Identifying Outcome Measures for Type 2 Diabetes Value-Based Contracting Using the Delphi Method

Lynn M. Neilson, PhD; Elizabeth C.S. Swart, BS; Chester B. Good, MD, MPH; William H. Shrank, MD, MSHS; Rochelle Henderson, PhD; Chronis Manolis, RPh; and Natasha Parekh, MD, MS

ABSTRACT

BACKGROUND: Value-based contracts (VBCs) between payers and pharmaceutical manufacturers link drug payments to predefined performance measures and require shared risk between both entities. It is unclear how outcome measures were selected in previously reported VBCs, and many VBCs have focused on surrogate endpoints often used in the conduct of clinical trials, which may not be valued by or of importance to patients.

OBJECTIVE: To identify outcome measures that are meaningful to key stakeholders and feasibly measured to inform VBCs for diabetes medications.

METHODS: We conducted a modified Delphi survey to incorporate views from patients (n=9), endocrinologists (n=5), primary care physicians (n=4), payers (n=3), pharmacy benefit managers (n=3), and pharmaceutical company representatives (n=2). A list of 12 diabetes-related outcome measures was generated from the literature and consultations with subject matter experts. Participants rated the importance of each outcome on a 5-point Likert scale and selected the 3 most meaningful outcomes. Nonpatient participants then used a Likert scale to rate the feasibility of collecting each outcome. Consensus was defined as $\geq 75\%$ agreement on the importance and feasibility of an outcome (Likert scores 4 or 5 or selection of an outcome as most meaningful). A 2-sample test of proportions was performed to examine differences between patient and nonpatient stakeholder rankings of outcomes.

RESULTS: All 12 outcomes reached consensus for importance on the Likert scale. The measure “reducing risk of heart attacks” was the most meaningful outcome (84%), while “reducing A1c levels” ranked second (68%). The 2 measures rated as most feasibly collected were “reducing A1c levels” and “reducing risk of hospitalizations from diabetes” (93.8% each). The measures “weight loss,” “reducing risk of diabetes-related kidney disease,” “reducing risk of emergency room visits from diabetes,” and “reducing risk of diabetes-related amputations and foot ulcers” also reached consensus for feasibility. There were statistically significant differences between patient and nonpatient stakeholders in the selection of “reducing A1c levels” (37.5% vs. 82.3%, respectively; $P=0.03$) and “reducing risk of diabetes-related kidney disease” (50.0% vs. 11.8%, respectively; $P=0.03$) as most meaningful outcomes.

CONCLUSIONS: The measures “reducing risk of heart attacks” and “reducing A1c levels” were identified as top priority diabetes outcome measures.

J Manag Care Spec Pharm. 2019;25(3):324-31

Copyright © 2019, Academy of Managed Care Pharmacy. All rights reserved.

What is already known about this subject

- The costs of insulin and other antidiabetic medications have risen dramatically over the past 2 decades, contributing to the consistent placement of diabetes as one of the top expenditures among therapeutic drug classes.
- Value-based contracts between payers and pharmaceutical manufacturers link drug payments to predefined outcome measures and require shared risk between both entities.
- Selection of clinical outcomes that are meaningful to all stakeholders and feasibly measurable via existing claims and medical data remains a challenge in the development of effective value-based contracts.

What this study adds

- A panel of key stakeholders that included patients, providers, payers, pharmacy benefit managers, and pharmaceutical company representatives identified “reducing risk of heart attacks” and “reducing A1c levels” as top priority diabetes outcome measures.
- Nonpatient stakeholders rated “reducing A1c levels” and “reducing risk of hospitalizations from diabetes” as the 2 most feasibly collected diabetes outcome measures.
- There was a statistically significant difference between patient and nonpatient stakeholders in the selection of several outcome measures as most meaningful.

The costs of insulin and other antidiabetic medications for the treatment of type 2 diabetes have risen dramatically over the past 2 decades, contributing to the consistent placement of diabetes as one of the top expenditures among therapeutic drug classes.¹⁻³ Traditionally, prices for diabetes medications have been based on oppositional negotiations between payers and manufacturers, often facilitated by pharmacy benefit managers. Amidst the transforming health care landscape where payment to providers for health care services is increasingly being linked to outcomes, however, there has been increased interest to better align incentives for payers and manufacturers by similarly tying medication payments to outcomes through value-based contracts. Value-based contracts between payers and manufacturers link drug payments to performance as measured by predefined patient outcomes

and require shared risk between both entities.⁴ Selection of clinical outcomes that are meaningful to all stakeholders and feasibly measurable via existing claims and medical data remains a challenge in the development of effective value-based contracts.⁴

Two value-based contracts for diabetes medications within the United States have been publicized over the past few years, one of which based pricing on whether the drug improved patient hemoglobin A1c (A1c) levels, and the other included manufacturer rebates if patients required further therapy due to drug ineffectiveness.^{5,6} Many have argued that the opacity of terms in value-based contracts serves as a barrier to shared learning and improvement in contract development.⁷ While a number of studies have used the Delphi method to reach provider expert consensus on clinical decision making and indicators in diabetes, to our knowledge, none have used this method to identify meaningful outcome measures among relevant stakeholders for the specific purpose of value-based contract development.⁸⁻¹¹

In this study, we incorporated multiple stakeholder perspectives, including those of patients, primary care physicians (PCPs), endocrinologists, payers, pharmacy benefit managers (PBMs), and pharmaceutical company representatives (PCRs), in a modified Delphi method to identify relevant and measurable diabetes clinical outcomes to inform value-based contracts for diabetes medications.

