


# Core outcome set for diabetes after pregnancy prevention across the life span: international Delphi study

Nancy Wu,<sup>1</sup> Sharleen O'Reilly,<sup>2</sup> Karoline Kragelund Nielsen,<sup>3</sup>  
Helle Terkildsen Maindal,<sup>3,4</sup> Kaberi Dasgupta <sup>1,5</sup>

**To cite:** Wu N, O'Reilly S, Nielsen KK, *et al*. Core outcome set for diabetes after pregnancy prevention across the life span: international Delphi study. *BMJ Open Diab Res Care* 2020;**8**:e001594. doi:10.1136/bmjdr-2020-001594

► Supplemental material is published online only. To view, please visit the journal online (<http://dx.doi.org/10.1136/bmjdr-2020-001594>).

Received 20 May 2020  
Revised 10 September 2020  
Accepted 28 September 2020

## ABSTRACT

**Introduction** Mothers with gestational diabetes mellitus (GDM) are at high risk of future diabetes. An active area of research examines health behavior change strategies in women within 5 years of a GDM pregnancy to prevent diabetes after pregnancy. We aimed to develop a core outcome set (COS) to facilitate synthesis and comparison across trials.

**Research design and methods** Candidate outcomes were identified through systematic review and scored for importance (1–9) by healthcare professionals, researchers, and women with prior GDM through an international two-round electronic-Delphi survey. Outcomes retained required round two scores above prespecified thresholds ( $\geq 70\%$  scoring 7–9) or expert panel endorsement when scores were indeterminate. The panel organized the COS by domain.

**Results** 115 stakeholders participated in the survey and 56 completed both rounds. SD of scores decreased by 0.24 (95%CI 0.21 to 0.27) by round 2, signaling convergence. The final COS includes 19 domains (50 outcomes): diabetes (n=3 outcomes), other related diseases (n=3), complications in subsequent pregnancy (n=2), offspring outcomes (n=3), adiposity (n=4), cardiometabolic measures (n=5), glycemia (n=3), physical activity (n=2), diet (n=4), breast feeding (n=2), behavior change theory (n=5), diabetes-related knowledge (n=2), health literacy (n=1), social support (n=1), sleep (n=1), quality of life (n=1), program delivery (n=4), health economic evaluation (n=2), and diabetes risk screening (n=2). The seven outcomes endorsed by  $\geq 90\%$  were diabetes development and GDM recurrence, attending the postpartum diabetes screening and completing oral glucose tolerance testing and/or other glycemia measures, weight and total energy intake, and health behaviors in general. Among the 15 at the 80%–90% endorsement level, approximately half were specific elements related to the top 7, while the remainder related to diabetes knowledge, personal risk perception, motivation for change, program element completion, and health service use and cost.

**Conclusion** Researchers should collect and report outcomes from the breadth of domains in the COS.

## INTRODUCTION

Gestational diabetes mellitus (GDM) is a critical public health issue, having both a high incidence and demonstrated associations with diabetes after pregnancy (DAP) in

## Significance of this study

### What is already known about this subject?

► Health behavior change can reduce diabetes incidence in women who are at high risk reflected by a history of gestational diabetes mellitus (GDM). An increasing number of trials are testing interventions soon after pregnancy to accomplish this. Researchers are measuring a variety of different outcomes.

### What are the new findings?

► Healthcare providers, researchers, and women with GDM specified a core outcome set (COS) that moves beyond biological measures to include behavioral, contextual, process, and economic measures.

### How might these results change the focus of research or clinical practice?

► By addressing the domains highlighted by this COS, strategies can be refined iteratively to achieve approaches that are both effective and feasible to translate into programs and practice for diabetes prevention.

mothers, offspring, and even fathers.<sup>1,2</sup> The US Diabetes Prevention Program trial proved that DAP in mothers is preventable through health behavior change in women whose pregnancy averaged a decade before enrollment.<sup>3</sup> However, most incident diabetes cases occur within 5 years of GDM.<sup>4</sup> We aimed to identify a core outcome set (COS) for the increasing number of trials testing health behavior change interventions for DAP prevention sooner after pregnancy.

A COS represents the standard set of outcomes to measure for a given population, in a particular field of interest. It facilitates comparison and synthesis, and allows for iterative advancement. In October 2017, we launched the international Core Outcome Set for Diabetes after Pregnancy Prevention Trials (COS-DAP) initiative. Our focus was the development and testing of interventions



© Author(s) (or their employer(s)) 2020. Re-use permitted under CC BY-NC. No commercial re-use. See rights and permissions. Published by BMJ.

For numbered affiliations see end of article.

### Correspondence to

Dr Kaberi Dasgupta;  
kaberi.dasgupta@mcgill.ca

that are efficacious for DAP prevention, compatible with the needs of women with young families, and feasible to be translated into programs. In accordance with Core Outcome Measures in Effectiveness Trials (COMET) guidelines, we registered and published both our COS development protocol and the systematic review from which we derived candidate outcomes.<sup>5–7</sup> We herein report the findings from our two-round electronic (e)-Delphi survey among a large international stakeholder group representing healthcare professionals (HCPs), researchers, and women with prior GDM; and the final COS.

