actions and workflows designed to deliver healthcare, such as care coordination and panel management)[6] significantly contribute to high-quality primary care and are feasible to implement, and which of these structures and processes do not contribute to meaningful/significant improvements in care or outcomes.

Studies have linked various care structures and processes to increased receipt of preventive services (eg, cancer screenings, depression screenings and vaccinations)[7 8] and improved patient outcomes (eg, blood pressure and cholesterol).[9–11] Studies referring to the patient-centred medical home (PCMH) model, developed to improve care quality and patient outcomes while reducing costs, are especially prolific among these types.[12] The model emphasises principles of team-based care, care coordination and patient-centred access. While the PCMH model has been linked to improvements in preventive care and health outcomes,[8 9] there is little evidence indicating which specific components are impactful and independent of a broader set of practice-level changes. Additionally, the Center for Medicare and Medicaid Innovations (CMMI) recommended a list of over 60 changes in care structures and processes needed to transform clinical practice; however, it is unclear which of these structures and processes are highest priority and which ones are empirically associated with improvements in preventive services and patient outcomes.[13] Moreover, because new structures and processes are typically implemented together, there is limited empirical evidence on which specific structures and processes, individually and uniquely, have the greatest impacts on preventive services and patient outcomes as well as being broadly feasible to implement at the practice level.[14 15]

The vast array and complexity of recommendations about which structures and processes might lead to the best outcomes in preventive and chronic disease care pose a challenge to practices that are chronically underfunded and understaffed and have little leeway to invest in large-scale changes such as the PCMH model and CMMI's list for practice transformation. It is important for practices to have guidance on which individual structures and processes are most impactful on care and feasible to implement. Given this, the objective of this study was to discern the relative importance of individual practice-level structures and processes for patient-level chronic disease management and screening and preventive care outcomes. We also sought to evaluate the feasibility of integrating these structures and processes into primary care practice. To do this, we conducted a modified Delphi study among actively practising PCPs, health system leaders and health services researchers. The results from this study can provide useful insights into specific structures and processes that experts believe should be prioritised for implementation in order to help practices deliver high-quality care, and which should not be prioritised.

METHODS

Study methods were developed and reported based on the Conducting and Reporting Delphi Studies (CREDES) guidelines[16] (online supplemental table 1). We conducted a novel application of a Delphi approach to assess expert opinion on the importance and feasibility of specific structures and processes. This study was part of a larger effort to develop and validate a tool to assess practice-level primary care structures and processes that are associated with better quality and patient outcomes. A flow chart illustrating the Delphi process can be found in online supplemental figure 1.

Delphi item generation

As part of a larger study, we conducted a literature review to identify a broad and comprehensive set of primary care structures and processes that were hypothetically or empirically associated with specific patient outcomes related to two key components of primary care practice, chronic disease management and preventive care. Outcomes included aspirin use, blood pressure control, cholesterol control and/or statin use, diabetes control, screening and referral for smoking cessation, mental health screening, cancer screening and obesity/body mass index.[14] We searched PubMed/MEDLINE, Embase, Cochrane CENTRAL and Web of Science for US-based research studies published in English between 1 January 2010 and 31 December 2018. We included any published study, commentary or editorial that mentioned practice-level primary care structures and/or processes. We initially identified 7763 articles, reviewers screened titles and abstracts for inclusion, and data were extracted from 221 articles. We identified 640 structures and processes from these articles. We then engaged in an iterative nominal review process to cull the list. In the first round, our study team discussed the understandability (ie, was the identified structure or process clear) and uniqueness (ie, was the identified structure or process sufficiently different from other structures and processes) of items and then voted on keeping or deleting items for inclusion in the Delphi study. This resulted in a reduced list of 480 structures and processes. In the next round, our study's external advisory committee, comprised PCPs, health services researchers and health system leaders joined the study team to review and reduce the list, applying the same criteria of understandability and uniqueness. Lastly, the study team reviewed outstanding expert panel recommendations and further reduced the structures and processes, through group discussion, to a final list of 258 unique items for inclusion in the Delphi study. No additional investigator-driven prioritisation or consolidation strategies were employed so that the list of items could be as comprehensive as possible (ie, the full set of clear and unique items), while retaining their specificity enough to guide implementation and/or intervention strategies.

Pilot testing and refinement

All Delphi study materials including recruitment materials, a frequently asked questions document and the survey were developed by the research team in consultation with the external advisory committee. Prior to each round of the Delphi process data collection, we pilot tested the procedures and materials with five individuals similar to our target population (ie, actively practising

providers, health system leaders and health services researchers). Small refinements were made to the materials to increase participation and clarity.