Methods

We conducted an electronic-based modified Delphi survey using widely accepted methods.^{12,13} The Delphi technique enables a purposeful sample of experts from multiple domains to reach consensus through an anonymous questionnaire process that prevents commandeering of the discussion by 1 or a few individuals.¹² After each respective round, a facilitator compiles the survey responses and incorporates the results into the design of the next round's survey, at which time participants are allowed to revise their previous answers in light of prevailing group opinions. Whereas a traditional Delphi study begins the first round with an open-ended question to gather a broad list of ideas from participants, this study was performed using a modified Delphi method in which an initial list of outcome measures was generated in advance of study onset. This study was deemed exempt by the University of Pittsburgh Institutional Review Board.

Selection of Outcome Measures

A comprehensive literature search was conducted by our study team to identify pertinent type 2 diabetes indicators used to assess quality of diabetes care.¹⁴⁻¹⁷ Following this literature review, a list of 12 diabetes-associated clinical outcomes was reviewed and unanimously achieved face validity without modification by an internal, multidisciplinary workgroup

that included PCPs, pharmacists, diabetes care managers, and 2 endocrinologists specializing in diabetes care.

Selection of Participants

Participants were selected to represent multiple value-based contract stakeholder perspectives, with an intentional majority stake given to patients and providers, since their viewpoints are often underrepresented in typical value-based contracts between payers and manufacturers.¹⁸ We targeted a diverse patient population with regard to duration of time since diabetes diagnosis, current medications used, and self-reported A1c. All patient participants were involved with the UPMC Health Plan Pharmacy Care Management Program, an outreach initiative focused on improving treatment adherence to reduce unplanned care among type 2 diabetes patients. A pharmacy care manager contacted patients by phone to inform them of the research study, answer any related questions, and obtain verbal consent for participation.

PCPs and endocrinologists with expertise in diabetes care from 2 large, distinct health care systems were recruited by email, as were representatives from 2 separate payer entities, 2 diabetes pharmaceutical manufacturers, and 1 PBM organization. All nonpatient panelists were identified as having experience in diabetes by either affiliate account representatives (PBMs and PCRs), clinical departmental leaders (PCPs and endocrinologists), or an internal workgroup (payers involved in care management initiatives for members with type 2 diabetes).

Patients were paid \$20 upon completion of the first-round survey and \$30 upon completion of the second-round survey. Other stakeholders were not financially compensated.

Survey Administration

Participants opted to complete either an online questionnaire via the SurveyMonkey platform or a paper questionnaire via U.S. mail (Appendix, available in online article). A unique identifier was assigned to each participant to ensure anonymity. We sent biweekly reminder emails to encourage participants to complete the survey within a 2-week time frame, and follow-up phone calls were made on a case-by-case basis to improve compliance. All data were collected between March and April 2018.

Round 1. The first-round questionnaire was prefaced by an introduction that provided information about the purpose of the study and use of the Delphi method to achieve study goals. Participants were asked to identify which stakeholder group they represented, informing them as to whether they were to complete 1 or both sections of the questionnaire.

In the first section, participants were asked to rate each of 12 diabetes outcome measures using a 5-point Likert scale, with 1 indicating that a respective measure was "not important" and 5 indicating that a respective measure was "very

FIGURE 1 Flowchart of Participants and Response Rates

PBMs = pharmacy benefit managers; PCRs = pharmaceutical company representatives.

important” in terms of their value in providing information about disease status. A plain language definition for each outcome measure was provided, as was a blank response field to collect additional outcome measures that participants deemed important but were not included in the original outcome list. In anticipation that stakeholders might rate all outcomes as “important” or “very important,” we then requested that participants select the 3 most meaningful outcomes to them.

Nonpatient stakeholders were then additionally prompted to complete the second section of the survey, which focused on the feasibility of collecting each outcome measure either through claims or electronic health record (EHR) data. Participants were asked to rank each item using a 5-point Likert scale, with 1 indicating that a respective outcome collection was “not at all feasible” and 5 indicating that a respective outcome collection was “very feasible.”

Round 2. Participants who completed the first-round survey were e-mailed or mailed the second-round survey. In this survey, the 12 outcome measures were listed in the same order as in the first-round survey. Based on previous Delphi study methodology, the collective group’s round 1 rating of each outcome measure was provided, alongside the individual’s

original round 1 rating.^{12,19} For example, if in round 1 a participant rated a given outcome as “important” and the aggregate mean for that outcome was “4.72,” the outcome in the second round would be followed by this text: “Group average from Round 1 was 4.72; Your response was 4.” Participants were then given the opportunity to modify their previous responses based on this new information.

Likewise, for the follow-up question asking participants to rank their 3 most meaningful outcomes, previous individual and popular group choices from the first-round survey were indicated next to each option in the second-round survey. For example, text may have indicated the following: “This option was the second most popular from the group from Round 1; You also chose this answer.”

Consensus and Statistical Analyses

Based on the literature, consensus was defined as $\geq 75\%$ agreement by participants on the importance or feasibility of an outcome measure (Likert scale scores 4 and 5) or selection of an outcome as most meaningful by $\geq 75\%$ of participants.²⁰ A 2-sample test of proportions was performed to examine differences between patient and nonpatient stakeholder rankings of outcomes.