COS development is a relatively new field and, as such, is subject to methodological debate. Concurrent with our initiative, another group developed a COS with a similar target population but a more general scope of evaluating metrics to monitor following a GDM pregnancy; subsequent to their e-Delphi survey, they removed a large number of highly endorsed outcomes at a consensus panel meeting in the interest of brevity, retaining indicators of glycemia, blood pressure, and breastfeeding status.<sup>8</sup> In contrast, we applied a priori scoring criteria for inclusion based on e-Delphi results, with only indeterminate outcomes voted on by a consensus panel. We thus present a COS that maintains fidelity with the e-Delphi survey, capturing biological, behavioral, contextual, process, and economic measures. Its comprehensive scope may not only permit the detection of biological impact but also facilitate progress toward the ultimate goal of translating findings into real-world programs.

## METHODS

We registered COS-DAP with the COMET initiative.

### Identification of candidate outcomes

As previously reported, we conducted a systematic review focused on health behavior change (ie, eating and physical activity) and diabetes screening intervention studies in women with previous GDM.<sup>7</sup> Briefly, two investigators independently screened titles and abstracts, and at least three reviewed the full text of selected articles and extracted candidate outcomes verbatim. A face-to-face meeting among investigators was held in Dublin, Ireland (March 2018), for deduplication, suggestion of potentially important outcomes not identified in the review, and outcome grouping, in preparation for the e-Delphi survey.

### Participants

Researchers and HCPs are knowledgeable in GDM management, risks, and DAP prevention interventions. Women with GDM can speak to the nuances of living with GDM and the experience of outcome assessments, particularly if they have participated in DAP prevention intervention studies. We recruited from all three of these stakeholder groups. As recommended for Delphi studies,<sup>9</sup> we aimed for 20–50 respondents per group. Email invitations and social media posts were disseminated through

pregnancy and diabetes organizations at international (eg, International Federation of Gynecology and Obstetrics), national (eg, Diabetes Canada and Diabetes Ireland), and regional levels (eg, motherhood groups). HCPs and researchers were asked to invite patients enrolled in previous GDM studies. Interested individuals were provided with a link to the survey website.

### Survey

The survey was hosted on the online COMET Delphi-Manager system<sup>10</sup> in English, French, and Danish. Round 1 was open from July to September 2018, and round 2 was open from October to December 2018. Registration and consent were completed electronically; general characteristics were queried. Participants rated each outcome on an ordinal scale (1=unimportant to 9=very important) or indicated inability to score. We provided a plain language definition for each outcome. During round 1, respondents could suggest additional outcomes. All who initiated round 1 were invited to complete a second round. During round 2, participants were asked to rerate all outcomes after considering their initial ratings and average ratings from each stakeholder group, which were displayed for each outcome. We held an investigator meeting after round 1 in Halifax, Canada (October 2018), alongside the 2018 Diabetes Canada Conference to plan the consensus meeting.

### Analysis

Adopting an approach described by other COS developers and COMET,<sup>11 12</sup> we retained outcomes for which  $\geq 70\%$  of participants scored 7–9 and  $\leq 15\%$  scored 1–3. We excluded outcomes that  $\geq 70\%$  of participants scored 1–3 and  $\leq 15\%$  scored 7–9. All others were considered indeterminate and brought forward to the consensus meeting. We analyzed the change in SD of each outcome's scores from round 1 to 2, one of the approaches to quantify convergence described by the COMET initiative.<sup>11</sup> All analyses were performed in R V.3.5.1.

### Consensus meeting

The meeting was a satellite event of the 10<sup>th</sup> International Symposium on Diabetes, Hypertension, Metabolic Syndrome, and Pregnancy held in Florence, Italy (May 2019). The 15 voting consensus panel members represented eight countries (Australia, Canada, Denmark, Hong Kong (China), Ireland, Israel, Norway, and Switzerland) with experts in medicine (eg, obstetrics, endocrinology, and internal medicine), nutrition, public health, and epidemiology. All but one had not participated in the e-Delphi survey. Patient representatives were absent.

The meeting followed a modified nominal group format. Panel members discussed each indeterminate outcome, then voted anonymously via FormPlus<sup>13</sup> to accept, reject, or indicate uncertainty on its inclusion in the COS. An indeterminate outcome was retained if  $\geq 70\%$  of members voted in favor. Panel members also discussed how to group outcomes into domains

**Table 1** Characteristics of survey initiators by stakeholder and survey round\* (% of n)

Characteristic	HCP		Researcher		Patient	
	n=54	n=34	n=27	n=21	n=34	n=12
Women	75.9	73.5	77.8	76.2	100	100
Age (years)						
<30	5.6	5.8	11.1	9.5	5.9	8.3
30–50	46.3	41.1	37	71.4	58.8	83.3
>50	48.1	52.9	51.8	19	35.3	8.3
Ethnicity						
European	61.1	61.8	63	61.9	61.8	58.3
East Asian	5.6	5.9	3.7	4.8	8.8	16.7
Latin American	1.9	0	7.4	9.5	5.9	8.3
Other†	29.5	26.4	25.9	23.8	23.5	16.7
No answer	1.9	5.9	0	0	0	0
Country						
Denmark	14.8	17.6	14.8	19	14.7	0
Canada‡	16.7	14.7	0	0	58.8	75
Australia	22.2	23.5	29.6	33.3	2.9	0
Ireland	9.3	8.8	11.1	9.5	17.6	16.7
Other§	37	35.4	44.5	38.2	6	8.3
Type of HCP						
Obstetrician/gynecologist	35.2	35.3	14.8	9.5		
Dietitian	25.9	23.5	11.1	9.5		
Endocrinologist	18.5	23.5	11.1	9.5		
Nurse	5.6	2.9	3.7	4.8		
Internist	1.9	2.9	0	0		
Other¶	7.3	3.1	18.5	9.5		
No answer	5.6	8.8	40.7	57.1		

\*Left and right columns within each stakeholder group represent n in the first and second rounds of the survey, respectively. Participants who identified as researchers were not explicitly queried for a healthcare profession background.

