



## Validation and endorsement of health system performance measures for opioid use disorder in British Columbia, Canada: A Delphi panel study ☆☆☆☆☆



Hasham Kamran<sup>a</sup>, Micah Piske<sup>b</sup>, Jeong Eun Min<sup>b</sup>, Lindsay A. Pearce<sup>c</sup>, Haoxuan Zhou<sup>b</sup>, Fahmida Homayra<sup>b</sup>, Linwei Wang<sup>c</sup>, Will Small<sup>a,d,e</sup>, Bohdan Nosyk<sup>a,b,\*</sup>

<sup>a</sup> Faculty of Health Sciences, Simon Fraser University, Burnaby, British Columbia, Canada

<sup>b</sup> Centre for Health Evaluation and Outcome Sciences, St. Paul's Hospital, Vancouver, British Columbia, Canada

<sup>c</sup> BC Centre for Excellence in HIV/AIDS, Vancouver, British Columbia, Canada

<sup>d</sup> British Columbia Centre on Substance Use, Vancouver, British Columbia, Canada

<sup>e</sup> Centre for Applied Research in Mental Health and Addiction, SFU Faculty of Health Sciences, Vancouver, British Columbia, Canada

### ARTICLE INFO

#### Keywords:

Opioid use disorder  
Performance measurement  
Health care quality improvement  
Delphi technique  
Qualitative analysis

### ABSTRACT

**Background:** Limited data exists on the performance of the healthcare system in opioid use disorder (OUD). We evaluated the face validity and potential risks of a set of health system performance measures for OUD collaboratively with clinicians, policymakers and people with lived experience of opioid use (PWLE) in the interest of establishing an endorsed set of measures for public reporting.

**Methods:** Through a two-stage Delphi-panel approach, a panel of clinical and policy experts validated and considered 102 previously constructed OUD performance measures for endorsement using information on measurement construction, sensitivity analyses, quality of evidence, predictive validity, and feedback from local PWLE. We collected quantitative and qualitative survey responses from 49 clinicians and policymakers, and 11 PWLE. We conducted inductive and deductive thematic analysis to present qualitative responses.

**Results:** A total of 37 measures of 102 were strongly endorsed (9/13 cascade of care, 2/27 clinical guideline compliance, 17/44 healthcare integration, and 9/18 healthcare utilization measures). Thematic analysis of responses revealed several themes regarding measurement validity, unintended consequences, and key contextual considerations. Overall, measures related to the cascade of care (excluding opioid agonist treatment dose tapering) received strong endorsements. PWLE highlighted barriers to accessing treatment, undignified aspects of treatment, and lack of a full continuum of care as their concerns.

**Conclusion:** We defined 37 endorsed health system performance measures for OUD and presented a range of perspectives on their validity and use. These measures provide critical considerations for health system improvement in the care of people with OUD.

### 1. Introduction

Canada continues to experience an ongoing opioid overdose crisis. Between January 2016 and December 2021, the country reported 29,052 opioid-related overdose deaths ([Government of Canada, 2022](#)).

In Canada's western-most province of British Columbia (BC) the rates of opioid-related overdose deaths remain the highest in the country despite offering perhaps the most robust range of services for opioid use disorder (OUD) nationwide. Additional efforts are therefore required to enhance these initiatives by monitoring and evaluating the support systems available for people with opioid use disorder (PWOD).

☆ **Funding Source:** Health Canada Substance Use and Addictions Program award no. 1819-HQ-000,036.

☆☆ **Role of the funding source:** The funding source was independent of the design of this study and did not have any role during its execution, analyses, interpretation of the data, writing, or decision to submit results. All authors had full access to the results in the study and take responsibility for the integrity of the data and accuracy of analysis.

\* **Disclaimer:** "All inferences, opinions, and conclusions drawn in this study are those of the authors, and do not reflect the opinions or policies of the Data Steward(s)."

\*\* **Declaration of competing interests:** BN is a Michael Smith Foundation for Health Research Scholar.

\* Corresponding author at: Faculty of Health Sciences, Simon Fraser University, BLU 9706, 8888 University Drive, Burnaby, BC, Canada V5A 1S6.