Recruitment

To identify Delphi potential panel participants, the study team, in conjunction with our study's external advisory committee, nominated individuals with expertise in primary care delivery. We then supplemented the list of nominees with publicly available membership data from the National Association of Community Health Centers, the American Academy of Family Physicians and the Committee on Implementing High Quality Primary Care. We sought to include actively practising PCPs (eg, Doctor of Medicine (MD), Doctor of Osteopathic Medicine (DO), Nurse Practitioner (NP) and Physician Assistant (PA)), health system leaders (eg, chief medical officer and medical director) and academic researchers in primary care health services. Because these categories were not mutually exclusive, each individual was assigned a primary designation by the study team through group consensus. We compiled a list of 88 actively practising providers, 81 health system leaders and 55 researchers. Consistent with recommendations in the literature, the goal was to recruit 10 people from each category for a total of 30 experts.[17 18] We randomly selected and invited 60 experts (20 from each category). We then refreshed the recruitment list with 34 new randomly selected names to increase overall participation. In total 94 experts were invited to participate, representing all individually nominated experts and a subset of experts who were randomly selected from the publicly available membership data described above. Experts received a brief introductory email inviting them to participate, with a high-level overview of the study, estimated time to complete, information about incentives and links to a frequently asked questions document and the Qualtrics survey. Potential participants received up to four email outreach attempts and one phone call to solicit participation.

Data collection

To reduce participant cognitive burden, we decided to employ a novel approach to the Delphi method; we did not feel it was feasible, reasonable or realistic to ask respondents to thoughtfully consider the importance and feasibility of 258 items. Rather than the typical approach of having all Delphi panel members review all items, we split the 258 structures and processes into three lists with an equal number of items (n=86) that contained different combinations of conceptually related item sets. Items were first grouped into 10 'domains' that were closely aligned with the chronic care and the patient-centred medical home models (ie, access, care coordination, care management, comprehensive care, continuity of care, patient-centred care, performance measurement, population management, substance use/mental health and team-based care) and then within the larger 'domains' items were grouped into 'subdomains' (eg, care management, pre-visit planning, telemedicine, care plans and home-based monitoring subdomains were all part of the larger care management domain). 'Domains' appeared on more than one list but 'subdomains' remained intact; no grouping appeared on more than one list. Our goal for grouping lists in this way was to increase comprehension and evaluation of more narrowly defined sets of items while also minimising participant consideration of what might be 'missing'. Experts were randomly assigned to review only one list. Groupings were displayed in the Qualtrics survey in random order and not named to combat recency or primacy bias.

The Delphi study employed an iterative procedure in which respondents received the survey in rounds until acceptable consensus was reached. This allowed for maximal data collection without unnecessary panellist burden, attrition or fatigue. Participants received a US$150 honorarium for each round of participation for an estimated 30 min of time to complete the survey.

In the first round, experts were asked to indicate for each item 'How much do the following characteristics or elements of primary care positively impact patient outcomes related to chronic disease management (eg, hypertension, diabetes)?' This question was repeated for screening and preventive care (eg, cancer screening, influenza vaccinations). Responses to the two questions (chronic disease management; screening and preventive care) were on a 6-point Likert scale from 'no impact' to 'very large impact'. Respondents were instructed to (1) use the full 6-point response scale so that we would be able to discriminate among many potential items believed to have the strongest relationships with patient outcomes and (2) use the 'don't know' option only when they did not understand what the item was, otherwise were instructed that they should use their best judgement about the relationship between the item and outcomes. Finally, because experts only reviewed a subset of items, we provided an open-ended question at the end of the survey where they could add additional structures and processes of primary care they believed to be associated with care quality and patient outcomes.

Based on our observed results in the first round, a second round was conducted. In round 2, experts viewed mean scores as well as their own scores for the list of 86 items they rated in round 1 and had an option of changing their original scores. Additionally, in round 2 they were asked, 'How feasible would it be for primary care practices to implement this characteristic or element into their practice?' Responses were on a 5-point Likert scale from 'not feasible for any practices' to 'feasible for all practices'.

Analyses

We analysed the data using SAS software, V.9.4.[19] For these analyses, we considered the Likert scale responses for impact and feasibility as continuous variables (0='no impact', 5='very large impact'). Similarly, feasibility response options were coded as 0 for 'not feasible for

any practices' and 4 for 'feasible for all practices'. Using data from round 2, we calculated univariate statistics to assess the degree of perceived impact, feasibility, as well as consensus among raters. We summarised the mean responses on impact and used this information to rank items. In the event of ties, we assigned the higher rank to tied items. We aimed to identify the 10 highest and lowest rated items within each category; however, we note that ties could cause the number of items to vary.