†Indigenous, African origin, South Asian, Southeast Asian, West Asian, or a combination of ethnicities.

‡HCPs from Canada with research activity chose not to identify as researchers.

§Argentina, Austria, Brazil, Croatia, Estonia, Germany, Hong Kong, India, Iran, Israel, Italy, Mexico, Portugal, Russia, Saudi Arabia, Spain, UK, USA, and Zimbabwe.

¶Diabetes educator; maternal–fetal medicine subspecialist; physician—general/family/primary care; critical care obstetrics.

HCP, healthcare professional.

and when to measure them. Suggestions for additional outcomes were discussed over email and voted for inclusion with the same criteria, through the online platform SurveyMonkey.<sup>14</sup>

## RESULTS

The systematic review identified 172 outcomes.<sup>7</sup> Collapse of overlapping constructs yielded the 121 included in the e-Delphi survey. Round 2 included an additional 33 outcomes suggested by participants during round 1.

One hundred thirty-four individuals from 23 countries registered for the e-Delphi survey. Round 1 was initiated by 115 stakeholders (HCP n=54, researcher n=27, women with prior GDM n=34). Across stakeholder groups, most were women >30 years of age and of European ethnicity

(table 1). Women with a GDM history were a median of 4 years past their last GDM pregnancy. Over half of them had  $\geq 2$  pregnancies, and 12% had developed type 2 diabetes in the years following the GDM pregnancy. Of the HCPs, 35% were obstetricians/gynecologists; 26% were dietitians; and 19% were endocrinologists. Over 50% of HCPs spent the majority of time on clinical activities. Among the researchers, 60% dedicated the majority of their time to research activities; some reported a clinical background in obstetrics/gynecology (15%), dietetics (11%), endocrinology (11%), and nursing (3.7%). Sixty-seven people initiated round 2, and 56 completed both rounds in entirety (49% of survey initiators).

The median SD of scores in round 1 was 1.84 (IQR 1.62–2.17) and that in round 2 was 1.61 (IQR 1.33–1.91).

The mean change in SD from round 1 to 2 was  $-0.24$  (95% CI  $-0.27, -0.21$ ). Forty-six outcomes met inclusion criteria by the end of round 2: among these, 22% related to behavior change and context, 17% to diet, 11% to glycemia and diabetes development, 7% to physical activity, and the remainder to a variety of other concepts (table 2). The 75 indeterminate outcomes were discussed and 8 were retained. Seven additional candidate outcomes arose from panel discussion, of which three were voted for inclusion (online supplemental material S1). The 57 outcomes that met COS inclusion criteria were collapsed into 50 non-overlapping outcomes. For example, 'carbohydrate intake', 'fat intake', 'fiber intake', 'saturated fat intake', and 'total energy intake' were collapsed into 'macronutrients'. The 50 outcomes were then classified by the panel into 19 domains: diabetes (n=3 outcomes), other related diseases (n=3), complications in subsequent pregnancy (n=2), offspring outcomes (n=3), adiposity (n=4), cardiometabolic measures (n=5), glycemia (n=3), physical activity (n=2), diet (n=4), breast feeding (n=2), behavior change theory (n=5), diabetes-related knowledge (n=2), health literacy (n=1), social support (n=1), sleep (n=1), quality of life (n=1), program delivery (n=4), health economic evaluation (n=2), and diabetes risk screening (n=2) (figure 1).

The seven outcomes endorsed by  $\geq 90\%$  at round 2 were type 2 diabetes development and GDM recurrence, attending the postpartum diabetes screening visit and completing oral glucose tolerance testing and/or other glycemia measures, weight and total energy intake, and health behaviors in general (online supplemental file 2).

The 15 at the 80%–90% endorsement level were knowledge related to diabetes and prevention behavior; perception of diabetes risk; and motivation to change; as well as completion of targets and activities of the program itself, and health service use and cost; and longer-term complications of diabetes like cardiovascular disease and more specific elements of the top 7, such as specific glycemia measures (OGTT, fasting glucose, and glycated hemoglobin); body mass index (BMI), which is complementary to weight; specific measures of health behaviors, such as sedentary time and moderate-to-vigorous physical activity; and dietary aspects other than total energy, including carbohydrate intake and dietary quality.

The consensus panel distinguished outcomes that were relevant, irrespective of time since delivery (13 domains and 34 outcomes) and outcomes pertinent to one of three specific time periods: weeks and months after delivery (four domains and eight outcomes), subsequent pregnancy (one domain and two outcomes), and later years (two domains and six outcomes) (figure 1).

## DISCUSSION

The COS-DAP initiative engaged an international group of 115 stakeholders and 15 consensus panel members. The mean SD across outcome scores decreased between e-Delphi rounds, signaling overall convergence of

opinion. The application of a priori criteria to the e-Delphi scores led to 50 outcomes. Importantly, these reflected not only disease development and physiological effects but also patient-oriented outcomes like sleep quality and quality of life; health behaviors and behavioral theory constructs; knowledge, literacy, and risk perception; participation, adherence, and process outcomes; and economic measures. The relevance of some outcomes was a function of time since delivery or intervention focus. We believe that a comprehensive approach is crucial to facilitate the development and implementation of effective DAP programs.