E-mail address: [bnosyk@sfu.ca](mailto:bnosyk@sfu.ca) (B. Nosyk).

<https://doi.org/10.1016/j.dadr.2022.100095>

Received 21 March 2022; Received in revised form 16 August 2022; Accepted 30 August 2022

2772-7246/© 2022 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY-NC-ND license

(<http://creativecommons.org/licenses/by-nc-nd/4.0/>)

Health system performance measurement offers clinicians, policy experts, and government officials a critical opportunity for accountability and provides actionable data to inform strategies and target resources to collaboratively improve healthcare delivery (Braithwaite et al., 2017; Fortney et al., 2014). Ongoing monitoring of care delivery and treatment engagement at the population level can help health officials determine whether quality of care is improving and whether further resources may be required to achieve desirable outcomes (Braithwaite et al., 2017; Fortney et al., 2014). Performance measures are therefore designed to establish a basis for data-driven public health action.

Performance measures to evaluate health services for OUD are limited around the world (Aarons et al., 2017; Scudder et al., 2017; Wu et al., 2016) and non-existent in BC. They are widely used for other chronic disease areas such as ischemic heart and chronic lung disease (Giovannetti et al., 2013) but have not yet been established in BC for OUD. The American Society of Addiction Medicine developed physician performance measures in OUD to evaluate standards of care for addiction specialist physicians capturing patient assessment and diagnosis, withdrawal management, treatment planning and management, and care coordination (ASAM Report on Performance Measures for Addiction Specialist Physicians, 2014). However, this document did not include measures of long-term treatment retention, clinical guideline compliance and receipt of harm reduction services or community-based care.

Leveraging the unique and comprehensive linked administrative databases available in BC, we previously developed 102 potential health system performance measures for OUD to provide provincial data on OUD care engagement, guideline compliance, and service provision verifying the effective predictive validity for each measure (Nosyk et al., 2022). Each performance measure reflects longitudinal population data from health administrative data sources and was informed by applicable guidelines for OUD treatment and concurrent disorders.

We executed a Delphi study for measure endorsement in collaboration with clinicians, policymakers and people with lived experience (PWLE) of OUD to support the measure development process. This article is intended to illustrate these views and provide greater context to the prior quantitative work that describes the initial measurement construction and quantification process (Nosyk et al., 2022). Our objective was to therefore collaboratively evaluate the face validity and potential risks of a set of health system performance measures for OUD in the interest of establishing endorsed measures for public reporting to inform efforts in response to the overdose public health emergency.

## 2. Methods

### 2.1. Initial performance measures development

The full performance measure development process was executed in two stages (i. Cohort and performance measures construction (including sensitivity analyses and predictive validity assessment) and ii. Two-stage Modified Delphi panel process for assessment of unintended consequences and face validity, and measure endorsement (Fig. 1). Detailed information regarding the first stage including cohort construction, performance measures construction, and analysis of predictive validity is described elsewhere (Nosyk et al., 2022). A full description of the datasets, measures, and sensitivity analyses is provided in **Supplementary Appendix 1**.

Performance measures were based on BC provincial health administrative data from January 1st, 1996 to November 30th, 2017, created according to collaborating stakeholders' stated needs and guided by theoretical (Donabedian, 1966) and applied (Braithwaite et al., 2018) studies in the peer-reviewed literature. A targeted review of the literature and consultation with care providers and local policymakers identified an initial set of 104 measures across four domains. We avoided health outcome measures and focused on mutable aspects of care delivery that may serve as targets for public health intervention. We organized mea-

asures according to four domains consisting of 21 groups: (1) the cascade of care for OUD (Groups 1–4; measures 1–13); (2) Opioid agonist treatment (OAT) clinical guideline compliance (Groups 5–13; measures 14–40); (3) healthcare integration: concurrent care and care pathways (Groups 14–19; measures 41–84); and (4) healthcare utilization (Groups 20–21; measures 85–104).