We assessed consensus on impact and feasibility ratings descriptively using the standard deviation (SD) of the ratings we obtained. Because of the lack of standard cut-offs for what a small SD is in the literature, we chose to approach consensus by transforming the SD of ratings into a measure of certainty in the mean rating. Therefore, we considered consensus as a function of the width of a confidence interval (CI) for the mean rating on both importance and feasibility. This novel approach was employed because, due to our data collection procedure, we did not have data from all panellists to all items. In the presence of this block missing data circumstance, we used a sampling distribution approach in which we treated each panellist as a random draw from the population of possible panellists and computed the variability in each mean score that we obtained. This variability is represented in the form of a CI, which we aimed to minimise in width in order to maximise precision.

We considered consensus to be achieved by a narrow enough CI to be within half a point of the mean in either direction or a total CI width of one or smaller. For items to be considered feasible, raters had to reach consensus based on this standard, and we required the mean score to be, at minimum, greater than or equal to 2, corresponding to a mean response of 'feasible for some practices', 'feasible for most practices' or 'feasible for all practices'. Note that our a priori definition of consensus is consistent with others who have used mean and SD to establish consensus.[20]

In addition to univariate descriptive analyses, we performed independent sample t-tests to compare the mean of the highest rated items (ie, the ~10 top ranked items) to the mean of all other items and the mean of the lowest rated items (ie, the ~10 bottom ranked items) to the mean of all other items in order to assess whether the means of the highest and lowest rated items were statistically significantly different than the other items.

Patient and public involvement

Patients and/or the public were not involved in the design, or conducts, or reporting, or dissemination plans of this research.

FINDINGS

29 individuals completed the first round of the modified Delphi study (30.9% response rate) and all participants also completed the second round. Table 1 shows the demographic characteristics of the 29 Delphi participants.

Table 1 Delphi respondent characteristics (N=29)

Characteristic	n	Percent
Gender		
Male	17	58.6
Female	12	41.4
Credential		
MD	20	69.0
MD and Master's degree	1	3.5
MD and PhD	1	3.5
PhD or other doctorate degree	3	10.3
Master's degree	4	13.8
Type of expert*		
Primary care provider	21	72.4
Health system leader	20	69.0
Health services researcher	21	72.4
Region		
Northeast	12	41.4
West	8	27.6
South	4	13.8
Southwest	3	10.3
Midwest	2	6.9

*Respondents were able to select more than one.

The majority of participants were male (58.6%). Participants had a mix of credentials with most having an MD (69.0%), followed by Master's degree (13.8%), PhD/other doctorate (10.3%), MD and Master's degree (3.5%), and MD and PhD (3.5%). The variety in credentials reflected our goal of including actively practising PCPs, health systems care leaders and primary care health services researchers, all of whom have different standard credentials. Almost three-quarters of respondents identified as PCPs (72.4%) and health services researchers (72.4%), while more than two-thirds identified as health systems leaders (69.0%). However, based on the study team's primary designations, 9 providers, 11 health systems leaders and 9 researchers served as Delphi panellists. Participants were located across the country: Northeast (41.6%), West (27.6%), South (13.8%), Southwest (10.3%) and the Midwest (6.9%).

Below we present the highest and lowest rated items for chronic disease management, followed by screening and prevention. Results from round 1 and change scores between rounds can be found in online supplemental tables 2 and 3. See online supplemental tables 4 and 5 for the findings related to all 258 items for chronic disease management and screening and prevention.

Chronic disease management

Although we aimed to identify the 10 highest performing items related to chronic disease management, 12 items

Table 2 Comparison of highest and lowest ranked items with all other items

	n	Impact mean (SD)	P value
Chronic disease management			
Highest ranked items	9	4.3 (0.15)	<0.001
All other items	249	3.3 (0.65)	
Lowest ranked items	10	1.6 (0.24)	<0.001
All other items	248	3.4 (0.59)	
Screening and preventive care			
Highest ranked items	12	4.2 (0.20)	<0.001
All other items	246	2.4 (0.82)	
Lowest ranked items	12	0.57 (0.21)	<0.001
All other items	246	2.5 (0.80)	

tied for the 10th place ranking. Consequently, we focused on the top nine items. After two rounds of the Delphi study, the mean impact rating across the nine highest rated items was 4.3 out of 5 (SD: 0.15), and the mean of these nine items was significantly different from the remaining 249 items (4.3 vs 3.3; p<0.001) (see table 2). Experts reached consensus on the potential impact on chronic disease management outcomes for six out of the nine items (see table 3). The mean feasibility of implementation score for all nine items was 2.4 out of 4 (SD: 0.41) (results not shown). Seven of the nine items were determined to be feasible based on having consensus among raters on implementation feasibility as well as meeting the minimum threshold score (mean ≥2). When we applied a priori definitions of reaching consensus on impact in addition to implementation feasibility (consensus plus minimum threshold score), four of the nine items were determined to have high impact and would also be feasible to implement. These items were (1) 'Providers use motivational interviewing to help patients set goals', (2) 'Practice has designated staff to manage patient panel', (3) 'Practice has onsite providers or staff that speak the most dominant, non-English language spoken by patients' and (4) 'Practice includes mental health providers and/or behavioural health specialists in care team'.