TABLE 1 Importance of Diabetes Outcome Measures

Outcome Measures ^a	Not Important/Slightly Important (1-2), n (%)		Moderately Important (3), n (%)		Important/Very Important (4-5), n (%)	
	Round 1	Round 2	Round 1	Round 2	Round 1	Round 2
Reducing risk of heart attacks	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	29 (100.0)	26 (100.0)
Reducing hemoglobin A1c levels	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	29 (100.0)	26 (100.0)
Avoiding hypoglycemic events/episodes of low blood sugar	0 (0.0)	0 (0.0)	1 (3.4)	0 (0.0)	28 (96.6)	26 (100.0)
Reducing risk of stroke	0 (0.0)	0 (0.0)	0 (0.0)	1 (3.9)	29 (100.0)	25 (96.2)
Reducing risk of diabetes-related amputations and foot ulcers	0 (0.0)	1 (3.9)	1 (3.4)	0 (0.0)	28 (96.6)	25 (96.2)
Reducing risk of diabetes-related kidney disease	0 (0.0)	0 (0.0)	1 (3.4)	1 (3.9)	28 (96.6)	25 (96.2)
Reducing risk of diabetes-related neuropathy/nerve damage	0 (0.0)	0 (0.0)	1 (3.4)	1 (3.9)	28 (96.6)	25 (96.2)
Reducing risk of emergency room visits from diabetes	0 (0.0)	0 (0.0)	1 (3.4)	1 (3.9)	28 (96.6)	25 (96.2)
Reducing risk of diabetes-related retinopathy/eye disease ^b	1 (3.6) ^b	0 (0.0)	0 (0.0) ^b	1 (3.9)	27 (96.4) ^b	25 (96.2)
Reducing risk of hospitalizations from diabetes	0 (0.0)	0 (0.0)	1 (3.4)	2 (7.7)	28 (96.6)	24 (92.3)
Improving quality of life	0 (0.0)	0 (0.0)	5 (17.2)	4 (15.4)	24 (82.8)	22 (84.6)
Weight loss	3 (10.3)	0 (0.0)	5 (17.2)	6 (23.1)	21 (72.4)	20 (76.9)

Note: Participants responded to the following prompt: "Below is a list of diabetes measures currently used to monitor diabetes. Please rate each of the measures in terms of their value in providing information about the status of diabetes."

^aFor all the measures listed, consensus ($\geq 75\%$ agreement) was reached after round 2.

^bOne participant (PCR) did not rate this item in round 1, so this item received 28 total responses rather than 29.

PCR=pharmaceutical company representative.

Results

Thirty-three individuals, including 12 (36%) patients, 12 (36%) providers, 4 (12%) payers, 3 (9%) PBMs, and 2 (6%) PCRs, conveyed interest in the study and were sent the first-round questionnaire (Figure 1). Of these, 29 (88%) responded to the first-round survey, and 26 of those 29 (90%) completed the second-round survey. The final expert panel included 9 (35%) patients, 9 (35%) providers, 3 (11%) payers, 3 (11%) PBMs, and 2 (8%) PCRs. There was diversity among the 9 patients regarding duration of time since diabetes diagnosis (range=0.5 to >20 years), current medications used (insulin, 56%; metformin, 67%; sulfonylurea 33%; sodium/glucose cotransporter 2 inhibitor, 33%; dipeptidyl peptidase 4 inhibitor, 11%), and self-reported A1c levels (range=6.1-13.8). Provider experts included 5 endocrinologists and 4 PCPs, each with at least 5 years of clinical diabetes care experience. No participants submitted additional indicators to include, and no outcomes were added or dropped between rounds. Consensus was reached after 2 rounds.

Value of Outcome Measures

After 2 Delphi rounds, all 12 diabetes outcome measures were rated as "important" or "very important" (4 or 5 on the Likert scale) by $\geq 75\%$ of participants (Table 1). Three of these outcomes ("reducing risk of heart attacks," "reducing A1c levels," and "avoiding hypoglycemic events/episodes of low blood sugar") received a rating of 4 or 5 by all 26 stakeholders in round 2 (100% agreement; Table 1). When asked to choose only the top 3 most meaningful outcomes, 84% of participants selected "reducing risk of heart attacks," reaching consensus

after the second round (Table 2, last column). Although consensus was not reached for any other measure in this group, "reducing A1c levels" ranked second (68% agreement) as a meaningful outcome (Table 2), aligning with the rankings of importance shown in Table 1.

A 2-sample test of proportions indicated a statistically significant difference between patient and nonpatient stakeholders in the selection of several outcomes as most meaningful. Specifically, 82.3% of nonpatient stakeholders chose "reducing A1c levels" as a most meaningful outcome, compared with 37.5% of patients ($P=0.03$; Table 2). Conversely, 50% of patients selected "reducing risk of diabetes-related kidney disease" as a most meaningful outcome, compared with just 11.8% of nonpatient stakeholders ($P=0.03$; Table 2).