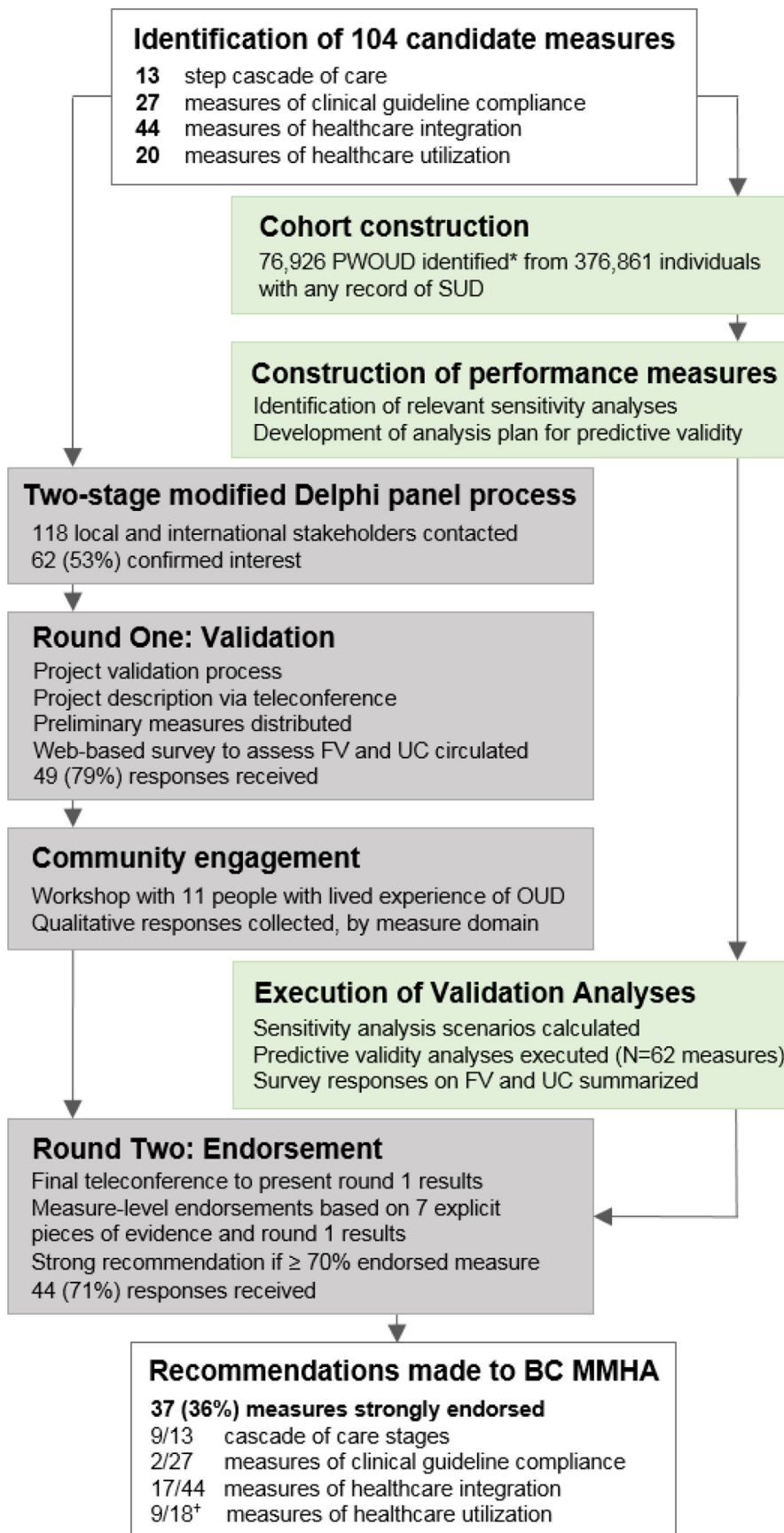
### 2.2. Delphi process and data collection

We conducted the second stage of the measure development process (assessment of unintended consequences and face validity, and measure endorsement) in collaboration with a panel of local and international experts in addiction medicine and healthcare policy, in addition to local community experts with lived experience. We requested engagement in each stage of the development process, with activities in the Delphi process distributed over three contact points during a six-month period (Fig. 1).