Among the 10 lowest rated items, the mean impact score was 1.6 out of 5 (SD: 0.24), and this mean was notably and significantly lower than the other 248 items (1.6 vs 3.4; p<0.001). Consensus on impact was reached for only 1 out of the 10 lowest rated items. The average feasibility of implementation for the 10 items was 2.2 out of 4 (SD: 0.54) (results not shown). Two of the 10 items were determined to be feasible to implement based on reaching consensus among raters and meeting a minimum implementation threshold score. None of the 10 lowest rated items met both definitions of consensus on impact and implementation feasibility.

Screening and preventive care

Due to ties, we identified 12 items with the highest impact scores for screening and prevention outcomes. Across these 12 items, the mean impact score was 4.2 out of 5 (SD: 0.20) and the mean score was statistically significantly higher than that of all other items (4.2 vs 2.4; p<0.001). Consensus on impact ratings was reached on 9 out of the 12 items (see table 4). The average feasibility of implementation score was 2.7 out of 4 (SD: 0.52) for all 12 items (results not shown). 10 of the 12 highest rated items were determined to be feasible to implement based on having consensus among raters on feasibility and meeting the minimum threshold score. Seven items met all a priori criteria for being impactful and feasible to implement. These items were (1) 'Practice utilizes standing protocols and orders', (2) 'Practice generates reports to alert clinicians to missed targets and to identify gaps in care, such as overdue visits, needed vaccinations, screenings or other preventive services', (3) 'Practice has designated staff to manage patient panel', (4) 'Practice sets performance goals and uses benchmarking to track quality of care', (5) 'Practice uses performance feedback to identify practice-specific areas of improvement', (6) Practice builds quality improvement activities into practice operations' and (7) 'Pre-visit planning data are reviewed during daily huddles'.

12 items were identified as having the least potential impact on screening and preventive care outcomes. Among the lowest rated items related to screening and preventive care, the mean impact score was 0.57 (SD: 0.21). When compared with all other items, the mean score was significantly lower (lowest items: 0.57 vs all other items: 2.5; p<0.001). The mean feasibility of implementation score was 2.3 (SD: 0.44) for all 12 items (results not shown). Half of the lowest rated items met the feasibility of implementation definition (consensus among raters on feasibility plus meeting a minimum threshold score). Based on expert ratings, there was consensus that three items would have low impact on screening and preventive care outcomes yet would be feasible to implement. Those items were as follows: (1) 'The EHR generates alerts for non-formulary medication choices', (2) 'Providers receive alerts when patients are discharged from the emergency department and hospital' and (3) 'Providers receive alerts when patients are admitted to the emergency department and hospital'.

DISCUSSION

In this study, we aimed to understand the relative impact of various practice-level primary care structures and processes on chronic disease management and screening and preventive care outcomes, as well as the feasibility of implementation for practices. Our panel of Delphi experts, consisting predominately of actively practising PCPs and health systems leaders, reviewed 258 unique structures and processes. After two Delphi rounds, 9 items had the highest average ratings for chronic disease management

Table 3 Chronic disease management highest and lowest ranked items by impact

Item	Impact				Feasibility			Consensus+ threshold*
	Rank	n†	Mean (SD)	Consensus‡	Mean (SD)	Consensus‡	Threshold§	
Highest ranked items¶								
Practice routinely engages in bidirectional communication between primary care providers and mental health specialists	1	9	4.6 (0.73)	+	2.4 (0.88)	–	+	–
Providers use motivational interviewing to help patients set goals	2	11	4.5 (0.69)	+	3.2 (0.60)	+	+	+
Practice works with inpatient case managers to identify high-risk patients who need follow-up support	3	11	4.4 (0.92)	–	2.1 (0.57)	+	+	–
Practice uses registry/EHR data to track gaps in care and outcomes	4	9	4.3 (1.00)	–	2.6 (0.73)	+	+	–
Practice generates reports to alert clinicians to missed targets and to identify gaps in care, such as overdue visits, needed vaccinations, screenings or other preventive services	4	9	4.3 (0.87)	–	2.4 (0.53)	+	+	–
Care coordinator contacts patients in need of follow-up	6	9	4.2 (0.44)	+	2.9 (0.78)	–	+	–
Practice has designated staff to manage patient panel	6	11	4.2 (0.60)	+	2.3 (0.79)	+	+	+
Practice has onsite providers or staff that speak the most dominant, non-English language spoken by patients	6	9	4.2 (0.67)	+	2.0 (0.71)	+	+	+
Practice includes mental health providers and/or behavioural health specialists in care team	6	9	4.2 (0.67)	+	2.0 (0.50)	+	+	+
Lowest ranked items								
Practice routinely asks about sexual orientation and gender identity	249	11	1.9 (1.30)	–	3.4 (0.67)	+	+	–
Practice contracts with external organisations to provide comprehensive care management for medically complex patients	249	11	1.9 (1.76)	–	1.4 (0.81)	+	–	–
Clinicians prescribe changes in health behaviours (eg, diet, physical activity) using prescription pads	251	11	1.8 (0.75)	+	2.5 (1.04)	–	+	–
Practice uses an interactive health record specifically to track preventive health	251	10	1.8 (1.23)	–	2.1 (0.88)	–	+	–
Clinician and staff salaries are linked to performance (eg, quality of care, patient experience, resource utilisation)	251	11	1.8 (1.33)	–	1.7 (0.79)	+	–	–
Practice routinely uses paraprofessionals (eg, community health workers, health advocates) in some capacity (eg, patient education, home visits, lifestyle coaching)	254	9	1.7 (2.00)	–	2.2 (0.67)	+	+	–
Practice sends newsletters to patients	255	9	1.4 (0.88)	–	2.4 (0.88)	–	+	–