Agreement increased between round 1 and round 2 for the top 4 ranked outcome measures (Table 2): "reducing risk of heart attacks," "reducing A1c levels," "reducing risk of stroke," and "reducing risk of hospitalizations from diabetes." Increased selection of these items was accompanied by decreased selection of several other outcomes, most notably "avoiding hypoglycemic events/episodes of low blood sugar," "quality of life," and "weight loss," by patients, providers, and PBMs in particular (Table 2).

Feasibility of Outcome Measures

Nonpatient stakeholders ranked "reducing A1c levels" and "reducing risk of hospitalizations from diabetes" as the 2 most feasibly collected diabetes outcome measures, with 93.8% agreement on each after 2 rounds (Table 3). The expert panel reached consensus on 4 additional clinical outcomes,

TABLE 2 Most Meaningful Diabetes Outcome Measures

Outcome Measures	Patients, n (%)		Providers, n (%)		Payers, n (%)		PBMs, n (%)		PCRs, n (%)		Total, n (%)	
	R1 (n = 10)	R2 (n = 8) ^a	R1 (n = 11)	R2 (n = 9)	R1 (n = 3)	R2 (n = 3)	R1 (n = 3)	R2 (n = 3)	R1 (n = 2)	R2 (n = 2)	R1 (N = 29)	R2 (N = 25) ^a
Reducing risk of heart attacks ^b	8 (80.0)	7 (87.5)	5 (45.5)	7 (77.8)	3 (100.0)	3 (100.0)	1 (33.3)	2 (66.7)	2 (100.0)	2 (100.0)	19 (65.5)	21 (84.0)
Reducing hemoglobin A1c levels ^c	4 (40.0)	3 (37.5)	7 (63.6)	8 (88.9)	2 (66.7)	2 (66.7)	3 (100.0)	3 (100.0)	1 (50.0)	1 (50.0)	17 (58.6)	17 (68.0)
Reducing risk of stroke	5 (50.0)	3 (37.5)	3 (27.2)	4 (44.4)	1 (33.3)	2 (66.7)	1 (33.3)	2 (66.7)	1 (50.0)	2 (100.0)	11 (37.9)	13 (52.0)
Reducing risk of hospitalizations from diabetes	0 (0.0)	1 (12.5)	2 (18.2)	3 (33.3)	0 (0.0)	1 (33.3)	1 (33.3)	1 (33.3)	2 (100.0)	1 (50.0)	5 (17.2)	7 (28.0)
Reducing risk of diabetes-related kidney disease ^c	3 (30.0)	4 (50.0)	3 (27.2)	1 (11.1)	1 (33.3)	0 (0.0)	1 (33.3)	1 (33.3)	0 (0.0)	0 (0.0)	8 (27.6)	6 (24.0)
Avoiding hypoglycemic events/episodes of low blood sugar	1 (10.0)	0 (0.0)	5 (45.5)	3 (33.3)	0 (0.0)	0 (0.0)	1 (33.3)	0 (0.0)	0 (0.0)	0 (0.0)	7 (24.1)	3 (12.0)
Improving quality of life	3 (30.0)	0 (0.0)	5 (45.5)	1 (11.1)	1 (33.3)	1 (33.3)	1 (33.3)	0 (0.0)	0 (0.0)	0 (0.0)	10 (34.5)	2 (8.0)
Reducing risk of diabetes-related amputations and foot ulcers	2 (20.0)	2 (25.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	2 (6.9)	2 (8.0)
Reducing risk of diabetes-related retinopathy/eye disease	0 (0.0)	2 (25.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	2 (8.0)
Weight loss	2 (20.0)	1 (12.5)	3 (27.2)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	5 (17.2)	1 (4.0)
Reducing risk of diabetes-related neuropathy/nerve damage	2 (20.0)	1 (12.5)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	2 (6.9)	1 (4.0)
Reducing risk of emergency room visits from diabetes	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	1 (33.3)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	1 (3.5)	0 (0.0)

Note: Participants responded to the following prompt: "If we had to limit the number of diabetes measures to 3 that are most meaningful, which ones would you choose to keep (please select only 3)?"

^aOne patient did not complete this question in round 2, so there were 25 total responses rather than 26.

^bConsensus ($\geq 75\%$ agreement) was reached after round 2.

^cStatistically significant difference found between patient and nonpatient stakeholder groups; $P=0.03$ for each of 2 outcomes noted.

PBMs=pharmacy benefit managers; PCRs=pharmaceutical company representatives; R1/R2=round 1/round 2.

rating them as feasible or very feasible to collect, including "weight loss" (87.5%), "reducing risk of diabetes-related kidney disease" (81.3%), "reducing risk of emergency room visits from diabetes" (75%), and "reducing risk of diabetes-related amputations and foot ulcers" (75%). Notably, "reducing risk of heart attacks" did not reach consensus as a feasible or very feasible outcome measure (43.8%).

Discussion

To our knowledge, this is the first Delphi-based study conducted with multiple stakeholders to select meaningful and measurable clinical diabetes outcomes for the development of pharmaceutical value-based contracts. After 2 rounds, our diverse expert panel reached consensus in ranking "reducing risk of heart attacks" as the most meaningful diabetes outcome measure and in rating "reducing A1c levels" and "reducing the risk of hospitalizations from diabetes" equally as the most feasibly collected diabetes outcome measures. The measure "reducing A1c levels" scored as the second most meaningful measure;

however, consensus was not reached when participants were asked to force-rank their top 3 choices.