After excluding two measures (Measures 93 and 95 were sensitivity analyses for measures 92 and 94, respectively, after Delphi round one) from the initial 104 measures, a total of 102 preliminary measures were validated and rated for endorsement over a modified two-stage Delphi process (Hsu and Sandford, 2007; RAND Corporation, 2016). We collaborated with the BC provincial overdose emergency task force to recruit provincial stakeholders with expertise in OUD and healthcare policy. We also recruited relevant international experts through snowball sampling using the professional networks of the senior authors and study collaborators. In total, 118 local stakeholders and international experts in addiction medicine and healthcare policy were contacted, yielding 49 respondents for the initial validation survey and 44 respondents for the endorsement stage. Detailed information on the survey is provided in **Supplementary Appendix 2 Tables A1-A2, and Figs. A1-A2**.

#### 2.2.1. Delphi round one: assessment of face validity and unintended consequences

The first round began with an introductory teleconference describing our approach to validate the measures with detailed instructions and definitions for assessing the face validity and potential unintended consequences in reporting. A web-based survey link was circulated following the teleconference. To ensure adequate response rates, reduce respondent burden, and to ensure the measures were within the respondent's area of expertise, we divided the measures into 21 groups and created separate surveys for clinical experts (groups 1–15; 51 measures in total) and policy experts (groups 1–4, 16–21; 65 measures total). All respondents assessed the cascade of care measures (groups 1–4). For each measure group, panelists scored face validity ("Does this group of indicators reflect what it intends to measure?") and potential for unintended consequences ("Does this group of indicators have potential to result in unintended, potentially harmful consequences in BC?") on 7-point Likert scales (1=low face validity/high potential for unintended consequences and 7=high face validity/low potential for unintended consequences – where a high score is favorable). The survey provided participants with a list of 20 potential unintended consequences identified in a review by Mannion and Braithwaite (2012), who constructed the listing according to whether the measures have potential to result in unintended, potentially harmful consequences to the system of care for PWOD. Examples of unintended consequences may include tunnel vision (focusing on aspects of care that can be measured while displacing other unmeasured aspects of care), misrepresentation (deliberate manipulation of data by staff), gaming (altering of behavior or reporting in order to obtain strategic advantage) and reduced staff morale. Further examples of unintended consequences are listed in Table A2. We additionally provided figures presenting 5-year population trends in BC for each measure, with a scientific rationale; a Grading of Recommendations Assessment Development and Evaluation (GRADE) evidence score (Guyatt et al., 2008); information on construction of the numerator and denominator; and descriptions of proposed alternate calculations



**Fig. 1. Flow diagram of performance measure development and endorsement.** Abbreviations: SUD: substance use disorder; FV: face validity, UC: unintended consequences, OUD: opioid use disorder, BC MMHA: BC Ministry of Mental Health and Addictions. \* Period Prevalence (1996–1997); defined according to a case finding algorithm, described in supplementary appendix 2. † Performance measures 93 and 95 were incorporated as sensitivity analysis scenarios for measures 92 and 94, respectively. Detailed information on the process of cohort construction, performance measure construction, and analysis of predictive validity is described in Nosyk et al., 2022.

for sensitivity analyses. Additionally, to adequately capture respondent feedback, open-text fields were also included as optional responses to provide justifications on ratings.

### 2.2.2. Community engagement

We conducted a workshop with a small group of community members with lived experience of OUD, identified through our partnership with a local community-based organization, the Vancouver Area Network of Drug Users (VANDU). The workshop was held as a focus-group discussion to gather discussion regarding individual experiences and perspectives on healthcare for OUD to inform the development of the measures. The questions were structured according to the four domains outlined in Section 2.1. Respondents were not asked to quantitatively rate the face validity or potential for unintended consequences of each measure (as this required an intensive review of data), but to provide insight into what aspects of care should be measured, and other contextual factors to consider with these measures based on their personal lived experience with substance use and the healthcare system. All responses were recorded, de-identified, and transcribed.