Continued

Table 3 Continued

Item	Impact				Feasibility			
	Rank	n†	Mean (SD)	Consensus‡	Mean (SD)	Consensus‡	Threshold§	Consensus+ threshold*
Practice offers an online algorithm for patients to assess symptoms	255	8	1.4 (1.06)	–	2.3 (1.04)	–	+	–
Practice uses interactive voice response (IVR) to monitor patient outcomes (eg, BP)	255	8	1.4 (0.92)	–	1.8 (0.89)	–	–	–
Staff notifies waiting times at registration or over the phone	258	11	1.3 (1.10)	–	2.2 (1.17)	–	+	–

*Includes consensus on importance and both measures of feasibility (ie, consensus and minimum threshold score of ≥2)
†n=Number of raters who reviewed the items
‡+=Reached consensus; –=Did not reach consensus
§+=Reached minimum feasibility threshold score of ≥2; –=Did not meet minimum feasibility threshold score of ≥2
¶This table includes only 9 highest ranked times because 12 items tied for the 10th ranking
BP, blood pressure; EHR, electronic health record.

outcomes and 12 items had the highest average ratings for screening and preventive care outcomes. Among these items, there was consensus regarding the potential impact and feasibility of implementation for only a subset of items. Specifically, when considering chronic disease management outcomes there was agreement among experts on four items while the experts separately came to consensus on seven items as they related to screening and prevention outcomes. Interestingly, only one item appeared on both lists: 'Practice has designated staff to manage patient panel'. Together these findings suggest that practices need to concentrate on implementing mostly distinct, rather than common, structures and processes to optimise chronic disease and preventive care. Given that high-quality primary care delivery necessitates that practices be able to provide both types of care well, it is essential that we be able to distinguish the most essential and impactful structures and processes associated with optimal clinical outcomes.

As noted, 'Practice has designated staff to manage patient panel' was at the top of both lists, signifying that it is most likely to have the greatest potential value to clinical care and that it would be feasible for practices to implement. Despite prevailing expert opinion in our study, there is a notable scarcity of empirical evidence to substantiate the importance or feasibility of designating a staff member for panel management. Two recent articles have assessed this item but findings were mixed. Schwartz *et al* found that incorporating panel management assistants was a feasible and valued strategy but was not associated with patient outcomes related to blood pressure, blood pressure control or smoking rates, although smokers had increased odds of receiving nicotine replacement therapy.[21] Chwastiak *et al* found a statistically greater change in hemoglobin A1c (HbA1c) in patients with diabetes who were enrolled in a collaborative care programme as compared with those not in the programme.[22] Unfortunately, because having designated staff to manage the patient panel was part of a larger multidisciplinary effort, it is not possible to attribute findings directly to this structural component. Given the mixed empirical evidence for this item's impact on chronic disease management, as well as screening and preventive care outcomes, and the fact that it was tested as part of multicomponent efforts, additional research is needed.

Among the 10 unique items where consensus on impact and feasibility of implementation was reached, we found they represented 6 different conceptually related groupings. The variation and array of primary care structures and processes that were judged to be important and feasible by experts illustrate that it is essential that providers and health system leaders think holistically about care. Furthermore, given the lack of substantial evidence supporting which specific structures and processes are genuinely associated with quality care and patient outcomes, consensus among experts is especially meaningful. In this study we have a vastly reduced list of