Byrne et al. (2017) conducted a similar type of Delphi study to identify a core set of outcomes that could be measured and reported for clinical trials for type 1 diabetes in young adults.¹⁹ Like our study, Byrne et al. incorporated a 2-round Delphi survey with multiple stakeholders (patients, health professionals, researchers, and policymakers); however, key value-based contracting stakeholders (payers, PBMs, and PCRs) were omitted. Their consensus panel of 12 individuals identified "measures of diabetes-related burden or stress" and "diabetes-related quality of life" as the top 2 important outcomes, but Byrne et al. noted that future work would be needed to determine the best way to accurately measure these variables, a key requirement of effective value-based contract implementation. Six other outcomes reached consensus in Byrne et al., including "number of events of severe hypoglycemia" (91.7%) and "objectively measured A1c" (83.3%), representing more commonly measured clinical disease indicators.¹⁹ The measure "reducing risk of heart attacks" was not

TABLE 3 Feasibility of Diabetes Outcome Measures

Outcome Measures	Not at All Feasible/Slightly Feasible (1-2), n (%)		Moderately Feasible (3), n (%)		Feasible/Very Feasible (4-5), n (%)	
	Round 1	Round 2	Round 1	Round 2	Round 1	Round 2
Reducing hemoglobin A1c levels (measured by lab values in EHR) ^a	1 (5.6)	0 (0.0)	2 (11.1)	1 (6.3)	15 (83.3)	15 (93.8)
Reducing risk of hospitalizations from diabetes (measured by claims and/or EHR) ^a	1 (5.6)	0 (0.0)	4 (22.2)	1 (6.3)	13 (72.2)	15 (93.8)
Weight loss (measured by EHR) ^a	3 (16.7)	1 (6.3)	2 (11.1)	1 (6.3)	13 (72.2)	14 (87.5)
Reducing risk of diabetes-related kidney disease (measured by claims and/or EHR) ^a	1 (5.6)	1 (6.3)	3 (16.7)	2 (12.5)	14 (77.8)	13 (81.3)
Reducing risk of emergency room visits from diabetes (measured by claims and/or EHR) ^a	0 (0.0)	0 (0.0)	5 (27.8)	4 (25.0)	13 (72.2)	12 (75.0)
Reducing risk of diabetes-related amputations and foot ulcers (measured by claims and/or EHR) ^a	1 (5.6)	1 (6.3)	6 (33.3)	3 (18.8)	11 (61.1)	12 (75.0)
Reducing risk of diabetes-related retinopathy/eye disease (measured by claims and/or EHR)	1 (5.6)	1 (6.3)	4 (22.2)	6 (37.5)	13 (72.2)	9 (56.3)
Reducing risk of heart attacks (measured by claims and/or EHR)	5 (27.8)	3 (18.8)	5 (27.8)	6 (37.5)	8 (44.4)	7 (43.8)
Reducing risk of stroke (measured by claims and/or EHR)	6 (33.3)	2 (12.5)	5 (27.8)	8 (50.0)	7 (38.9)	6 (37.5)
Avoiding hypoglycemic events/episodes of low blood sugar (measured by claims and/or EHR)	2 (11.1)	2 (12.5)	7 (38.9)	9 (56.3)	9 (50.0)	5 (31.3)
Reducing risk of diabetes-related neuropathy/nerve damage (measured by claims and/or EHR)	3 (16.7)	3 (18.8)	9 (50.0)	10 (62.5)	6 (33.3)	3 (18.8)
Improving quality of life (measured by EHR)	8 (44.4)	9 (56.3)	5 (27.8)	5 (31.3)	5 (27.8)	2 (12.5)

Note: Participants responded to the following prompt: "Please rate each measure below on how feasible they are to collect in an integrated delivery and finance system where both administrative claims and electronic health record (EHR) data are accessible."

^aConsensus ($\geq 75\%$ agreement) was reached after round 2.

EHR = electronic health record.

included in their core set of outcomes, possibly because it is less of an imminent concern in a younger type 1 diabetes population.

Several studies have used the Delphi method to reach provider expert consensus on clinical decision making and indicators for type 2 diabetes, but none have focused on identifying meaningful and measurable outcomes to incorporate into value-based contracts.⁸⁻¹¹ In 1 study, a Canadian research team used a modified Delphi method with a panel of 12 experts (which included diabetes clinicians, researchers, and policy advisors, but not patients or other value-based contract stakeholders) to identify important diabetes indicators as part of a larger, programmatic evaluation tool.⁷ The Canadian team reached consensus on an extensive, "all-inclusive" set of 33 indicators that could be used to guide comprehensive evaluation of diabetes programs, including several relevant health status and quality of life measures that were included in our study (i.e., A1c, kidney disease, and cardiovascular complications). While these 33 indicators serve an important role in the intended context, their feasibility of measurement was not addressed.