Disease outcomes such as incident type 2 diabetes and cardiovascular disease were endorsed, as expected. Other disease outcomes included recurrent GDM and gestational hypertension in a subsequent pregnancy, as well as type 1 diabetes; indeed, a subset of women with GDM develop type 1 diabetes rather than type 2 diabetes in the years following pregnancy.<sup>15</sup> Diabetes in the offspring was also endorsed; both type 2 diabetes and type 1 diabetes in offspring have been associated with GDM in the mother in previous studies.<sup>16 17</sup>

Our stakeholders endorsed quality of life and sleep quality, which are patient-oriented outcomes that are also related to diabetes risk. Lack of sleep and psychosocial stress may increase risk of insulin resistance,<sup>18 19</sup> and maternal sleep disordered breathing is associated with GDM development.<sup>20</sup> Our COS is notable for focusing not only on biological effect measures but also on the behaviors that are expected to lead to these effects, namely, diet, physical activity, and breast feeding; these three domains captured outcomes such as sedentary time, macronutrients, and breastfeeding length. Complementing these were the psychological factors (eg, perceived diabetes risk and motivation to change) that need to shift for behavioral change to occur. A systematic review of weight loss interventions determined that such factors were more strongly associated with program attrition than demographic factors.<sup>21</sup> Some DAP prevention studies assess these factors at baseline,<sup>22</sup> but their improvement over the course of an intervention may signal potential for behavioral and therefore metabolic change. Similarly, our stakeholders emphasized the importance of health literacy, diabetes knowledge, diabetes risk perception, and social support. Again, these may be not only baseline contextual factors but also factors that evolve during an intervention.

Our stakeholders also underscored program participation and program delivery outcomes. Previous DAP prevention studies have exhibited variation in both participation rates and their reporting<sup>23 24</sup>; this metric is crucial to justify the implementation of sustainable programs. Frameworks exist to encourage reporting of program delivery process measures,<sup>25 26</sup> but COS-DAP represents a mandate to employ them. Finally, this COS recognizes the importance of collecting data to ascertain costs and cost effectiveness in order to achieve real-world implementation.

**Table 2** Delphi survey results for outcomes that met inclusion criteria: % of n participants who rated the outcome 7–9

Outcome	Round 1				Round 2			
	All	HCP	R	P	All	HCP	R	P
	(n=115)	(n=54)	(n=27)	(n=34)	(n=67)	(n=34)	(n=21)	(n=12)
Diabetes								
Diagnosis of type 1 diabetes	64.3	70	59.2	57.1	77.8	85.3	73.7	60
Diagnosis of type 2 diabetes	91.8	98	92.6	80.9	96.8	100	100	80
Complications in subsequent pregnancy								
Gestational diabetes mellitus recurrence	88.7	92	88.9	76.2	98.3	100	100	90
Other related disease								
Cardiovascular disease	64	72	66.6	56.5	82.5	91.2	89.6	60
Hypertension	61	64	66.6	52.2	71.4	73.5	79	50
Cardiometabolic risk								
Diastolic blood pressure	56	58	59.2	52.1	74.6	79.3	84.3	50
HDL cholesterol	56	56	55.5	47.8	69.9	67.7	78.9	50
LDL cholesterol	57	58	66.6	52.1	74.6	79.4	84.3	50
Systolic blood pressure	56	58	59.2	52.1	71.4	73.5	84.2	50
Glycemia measures								
Fasting blood glucose	75	73.6	82.1	66.6	89.2	91.2	90	63.7
HbA1c test	68.5	69.8	78.6	62.9	84.5	91.1	90	54.6
OGTT	82.4	90.5	82.2	70.3	89.3	97	90	81.9
Adiposity								
Body mass index	69.5	71.4	82.2	49.9	88	91.2	100	58.4
Postpartum weight retention	70.4	71.5	85.8	58.8	79.1	79.4	95.2	50
Weight	72.8	73.3	89.3	61.8	94	100	100	66.7
Waist circumference	53.1	53.7	64.3	54.6	71.7	73.5	81	50
Dietary								
Carbohydrate intake	69.1	63.9	65.4	81	84.5	87.1	84.2	100
Diet quality	71.3	68.1	73	90.5	87.9	87.2	84.3	100
Fat intake	64.9	63.9	57.7	76.2	79.3	80.6	84.2	87.5
Fiber intake	63.8	61.7	57.7	76.1	75.9	71	73.7	87.5
Food group servings	53.1	44.7	53.8	61.9	72.4	67.7	52.7	87.5
Glycemic load	62.7	61.7	53.9	71.4	72.4	67.8	68.5	87.5
Saturated fat intake	60.6	57.3	53.8	71.4	77.7	74.2	84.2	87.5
Total energy intake	77.6	78.7	73	90.5	93.1	93.6	94.8	100