Table 4 Screening and preventive care highest and lowest ranked items by impact*

Item	Impact				Feasibility			Consensus + Threshold†
	Rank	n‡	Mean (SD)	Consensus§	Mean (SD)	Consensus§	Threshold¶	
Highest ranked items								
Practice uses registry/EHR data to track gaps in care and outcomes	1	9	4.6 (1.01)	–	2.6 (0.73)	+	+	–
Practice's patients are divided into patient panels assigned to specific provider teams	2	11	4.5 (0.69)	+	3.0 (0.89)	–	+	–
Practice utilizes standing protocols and orders	3	9	4.3 (0.71)	+	3.3 (0.71)	+	+	+
Practice generates reports to alert clinicians to missed targets and to identify gaps in care, such as overdue visits, needed vaccinations, screenings or other preventive services	3	9	4.3 (0.71)	+	2.4 (0.53)	+	+	+
Practice has designated staff to manage patient panel	5	11	4.2 (0.40)	+	2.3 (0.79)	+	+	+
Practice sets performance goals and uses benchmarking to track quality of care	6	11	4.1 (0.54)	+	3.3 (0.65)	+	+	+
Practice uses performance feedback to identify practice-specific areas for improvement	6	11	4.1 (0.30)	+	2.9 (0.83)	+	+	+
EHR and panel management software are interoperable (ie, exchange data)	6	11	4.1 (0.70)	+	1.5 (0.53)	+	–	–
Practice builds quality improvement activities into practice operations	9	9	4.0 (0.71)	+	3.0 (0.71)	+	+	+
Pre-visit planning data are reviewed during daily huddles	9	11	4.0 (0.77)	+	3.0 (0.45)	+	+	+
Practice is open during evening and/or weekend hours	9	11	4.0 (0.89)	–	2.9 (0.83)	+	+	–
Providers can directly order diagnostic tests, medication and counselling from clinical decision support tool	9	9	4.0 (1.00)	–	2.3 (0.50)	+	+	–
Lowest ranked items								
Practice offers facilitated group visits on topics related to self-management of chronic disease(s)	247	10	0.8 (1.03)	–	2.0 (0.63)	+	+	–
Practice combines home-based monitoring (eg, blood pressure, blood glucose) with nurse-administered behavioural interventions	247	9	0.8 (1.09)	–	2.0 (0.71)	+	+	–
Staff schedule routine visits with a pharmacist for patient	247	11	0.8 (0.87)	–	1.7 (0.47)	+	–	–
Practice has a designated person who performs medication reconciliation	250	11	0.7 (0.79)	+	2.7 (0.90)	–	+	–
Practice routinely uses paraprofessionals (eg, community health workers, health advocates) in some capacity (eg, patient education, home visits, lifestyle coaching)	250	9	0.7 (1.41)	–	2.2 (0.67)	+	+	–
The EHR has automated tools to monitor and alert clinicians regarding patient medication adherence	252	11	0.6 (1.12)	–	1.6 (0.92)	–	–	–
Clinicians assess and track medication adherence in EHR	253	11	0.5 (1.21)	–	2.4 (1.03)	–	+	–

Continued

Table 4 Continued

Item	Impact				Feasibility			Consensus + Threshold†
	Rank	n‡	Mean (SD)	Consensus§	Mean (SD)	Consensus§	Threshold¶	
The EHR generates alerts for non-formulary medication choices	253	11	0.5 (0.82)	+	2.3 (0.65)	+	+	+
Practice uses interactive voice response (IVR) to monitor patient outcomes (eg, BP)	253	8	0.5 (0.76)	–	1.8 (0.89)	–	–	–
The EHR alerts the provider about drug–drug interactions	256	5	0.4 (0.89)	–	2.9 (1.04)	–	+	–
Providers receive alerts when patients are discharged from the emergency department and hospital	256	11	0.4 (0.81)	+	2.8 (0.60)	+	+	+
Providers receive alerts when patients are admitted to the emergency department and hospital	258	11	0.1 (0.30)	+	2.6 (0.67)	+	+	+

*This table includes 12 highest and lowest ranked items due to ties.
†Includes consensus on importance and both measures of feasibility (ie, consensus and minimum threshold score of ≥2)
‡n=Number of raters who reviewed the items
§+=Reached consensus; –=Did not reach consensus
¶+=Reached minimum feasibility threshold score of ≥2; –=Did not meet minimum feasibility threshold score of ≥2
BP, blood pressure; EHR, electronic health record.

‘higher priority’ structures and processes based on systematic PCP, health system leader and researcher input. We submit that these items rated highly with expert consensus are those that practices should prioritise for implementation and/or intervention. The lack of consensus on quite a few of the highly rated items underscores both the lack of, and need for, rigorous empirical studies of the associations between these specific structures and processes and quality of care metrics and patient outcomes.