When developing value-based contracts, transparency is important in justifying the selection of specific outcomes, building trust between payers and manufacturers, and enabling shared learning with others. By using the Delphi method to identify outcomes for value-based contracting, we demonstrated the use of research methodology to inform business practice. It

is unclear how previously reported value-based contracts have selected their outcomes, and a number of value-based contracts have focused on surrogate endpoints often used in the conduct of clinical trials that may not be valued by or of importance to patients. To this point, 1 study found that only 46% of registered diabetes trials included "patient-important" outcomes (defined as stroke, myocardial infarction, amputation, loss of vision, end-stage renal disease, hypoglycemic events, delayed wound healing, infection, visual disturbances, pain, and functional status) as a primary or secondary outcome.¹⁷

A1c measurement is a common surrogate endpoint used for diabetes monitoring and often serves as a primary outcome in diabetes clinical trials.¹⁷ Our study results reflect the perceived value of this outcome measure, since all 26 stakeholders rated "reducing A1c levels" as important/very important on the Likert scale, and consensus was reached among nonpatient stakeholders on its feasibility to collect. However, our finding that a significantly greater proportion of nonpatient stakeholders ranked "reducing A1c levels" as a most meaningful outcome measure compared with patients should be taken into consideration when developing innovative value-based contracts. Interestingly, more patients selected "reducing risk of diabetes-related kidney disease" as a most meaningful outcome than "reducing A1c levels."

Although a significantly smaller proportion of nonpatient stakeholders deemed "reducing risk of diabetes-related kidney

disease” as a most meaningful outcome compared with patient stakeholders, consensus was reached on its importance and feasibility using the Likert scale, suggesting the potential value of incorporating this measure into value-based contracts for diabetes medications.

The measure “reducing risk of heart attacks” was the only outcome to reach consensus as a most meaningful measure among all stakeholders but was not generally perceived as feasible to collect. It is possible that participants interpreted “risk” as a more complex assessment that is not readily available through existing data, whereas phrasing the outcome as “reducing heart attacks” may have elicited a different response, since that information is readily available in administrative claims and EHR data. Notably, EHR data can be used to calculate atherosclerotic cardiovascular disease (ASCVD) risk, a measure that predicts risk of heart attacks. If calculated ASCVD scores were incorporated into standard EHR readouts and were readily available to patients, providers, payers, and other value-based contracting stakeholders, this might strengthen the feasibility of data collection for this meaningful diabetes outcome measure.

In observing the changes in stakeholder responses between round 1 and round 2, the strength of this anonymous, multi-round Delphi technique is demonstrated. For example, 34.5% of participants ranked “improving quality of life” as a most meaningful measure in round 1. However, likely due to receipt of feedback that the collective group ranked “improving quality of life” second-to-last in importance in question 1 of the first-round survey, 7 participants opted to select a different outcome as 1 of their top 3 most meaningful measures in the second-round survey. This result exemplifies the purpose of the Delphi method—allowing individuals to consider the prevailing group opinion and reevaluate their initial response, with the ultimate goal of achieving collective consensus among multiple stakeholders with diverse backgrounds and viewpoints.

Based on the findings of our Delphi study that incorporated the viewpoints of a diverse stakeholder panel, several diabetes outcome measures may be effectively used for the development of pharmaceutical value-based contracts that aim to improve patient outcomes while reducing costs.

Limitations

Our study has some limitations to consider. An inherent limitation was that all patient participants were recruited from our internal Pharmacy Care Management Program, potentially introducing selection bias among our patient panelists. For example, this care management program outreaches to patients who are considered nonadherent with their diabetes medications based on administrative claims data; therefore, this cohort may have poorer disease control than the general diabetes population. On the other hand, patients who participated in the care management program and responded to our survey may be more engaged than others. The directional effect of this potential selection bias is therefore unclear. Nevertheless, we made a focused effort to include diverse patients by recruiting those who were diagnosed with diabetes relatively recently, as

well as patients living with diabetes for many years, patients with varying degrees of glycemic control, and patients taking medications from different drug classes.

Another limitation was that all providers on the panel were based in southwestern Pennsylvania. While there was a lack of geographic diversity among providers, PCPs and endocrinologists were included in the study to capture potential differences in opinion between generalists and specialists with expertise in diabetes care. Furthermore, not all participants who originally expressed interest in the study responded to both rounds of the survey; therefore, a subset of viewpoints were potentially not represented in the final results. Notably, our response rates of 88% in round 1 and 90% in round 2 are consistent with what is typically reported in the literature, and participation was sufficient to achieve our study goals.¹²

Finally, while our Delphi panel was similarly sized, if not larger, as those described in other published studies, it included a relatively small number of nonpatient, nonprovider representatives, and it is possible that had we surveyed larger groups of individuals in each stakeholder group, our findings may have changed.¹² Nevertheless, we intentionally oversampled patients and providers proportional to other stakeholder groups, since patient and provider viewpoints are often underrepresented in typical value-based contracts between payers and manufacturers.

Conclusions

It is important that value-based contracts include outcomes that are identified as meaningful and measurable by key stakeholders. In this study, we demonstrated use of the Delphi method to inform business practice, where a panel of experts indicated that “reducing risk of heart attacks” and “reducing A1c levels” are the most valuable diabetes outcome measures, among several other priority outcomes. These findings should inform development of effective value-based contracts that will improve patient outcomes while lowering costs of diabetes drugs.

Authors

LYNN M. NEILSON, PhD; ELIZABETH C.S. SWART, BS; and WILLIAM H. SHRANK, MD, MSHS, Center for Value-Based Pharmacy Initiatives, UPMC Health Plan, Pittsburgh, Pennsylvania. CHESTER B. GOOD, MD, MPH, and NATASHA PAREKH, MD, MS, Center for Value-Based Pharmacy Initiatives, UPMC Health Plan, and Division of General Internal Medicine, University of Pittsburgh, Pittsburgh, Pennsylvania. ROCHELLE HENDERSON, PhD, Express Scripts, St. Louis, Missouri, and CHRONIS MANOLIS, RPh, Center for Value-Based Pharmacy Initiatives and Pharmacy Division, UPMC Health Plan, Pittsburgh, Pennsylvania.