Continued

Table 2 Continued

Outcome	Round 1				Round 2			
	All	HCP	R	P	All	HCP	R	P
	(n=115)	(n=54)	(n=27)	(n=34)	(n=67)	(n=34)	(n=21)	(n=12)
Physical activity								
Intensity of exercise	55.4	47.8	50	60	70.6	64.5	73.8	87.5
Moderate-to-vigorous activity	69.6	63.1	76.9	75	81.1	71	79	100
Sedentary time	77.1	76.1	77	80	84.5	83.9	94.7	100
Quality of life								
Quality of life	63.7	60	70.3	77.3	76.2	73.5	73.7	90
Sleep								
Sleep quality	60.9	45.7	69.2	85	70.6	58.1	84.2	87.5
Knowledge								
Diabetes knowledge	76.1	73.8	65.4	95	89.5	86.7	89.5	100
Diabetes prevention behavior knowledge	82.6	82.7	73.1	95	89.5	83.3	89.5	100
Behavior change theory								
Diet self-efficacy	66.2	67.4	65.4	85	75.4	73.3	84.2	100
Weight self-efficacy	64.1	65.2	53.9	75	72	63.4	78.9	87.5
Motivation to change	81.5	82.7	69.2	95	89.5	86.7	94.8	100
Perceived barriers	60.8	65.2	53.8	70	73.7	70	84.2	100
Perceived diabetes risk	76.1	74	65.4	95	86	83.3	84.3	100
Healthy lifestyle behaviors	84.8	87	73.1	95	92.9	90	94.8	100
Health literacy								
Health literacy	64.2	65.2	57.7	85	72	63.3	73.7	87.5
Social support								
Social support	64.2	65.2	57.7	75	79	76.6	84.3	100
Program delivery								
Completion of targets/activities	65.2	63.1	69.2	70	87.5	80	100.1	87.5
Intervention components engaged with	57.6	63.1	57.7	42.2	76.7	73.3	94.4	100
Whether goals were set for the woman and achieved*	0	0	0	0	73.1	70.1	88.9	87.5
Offspring outcomes								
Neonatal mortality/stillbirth	57.3	59.2	65.3	38.1	75.1	75.8	89.5	75
Health economic evaluation								
Health service use and cost	60.9	60.8	69.2	50	82.2	83.4	83.3	75

Continued

**Table 2** Continued

Outcome	Round 1				Round 2			
	All	HCP	R	P	All	HCP	R	P
	(n=115)	(n=54)	(n=27)	(n=34)	(n=67)	(n=34)	(n=21)	(n=12)
Risk screening interventions								
Attending postpartum diabetes screening test	85.9	84	92.6	68.2	95.2	94.1	100	90
Completion of OGTT/blood glucose/HbA1c postpartum	86.9	92	92.6	81.8	92.1	94.1	100	80

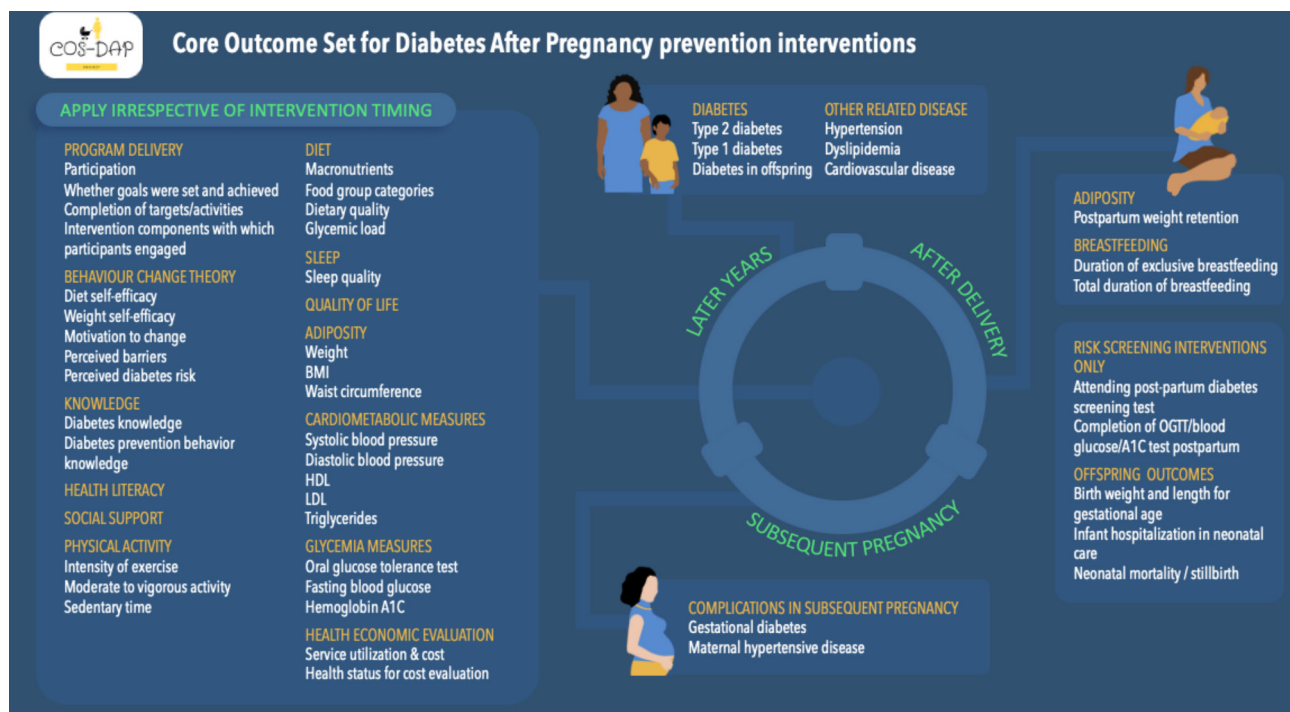
\*This was suggested as a candidate outcome following completion of the first survey round.

HbA1c, glycated hemoglobin; HCP, healthcare professionals; HDL, high-density lipoprotein; LDL, low-density lipoprotein; OGTT, oral glucose tolerance test; P, patients; R, researchers.