Delphi panellists also identified structures and processes that were deemed to have lower value for implementation and/or intervention. 10 and 12 items received the lowest ratings on impact for chronic disease management and screening and preventive care outcomes respectively. Consensus was reached on four items (one for chronic disease outcomes; three for screening and prevention outcomes), though none appeared on both lists. While we also assessed feasibility of implementation, this concept is less relevant as these items were determined by experts to be lower priority structures or processes for implementation in primary care practices. Based on expert opinion, these low rated structures and processes should not be prioritised for implementation by practices in favour of other more impactful and feasible items identified in this study.

To our knowledge, this is the first study to identify which care structures and processes are important and feasible to implement in primary care settings in the USA. A previous Delphi study among Canadian providers assessed the importance of organisational structures and processes to family practice.[23] Our study differs from this earlier research in a number of ways including the fact that our study is based on healthcare delivery in the USA, is more contemporary and follows important healthcare transformations including the Affordable Care Act and Health Information Technology for Economic and Clinical Health Act, and we were focused on more specific structures and processes rather than implementing outcomes, limiting the comparability.

Strengths and limitations

A major strength of this study is that two-thirds of the Delphi experts are actively practising primary care physicians and/or health system leaders who have experience ‘in the field’ and regularly make decisions about investments in primary care structures and processes. Inclusion of these experts in the Delphi study is aligned with recommendations for greater practice-based evidence and the need to bring research closer to real-life practice settings.[24] However, this study is not without limitations. First, although this study is data driven, we did not assess the reliability of our results. That is, had different experts reviewed the same set of items the ratings may have been different and we might have arrived at a different consensus about what is impactful and feasible. Second, we did not rotate items across survey administrations resulting in experts reviewing only subset of items rather than all 258 items. While this methodological

decision was made to reduce participant burden, it may be possible that our findings would have been different if all experts reviewed all items. Third, we have limited information about the people who declined to participate; therefore, our results are sensitive to completers of our modified Delphi study. Fourth, additional primary care characteristics suggested by experts in response to the open-ended question were not rated for importance or feasibility by the other experts; however, after reviewing write-in responses we determined that only three new items were suggested and by only one person each time. The limited number of new write-in responses illustrates that our original list of items was comprehensive and lessens concern about important items not being rated by multiple Delphi participants. Lastly, our definition of consensus is novel and has not previously been defined by others in the Delphi literature in the same way, though it is consistent with those who have used mean and SD to establish consensus.

Conclusion

Amidst the extensive literature on primary care interventions, there is a conspicuous absence of robust evidence about which elements of primary care delivery are uniquely associated with higher quality and better patient outcomes. This study addresses this knowledge gap to some degree by identifying care structures and processes that represent high and low value for implementation and/or intervention as judged by experts in primary care. By including a focus on both clinical significance and feasibility for practices, our hope is that these findings can inform studies of implementation strategies to improve quality of care by guiding the targets for practice change.

Author affiliations
[1]Department of Population Health, New York University Grossman School of Medicine, New York, New York, USA
[2]Department of Public Health Policy and Management, New York University College of Global Public Health, New York, New York, USA
[3]Robert D. and Patricia E. Kern Center for the Science of Health Care Delivery, Mayo Clinic, Phoenix, Arizona, USA
[4]Center for State Health Policy, Rutgers, New Brunswick, New Jersey, USA
[5]Applied Statistics, Social Science, Humanities, New York University, New York, New York, USA
[6]PRIISM Applied Statistics Center, New York University, New York, New York, USA
[7]Department of Health Services, University of Washington, Seattle, Washington, USA
[8]Department of Family Medicine, Oregon Health and Science University, Portland, Oregon, USA
[9]VA Ann Arbor Health System Center for Clinical Management Research, Ann Arbor, Michigan, USA

Acknowledgements We thank the Delphi participants for their time and effort. We would also like to thank Rachel Thornton and Mike Barry for their help with recruiting participants and administering the Delphi survey, Steven Friedman for his help with managing and conducting preliminary analyses of the data and Kayla Fennelly for her help with preparing the supplementary materials for submission.

Contributors SLA is the guarantor of the study. SLA, CAB and DRS conceptualised and designed the study with support from MMP, DH, CMC, SBB, AMN, BJW, LD and DJC. SLA and CAB conceptualised and led data acquisition. SLA and LK analysed the data with support from DH and CMC. SLA and LK drafted the initial manuscript. All authors critically reviewed the manuscript drafts and approved the final version. We attest all authors meet ICMJE requirements for authorship.

Funding This work was supported by the National Center for Advancing Translational Science (NCATS) grant number (5U01TR002008).

Competing interests None declared.

Patient and public involvement Patients and/or the public were not involved in the design, or conduct, or reporting, or dissemination plans of this research.

Patient consent for publication Not applicable.

Ethics approval This study was deemed exempt by the New York University (NYU) Grossman School of Medicine Institutional Review Board; therefore, written consent was not necessary and completion of the survey was considered participant consent.

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement Data are available on reasonable request.