AUTHOR CORRESPONDENCE: Natasha Parekh, MD, MS, Center for Value-Based Pharmacy Initiatives, UPMC Health Plan, 600 Grant St., 40th Fl., Pittsburgh, PA 15219. Tel.: 772.708.1047; E-mail: parekh@upmc.edu.

DISCLOSURES

Express Scripts provided research funding for this study to the UPMC Center for Value-Based Pharmacy Initiatives. Henderson is employed by Express Scripts and was involved in the conception and design of the study and manuscript approval. The other authors are employed by the UPMC Center for Value-Based Pharmacy Initiatives and have nothing to disclose.

REFERENCES

1. Hua X, Carvalho N, Tew M, et al. Expenditures and prices of anti-hyperglycemic medications in the United States: 2002-2013. *JAMA*. 2016;315(13):1400-02.
2. Zhuo X, Zhang P, Kahn HS, et al. Change in medical spending attributable to diabetes: national data from 1987 to 2011. *Diabetes Care*. 2015;38(4):581-87.
3. Express Scripts. 2017 drug trend report. Available at: <http://lab.express-scripts.com/lab/drug-trend-report/2017-dtr>. Accessed January 5, 2019.
4. AMCP Partnership Forum: advancing value-based contracting. *J Manag Care Spec Pharm*. 2017;23(11):1096-102. Available at: <https://www.jmcp.org/doi/10.18553/jmcp.2017.17342>.
5. Carlson JJ, Sullivan SD, Garrison LP, Neumann PJ, Veenstra DL. Linking payments to health outcomes: a taxonomy and examination of performance-based reimbursement schemes between healthcare payers and manufacturers. *Health Policy*. 2010;96(3):179-90.
6. Yu JS, Chin L, Oh J, Farias J. Performance-based risk-sharing arrangements for pharmaceutical products in the United States: a systematic review. *J Manag Care Spec Pharm*. 2017;23(10):1028-40. Available at: <https://www.jmcp.org/doi/10.18553/jmcp.2017.23.10.1028>.
7. Nazareth T, Ko JJ, Sasane R, et al. Outcomes-based contracting experience: research findings from U.S. and European stakeholders. *J Manag Care Spec Pharm*. 2017;23(10):1018-26. Available at: <https://www.jmcp.org/doi/10.18553/jmcp.2017.23.10.1018>.
8. Paquette-Warren J, Tyler M, Fournie M, Harris SB. The Diabetes Evaluation Framework for Innovative National Evaluations (DEFINE): construct and content validation using a modified Delphi method. *Can J Diabetes*. 2017;41(3):281-96.
9. Avogaro A, Giaccari A, Fioretto P, et al. A consensus statement for the clinical use of the renal sodium-glucose co-transporter-2 inhibitor dapagliflozin in patients with type 2 diabetes mellitus. *Expert Rev Clin Pharmacol*. 2017;10(7):763-72.
10. Calvo Gómez C, Gómez-Huelgas R, Lima Ruiz J, et al. Expert consensus on the use of liraglutide in the treatment of diabetes in internal medicine. *Int J Clin Pharmacol Ther*. 2016;54(11):880-90.
11. Magnan EM, Gittelsohn R, Bartels CM, et al. Establishing chronic condition concordance and discordance with diabetes: a Delphi study. *BMC Fam Pract*. 2015;16:42.
12. Boulkedid R, Abdoul H, Loustau M, Sibony O, Albeti C. Using and reporting the Delphi method for selecting healthcare quality indicators: a systematic review. *PLoS One*. 2011;6(6):e20476.
13. Hsu C, Sandford BA. The Delphi technique: making sense of consensus. *Pract Assess, Res & Eval*. 2007;12(10). Available at: <https://pareonline.net/getvn.asp?v=12&n=10>. Accessed January 5, 2019.
14. O'Connor PJ, Bodkin NL, Fradkin J, et al. Diabetes performance measures: current status and future directions. *Diabetes Care*. 2011;34(7):1651-59.
15. Calsbeek H, Ketelaar NA, Faber MJ, Wensing M, Braspenning J. Performance measurements in diabetes care: the complex task of selecting quality indicators. *Int J Qual Health Care*. 2013;25(6):704-09.
16. Aron DC. Quality indicators and performance measures in diabetes care. *Curr Diab Rep*. 2014;14(3):472.
17. Gandhi GY, Murad MH, Fujiyoshi A, et al. Patient-important outcomes in registered diabetes trials. *JAMA*. 2008;299(21):2543-49.
18. Seeley E, Kesselheim AS. Outcomes-based pharmaceutical contracts: an answer to high U.S. drug spending? *Issue Brief (Commonw Fund)*. 2017;2017:1-8.
19. Byrne M, O'Connell A, Egan AM, et al. A core outcomes set for clinical trials of interventions for young adults with type 1 diabetes: an international, multi-perspective Delphi consensus study. *Trials*. 2017;18(1):602.
20. Diamond IR, Grant RC, Feldman BM, et al. Defining consensus: a systematic review recommends methodologic criteria for reporting of Delphi studies. *J Clin Epidemiol*. 2014;67(4):401-09.