Overall, a large number of outcomes met survey inclusion criteria, with round 2 scores ranging from 69.9% to 98.3%. However, the diversity of outcomes is represented among even the top-scoring (ie, above 84.5%) outcomes (online supplemental material S2). For example, although moderate-to-vigorous activity did not receive a score above 84.5%, sedentary time and health lifestyle behaviors did. Similarly, fiber and saturated fat intake did not receive scores above 84.5%, but diet quality and carbohydrate intake scored highly. The ideal short-term study testing diabetes prevention interventions should thus consider impact on the breadth of outcomes in this COS, namely, disease outcomes, body weight/BMI, overall diet, physical activity, and completion of clinical tests. It should report on program-related process measures, such as the patient's completion of targets/activities. Finally, the study should also consider some impact on patient knowledge and motivation, which are measures that may indicate impact beyond the study.

One main objective of the Delphi process is to be inclusive of perspectives from a variety of stakeholders.<sup>12</sup> It was thus important that we included three key stakeholder groups. HCPs did not endorse sleep quality highly, but researchers and women with GDM did, allowing this outcome to cross inclusion thresholds. Participation of women with GDM in the Delphi process was important in retaining 'patient-oriented' outcomes such as social support, quality of life, and dietary outcomes—their high endorsement of these outcomes in round one prompted HCPs and researchers to change their scores to similarly high numbers in round 2, or their high scoring of these outcomes pushed an otherwise lower total score past the inclusion threshold. In contrast, weight and adiposity measures were retained because of high endorsement by HCPs and researchers, but were not endorsed by women with GDM. Women's low scores likely reflect the stigma of postpartum weight retention, which must be addressed so that inclusion of these important outcomes do not impact follow-up.

Our emphasis on inclusiveness of perspectives is also reflected by our application of a priori criteria following the e-Delphi survey. In contrast, as previously discussed, Bogdanet and colleagues developed a COS focusing on follow-up in women with medication-treated GDM, with an emphasis on brevity. During their consensus meeting, even outcomes with high endorsement during the e-Delphi process ( $\geq 70\%$  of participants scoring 7–9) were considered for exclusion. A consensus panel is, by design, smaller and less diverse than a group of e-Delphi survey respondents. We argue that removing outcomes highly endorsed by the survey respondents gives the smaller consensus panel a disproportionate voice in determining what is core. Therefore, we emphasized fidelity to the e-Delphi process and comprehensiveness of the outcome set. This is important not only for this particular COS but also for COS methodology in general.



**Figure 1** Core outcome set for diabetes after pregnancy prevention interventions. Fifty outcomes are grouped into 19 domains and arranged according to when they are measured relative to time since pregnancy during the intervention. Health literacy, social support, and quality of life are individual outcomes. HDL, high-density lipoprotein; LDL, low-density lipoprotein.

### Strengths and limitations

COS-DAP gathered perspectives in three different languages from 23 countries. Most came from high-income and upper middle-income countries, which enhances the COS' applicability in these settings. However, it may be less applicable in low-income or middle-income countries. Of the women with GDM, most came from Canada and had previous experience participating in diabetes prevention studies.<sup>27 28</sup> Although not representative of the broader target population, they had the background knowledge required to offer meaningful input.

Survey attrition was high, which aligns with reported higher attrition in surveys with greater than 50 items.<sup>29</sup> Our attrition was primarily among women with GDM. A COS project for epilepsy in pregnancy also reported difficulty engaging patients in the Delphi process.<sup>30</sup> As per ongoing discussions on COS methodology, patient input may be better captured with qualitative methods.<sup>11</sup> Their input was nonetheless valuable, influencing the responses of other stakeholders during round 2 of the e-Delphi survey.

We quantified convergence of opinion in the Delphi process by examining the SD for each outcome's scores. This approach is similar to Brookes and colleagues' in their study of different Delphi feedback methods.<sup>31</sup> COS development guidelines emphasize methodological integrity if a project intends to combine perspectives from a heterogeneous group of participants.<sup>11 12</sup> As previously mentioned, we were consistent with our published protocol and applied a priori survey criteria for outcome inclusion, only voting on indeterminate outcomes at the consensus meeting to avoid a COS biased toward the views of select individuals.

We acknowledge that the final COS captures a large number of outcomes, which may appear to increase participant burden. However, many are process measures and costs, captured through tracking during the intervention. Some are relevant to only specific time periods relative to pregnancy. Therefore, adopting the COS may not be as difficult as the total number of outcomes may suggest. Further, there are key potential benefits—they include identifying aspects of the strategy that may require optimization to enhance efficacy and determining whether implementation is feasible.

COS is subject to refinement as evidence evolves. Next steps may include engagement of women with GDM in a qualitative study to better discern their priorities, as well as a Consensus-based Standards for the Selection of Health Measurement Instruments initiative to identify how best to measure the outcomes in COS-DAP.

### CONCLUSION

A growing body of health behavior change trials aim to prevent DAP in women at known high risk of diabetes. The COS-DAP initiative executed a methodologically rigorous process to define a COS that captures not only glycemia and cardiometabolic risk factors but also behavioral, process, contextual, and economic measures. COS-DAP mandates that researchers consider a breadth of meaningful outcomes as core in this area of research. If research fails to consistently measure outcomes that are important to delivering impact on the health and well-being of this population, then the COS has failed to drive



the change it states is at the center of its mandate. We call on the scientific community to apply COS-DAP in future trials to build a meaningful and actionable evidence base, and ultimately to improve health outcomes.