ORCID iD
Stephanie L Albert http://orcid.org/0000-0003-3233-3742

REFERENCES

1 Institute of Medicine.Defining primary care. In: Donaldson MS, Yordy KD, Lohr KN, et al, eds. *Primary Care: America's Health in a New Era*. Washington, DC: The National Academies Press, 1996.
2 Santo L, Okeyode T. National Ambulatory Medical Care Survey: 2018 National Summary Tables, Available: https://www.cdc.gov/nchs/data/ahcd/namcs_summary/2018-namcs-web-tables-508.pdf
3 Rao A, Kale MS. Characterizing the Role of the Primary Care Provider in Preventive Health Exams: NAMCS 2011-2014. *J Gen Intern Med* 2019;34:1385–7.
4 Willis J, Antono B, Bazemore A, *et al*. The state of primary care in the United States: A chartbook of facts and statistics. 2020.
5 Grimshaw JM, Eccles MP, Lavis JN, *et al*. Knowledge translation of research findings. *Implementation Sci* 2012;7:50.
6 Donabedian A. The quality of care. How can it be assessed? *JAMA* 1988;260:1743–8.
7 Mader EM, Fox CH, Epling JW, *et al*. A Practice Facilitation and Academic Detailing Intervention Can Improve Cancer Screening Rates in Primary Care Safety Net Clinics. *J Am Board Fam Med* 2016;29:533–42.
8 Pourat N, Charles SA, Snyder S. Availability of Care Concordant With Patient-centered Medical Home Principles Among Those With Chronic Conditions. *Med Care* 2016;54:262–8.
9 Tai-Seale M, Wilson CJ, Panattoni L, *et al*. Leveraging Electronic Health Records to Develop Measurements for Processes of Care. *Health Serv Res* 2014;49:628–44.
10 Mundt MP, Gilchrist VJ, Fleming MF, *et al*. Effects of Primary Care Team Social Networks on Quality of Care and Costs for Patients With Cardiovascular Disease. *Ann Fam Med* 2015;13:139–48.
11 Zhou YY, Kanter MH, Wang JJ, *et al*. Improved Quality At Kaiser Permanente Through E-Mail Between Physicians And Patients. *Health Aff (Millwood)* 2010;29:1370–5.
12 Philip S, Govier D, Pantely S. *Patient-Centered Medical Home: Developing the Business Case from a Practice Perspective*; 2019. Available: https://www.ncqa.org/wp-content/uploads/2019/06/06142019_WhitePaper_Milliman_BusinessCasePCMH.pdf
13 Transforming Clinical Practice Intiative (TCPi). TCPI Change Package: Transforming Clinical Practice, 2016. Available: https://www.pcpcc.

org/sites/default/files/resources/TCPI%20Change%20Package_Color_March%2016_v2.0.pdf
14 Paul M, Shelley D, Albert SL, *et al*. Linking practices and outcomes: Evidence for discretepractice-level strategies to improve primary care outcomes. presented at:AcademyHealth Science of Dissemination and Implementation Annual Conference. 2021.
15 Nguyen AM, Paul MM, Shelley DR, *et al*. Ten Common Structures and Processes of High-Performing Primary Care Practices. *J Public Health Manag Pract* 2022;28:E639–44.
16 Jünger S, Payne SA, Brine J, *et al*. Guidance on Conducting and REporting DElphi Studies (CREDES) in palliative care: Recommendations based on a methodological systematic review. *Palliat Med* 2017;31:684–706.
17 Clayton MJ. Delphi: a technique to harness expert opinion for critical decision-making tasks in education. *Educ Psychol (Lond)* 1997;17:373–86.
18 de Villiers MR, de Villiers PJT, Kent AP. The Delphi technique in health sciences education research. *Med Teach* 2005;27:639–43.
19 SAS Institute Inc. *SAS software. Version 9.4*. 2013.
20 von der Gracht HA. Consensus measurement in Delphi studies: Review and implications for future quality assurance. *Technol Forecast Soc Change* 2012;79:1525–36.
21 Schwartz MD, Jensen A, Wang B, *et al*. Panel Management to Improve Smoking and Hypertension Outcomes by VA Primary Care Teams: A Cluster-Randomized Controlled Trial. *J Gen Intern Med* 2015;30:916–23.
22 Chwastiak LA, Jackson SL, Russo J, *et al*. A collaborative care team to integrate behavioral health care and treatment of poorly-controlled type 2 diabetes in an urban safety net primary care clinic. *Gen Hosp Psychiatry* 2017;44:10–5.
23 Barnsley J, Berta W, Cockerill R, *et al*. Identifying performance indicators for family practice: assessing levels of consensus. *Can Fam Physician* 2005;51:700–1.
24 Green LW. Making research relevant: if it is an evidence-based practice, where's the practice-based evidence? *Fam Pract* 2008;25:i20–4.