APPENDIX Mailed Delphi Survey: Round 1

The purpose of this research is to evaluate diabetes measures currently used to monitor diabetes and to understand which measures are the most meaningful and valuable to patients living with diabetes. We will be consulting a panel of people considered to be experts on diabetes, and will ask them to complete a brief online survey (below). This survey will use the Delphi Method, where experts answer questionnaires in two or more rounds. If you are willing to participate, the survey will ask about different diabetes measures and how different measures are perceived by the participant.

1. Are you a patient, caregiver, provider, payer, pharmaceutical company representative, or pharmacy benefits manager representative (Please select one)?

☐ Patient	☐ Payer
☐ Provider-Endocrinologist	☐ Pharmaceutical Company Representative
☐ Provider-Primary Care Physician	☐ Pharmacy Benefits Manager Representative

2. Below is a list of diabetes measures currently used to monitor diabetes. Please rate each of the measures in terms of their value in providing information about the status of diabetes.

	Not Important (1)	Slightly Important (2)	Moderately Important (3)	Important (4)	Very Important (5)
Reducing hemoglobin A1c levels	☐	☐	☐	☐	☐
Reducing risk of heart attacks	☐	☐	☐	☐	☐
Reducing risk of stroke	☐	☐	☐	☐	☐
Reducing risk of diabetes-related retinopathy/eye disease	☐	☐	☐	☐	☐
Reducing risk of diabetes-related kidney disease	☐	☐	☐	☐	☐
Reducing risk of diabetes-related neuropathy/nerve damage	☐	☐	☐	☐	☐

continued on next page

APPENDIX Mailed Delphi Survey: Round 1 (continued)

2. Below is a list of diabetes measures currently used to monitor diabetes. Please rate each of the measures in terms of their value in providing information about the status of diabetes.

	Not Important (1)	Slightly Important (2)	Moderately Important (3)	Important (4)	Very Important (5)
Reducing risk of diabetes-related amputations and foot ulcers	☐	☐	☐	☐	☐
Weight loss	☐	☐	☐	☐	☐
Improving quality of life	☐	☐	☐	☐	☐
Avoiding hypoglycemic events/episodes of low blood sugar	☐	☐	☐	☐	☐
Reducing risk of hospitalizations from diabetes	☐	☐	☐	☐	☐
Reducing risk of emergency room visits from diabetes	☐	☐	☐	☐	☐

*** 3. If we had to limit the number of diabetes measures to 3 that are most meaningful, which ones would you choose to keep (please select only 3)?**

☐ Reducing hemoglobin A1c levels	☐ Reducing risk of diabetes-related amputations and foot ulcers
☐ Reducing risk of heart attacks	☐ Weight loss
☐ Reducing risk of stroke	☐ Improving quality of life
☐ Reducing risk of diabetes-related retinopathy/eye disease	☐ Avoiding hypoglycemic events/episodes of low blood sugar
☐ Reducing risk of diabetes-related kidney disease	☐ Reducing risk of hospitalizations from diabetes
☐ Reducing risk of diabetes-related neuropathy/nerve damage	☐ Reducing risk of emergency room visits from diabetes

continued on next page

APPENDIX Mailed Delphi Survey: Round 1 (continued)

4. Are there any important measures for diabetes we are missing?**5. Please rate each measure below on how feasible they are to collect in an integrated delivery and finance system where both administrative claims and electronic health record (EHR) data are accessible.**

	Not At All Feasible (1)	Slightly Feasible (2)	Moderately Feasible (3)	Feasible (4)	Very Feasible (5)
Reducing hemoglobin A1c levels (measured by lab values in EHR)	☐	☐	☐	☐	☐
Reducing risk of heart attacks (measured by claims and/or EHR)	☐	☐	☐	☐	☐
Reducing risk of stroke (measured by claims and/or EHR)	☐	☐	☐	☐	☐
Reducing risk of diabetes-related retinopathy/eye disease (measured by claims and/or EHR)	☐	☐	☐	☐	☐
Reducing risk of diabetes-related kidney disease (measured by claims and/or EHR)	☐	☐	☐	☐	☐
Reducing risk of diabetes-related neuropathy/nerve damage (measured by claims and/or EHR)	☐	☐	☐	☐	☐
Reducing risk of diabetes-related amputations and foot ulcers (measured by claims and/or EHR)	☐	☐	☐	☐	☐

continued on next page

APPENDIX Mailed Delphi Survey: Round 1 (continued)

5. Please rate each measure below on how feasible they are to collect in an integrated delivery and finance system where both administrative claims and electronic health record (EHR) data are accessible.

	Not At All Feasible (1)	Slightly Feasible (2)	Moderately Feasible (3)	Feasible (4)	Very Feasible (5)
Weight loss (measured by EHR)	☐	☐	☐	☐	☐
Improving quality of life (measured by EHR)	☐	☐	☐	☐	☐
Avoiding hypoglycemic events/episodes of low blood sugar (measured by claims and/or EHR)	☐	☐	☐	☐	☐
Reducing risk of hospitalizations from diabetes (measured by claims and/or EHR)	☐	☐	☐	☐	☐
Reducing risk of emergency room visits from diabetes (measured by claims and/or EHR)	☐	☐	☐	☐	☐