#### Author affiliations

<sup>1</sup>Centre for Outcomes Research and Evaluation, Research Institute of the McGill University Health Centre, Montreal, Quebec, Canada

<sup>2</sup>UCD Centre for Perinatal Research and Institute of Food and Health, School of Agriculture and Food Science, University College Dublin, Dublin, Ireland

<sup>3</sup>Health Promotion Research, Steno Diabetes Center Copenhagen, Gentofte, Denmark

<sup>4</sup>Section for Health Promotion and Health Services Research, Aarhus University, Aarhus, Denmark

<sup>5</sup>Department of Medicine, Faculty of Medicine, McGill University, Montreal, Quebec, Canada

**Acknowledgements** The COS-DAP team members are grateful to the women with GDM, healthcare professionals, and researchers who took the Delphi survey. We thank both the librarian and Yvonne Leonard (MSc) from University College Dublin for their work on the systematic literature review. Our Canadian collaborators who attended the October 2018 clarification meeting included Anne-Sophie Brazeau (RD, PhD), assistant professor, McGill University; Ilana Halperin (MD, MSc), assistant professor, University of Toronto; Geetha Mukerji (MD, MSc), assistant professor, University of Toronto; Kara Nerenberg (MD, MSc), associate professor, University of Calgary; Ravi Retnakaran (MD, MSc), professor, University of Toronto; Baiju Shah (MD, PhD), associate professor, University of Toronto; Brandy Wicklow (MD, MSc), associate professor, University of Manitoba. Our international consensus panel members were Sonia Butalia (MD, MSc), assistant professor, University of Calgary, Canada; Anne Cathrine Staff (MD, PhD), professor, University of Oslo, Norway; Lorraine Lipscombe (MD, MSc), associate professor, University of Toronto, Canada; Ronald CW Ma (MD), professor, The Chinese University of Hong Kong, Hong Kong; Fionnuala McAuliffe (MD), professor, University College Dublin, Ireland; David McIntyre (MD), professor, University of Queensland, Australia; Sara J. Meltzer (MD), associate professor, McGill University, Canada; Per Ovesen (MD, PhD), professor, Aarhus University, Denmark; Liona CY Poon (MD), professor, The Chinese University of Hong Kong, Hong Kong; Eyal Sheiner (MD, PhD), professor, Ben-Gurion University of the Negev, Israel; and Umberto Simeoni (MD), professor, University of Lausanne, Switzerland. We also thank Deborah Chan (McGill University Health Centre), Nanna Husted Jensen (PhD, Aarhus University), and Sarah Louise Killeen (PhD, University College Dublin) for their assistance in organizing the meetings. We owe special thanks to Dr Moshe Hod (professor, Tel-Aviv University, Israel), who facilitated execution of the COS-DAP Delphi survey internationally and scheduling of the consensus meeting alongside the 10th International Symposium on Diabetes, Hypertension, Metabolic Syndrome, and Pregnancy (May 29–June 1, 2019, Florence, Italy). Cartoon icons of women and children in figure 1 were obtained through the graphic design platform Canva.com and were used in accordance with their respective license agreements. Elements are by iconology (Canva Free Media License Agreement) and DAPA Images (Canva One Design Use License Agreement).

**Contributors** SO'R, KKN, HTM, and KD conceived the original project. Project design was further developed with NW. NW and KD executed the project with assistance and input at various stages from SOR, KKN, and HTM. NW collected, cleaned, analyzed, and interpreted the data with oversight from KD. NW and KD drafted the manuscript. All authors critically reviewed the manuscript and approved the final version as submitted.

**Funding** It is supported by the Canadian Institutes of Health Research Planning and Dissemination Grant (Funding Reference #: PCS-155268).

**Disclaimer** The funder had no input into the conduct of the research or article preparation.

**Competing interests** None declared.

**Patient consent for publication** Not required.

**Ethics approval** The study design and protocol were approved by the research ethics board of the McGill University Health Centre on July 31, 2018 (reference number MP-37-2019-4765). Informed consent was obtained from all survey participants using a procedure that was approved by the ethics board.

**Provenance and peer review** Not commissioned; externally peer reviewed.

**Data availability statement** All data relevant to the study are included in the article or uploaded as supplemental information.

**Supplemental material** This content has been supplied by the author(s). It has not been vetted by BMJ Publishing Group Limited (BMJ) and may not have been peer-reviewed. Any opinions or recommendations discussed are solely those of the author(s) and are not endorsed by BMJ. BMJ disclaims all liability and responsibility arising from any reliance placed on the content. Where the content includes any translated material, BMJ does not warrant the accuracy and reliability of the translations (including but not limited to local regulations, clinical guidelines, terminology, drug names and drug dosages), and is not responsible for any error and/or omissions arising from translation and adaptation or otherwise.

**Open access** This is an open access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited, appropriate credit is given, any changes made indicated, and the use is non-commercial. See: <http://creativecommons.org/licenses/by-nc/4.0/>.

#### ORCID iD

Kaberi Dasgupta <http://orcid.org/0000-0002-2447-3553>

#### REFERENCES

- Dabelea D, Snell-Bergeon JK, Hartsfield CL, *et al*. Increasing prevalence of gestational diabetes mellitus (GDM) over time and by birth cohort: Kaiser Permanente of Colorado GDM screening program. *Diabetes Care* 2005;28:579–84.
- Nguyen CL, Pham NM, Binns CW, *et al*. Prevalence of gestational diabetes mellitus in eastern and southeastern Asia: a systematic review and meta-analysis. *J Diabetes Res* 2018;2018:1–10.
- Ratner RE, Christophi CA, Metzger BE, *et al*. Prevention of diabetes in women with a history of gestational diabetes: effects of metformin and lifestyle interventions. *J Clin Endocrinol Metab* 2008;93:4774–9.
- Kim C, Newton KM, Knopp RH. Gestational diabetes and the incidence of type 2 diabetes: a systematic review. *Diabetes Care* 2002;25:1862–8.
- COMET Initiative Database. Registration: core outcome set and core contextual factors in diabetes after pregnancy prevention trials n.d. Available: <http://www.comet-initiative.org/Studies/Details/1083>
- Nielsen KK, O'Reilly S, Wu N, *et al*. Development of a core outcome set for diabetes after pregnancy prevention interventions (COS-DAP): a study protocol. *Trials* 2018;19:708.
- O'Reilly SL, Leonard Y, Dasgupta K, *et al*. Diabetes after pregnancy prevention trials: systematic review for core outcome set development. *Matern Child Nutr* 2020;16:e12947.
- Bogdanet D, Reddin C, Macken E, *et al*. Follow-up at 1 year and beyond of women with gestational diabetes treated with insulin and/or oral glucose-lowering agents: a core outcome set using a Delphi survey. *Diabetologia* 2019;62:2007–16.
- Hsu C-C, Sandford Brian A. The Delphi technique: making sense of consensus. *Practical Assessment, Research & Evaluation* 2007;12.
- COMET DelphiManager. Available: <http://www.comet-initiative.org/delphimanager/> [Accessed 30 Oct 2019].
- Williamson PR, Altman DG, Bagley H, *et al*. The COMET Handbook: version 1.0. *Trials* 2017;18:1–50.
- Williamson PR, Altman DG, Blazeby JM, *et al*. Developing core outcome sets for clinical trials: issues to consider. *Trials* 2012;13:132.
- Formplus. Online & offline data collection tool. Available: <https://www.formpl.us/> [Accessed 30 Oct 2019].
- SurveyMonkey. The World's Most Popular Free Online Survey Tool. Available: <https://www.surveymonkey.com/> [Accessed 1 Nov 2019].
- Kim C. Maternal outcomes and follow-up after gestational diabetes mellitus. *Diabet Med* 2014;31:292–301.
- Wicklow BA, Sellers EAC, Sharma AK, *et al*. Association of gestational diabetes and type 2 diabetes exposure in utero with the development of type 2 diabetes in first nations and Non-First nations offspring. *JAMA Pediatr* 2018;172:724–31.
- Blotsky AL, Rahme E, Dahhou M, *et al*. Gestational diabetes associated with incident diabetes in childhood and youth: a retrospective cohort study. *CMAJ* 2019;191:E410–7.
- Spiegel K, Knutson K, Leproult R, *et al*. Sleep loss: a novel risk factor for insulin resistance and type 2 diabetes. *J Appl Physiol* 2005;99:2008–19.
- Räikkönen K, Keltikangas-Järvinen L, Adlercreutz H, *et al*. Psychosocial stress and the insulin resistance syndrome. *Metabolism* 1996;45:1533–8.

- 20 Pamidi S, Kimoff RJ. Maternal sleep-disordered breathing. *Chest* 2018;153:1052–66.
- 21 Moroshko I, Brennan L, O'Brien P. Predictors of dropout in weight loss interventions: a systematic review of the literature. *Obes Rev* 2011;12:912–34.
- 22 O'Reilly SL, Dunbar JA, Versace V, et al. Mothers after gestational diabetes in Australia (MAGDA): a randomised controlled trial of a postnatal diabetes prevention program. *PLoS Med* 2016;13:e1002092.
- 23 Dasgupta K, Terkildsen Maindal H, Kragelund Nielsen K, et al. Achieving penetration and participation in diabetes after pregnancy prevention interventions following gestational diabetes: a health promotion challenge. *Diabetes Res Clin Pract* 2018;145:200–13.
- 24 Lim S, Liang X, Hill B, et al. A systematic review and meta-analysis of intervention characteristics in postpartum weight management using the TIDieR framework: a summary of evidence to inform implementation. *Obes Rev* 2019;20:1045–56.
- 25 Hoffmann TC, Glasziou PP, Boutron I, et al. Better reporting of interventions: template for intervention description and replication (TIDieR) checklist and guide. *BMJ* 2014;348:g1687.
- 26 Pronk NP. Designing and evaluating health promotion programs. *Dis Manage Health Outcomes* 2003;11:149–57.
- 27 Brazeau A-S, Meltzer SJ, Pace R, et al. Health behaviour changes in partners of women with recent gestational diabetes: a phase IIA trial. *BMC Public Health* 2018;18:575.
- 28 Brazeau A-S, Leong A, Meltzer SJ, et al. Group-based activities with on-site childcare and online support improve glucose tolerance in women within 5 years of gestational diabetes pregnancy. *Cardiovasc Diabetol* 2014;13:104.
- 29 Gargon E, Crew R, Burnside G, et al. Higher number of items associated with significantly lower response rates in COS Delphi surveys. *J Clin Epidemiol* 2019;108:110–20.
- 30 Al Wattar BH, Tamilselvan K, Khan R, et al. Development of a core outcome set for epilepsy in pregnancy (E-CORE): a national multi-stakeholder modified Delphi consensus study. *BJOG* 2017;124:661–7.
- 31 Brookes ST, Macefield RC, Williamson PR, et al. Three nested randomized controlled trials of peer-only or multiple stakeholder group feedback within Delphi surveys during core outcome and information set development. *Trials* 2016;17:409.