### 2.2.3. Delphi round two: endorsement

In the second stage, via teleconference, we discussed responses and feedback received from PWLE, summarized validation results (including qualitative responses) received from round one and explained the endorsement methodology. In all, seven pieces of supporting evidence (the measures themselves, with the numerator, denominator and a 5-year plot, GRADE evidence quality ranking, sensitivity analyses, and quantitative results of face validity, unintended consequences, and predictive validity analyses) were summarized for each measure. Respondents then provided endorsement ratings for each assigned measure on 7-point Likert scales (“How strongly do you endorse this measure for reporting, given all of the information presented?”). A strong recommendation required at least 70% of respondents to endorse a measure (score of 5–7), 50–69% was considered a moderate recommendation, and measures receiving less than 50% endorsement were not recommended (Bradley et al., 2013; Grunfeld et al., 2008; Rewa et al., 2018; Salavati et al., 2017).

Similar to round one, we included optional open-text fields for respondents to provide justifications on endorsements and feedback. Qualitative responses obtained in both Delphi rounds were included for analysis.

## 2.3. Data analysis

### 2.3.1. Quantitative responses

First, we presented the breakdown and characteristics of respondents during both rounds of the Delphi process. Second, we presented respondents’ Likert ratings on face validity and unintended consequences, as well as measurement ranking (according to group mean – categorized as high if mean group score fell within top 1/3 of mean scores, mid if in the middle 1/3 of mean scores, and low if in the bottom 1/3 of mean scores), in addition to measurement endorsement ratings as previously defined. Finally, to investigate whether higher ratings of face validity and lower perceived risks of unintended consequences were associated with higher levels of endorsement, we quantitatively assessed the independent effects of validation results (% responses providing a strong endorsement – Likert scale responses  $\geq 5$ ) by conducting a multiple linear regression on each measures’ endorsement level.

### 2.3.2. Qualitative responses

We conducted qualitative analysis of the text responses provided by clinicians and policy experts in both Delphi rounds, as well as the responses of PWLE during the focus-group discussion. We coded and analyzed qualitative responses using NVivo 12.

Data included over 50,000 words of responses from the two stages and the community focus group discussion. Numerous reviews and re-reviews of qualitatively analyzing the data were part of the process,

leading to the developed approach. The iterative approach was defined by grouping comments by performance measure groups and thematically analyzing each comment. A mixture of inductive and deductive thematic analysis was performed. New themes were created for face validity and community engagement responses, while themes for unintended consequences were a mixture of emerging and *a priori* themes based upon examples of unintended consequences in the survey. We stratified the data on endorsements by individual measures. A qualitative expert cross-checked codes, themes, and representative quotes for accuracy and inter-rater reliability.

## 2.4. Ethical approval

This work was mandated by the Government of British Columbia as part of the response to the provincial opioid overdose public health emergency and classified as a quality improvement initiative. It was determined by the Providence Health Care Research Institute and the Simon Fraser University Office of Research Ethics that the analysis met criteria for exemption per Article 2.5 of the Tri-Council Policy Statement: Ethical Conduct for Research Involving Humans (Canadian Institutes of Health Research, 2010).

## 3. Results

### 3.1. Respondent characteristics

In the first stage, we received 49 responses from 26 policymakers and 23 clinicians (Table 1). In the second stage, we received 44 responses from 24 policymakers and 20 clinicians. We additionally collected qualitative responses from a total of 11 local PWLE.

### 3.2. Validation

#### 3.2.1. Quantitative ratings

Six measure groups (30 measures in total) had highly ranked validity with respect to face validity and unintended consequences (1. Opioid agonist treatment (OAT) engagement: Cascade of Care- Ever on OAT and 2. OAT retention; 3. Clinical guideline compliance: Medication switching and 4. Missed Doses; 5. Healthcare Integration: Concurrent Care-Chronic diseases; and 6. Healthcare utilization: Supply of health services for OUD) (Table 2).

A total of four performance measure groups had uniformly low rankings for face validity and unintended consequences (1. OAT engagement: Cascade of Care- Successful OAT taper; 2. Clinical guideline compliance: Baseline assessment; 3. Take-home dosing eligibility and 4. Dose tapering). Complete quantitative results from the validation (and endorsement) process are provided in **Supplementary Appendix 2 Tables A3-A6** (complete list of endorsed measures provided in Table A6).

#### 3.2.2. Qualitative responses

**3.2.2.1. Face validity.** Key themes on each measure group’s face validity that emerged from qualitative responses from clinical and policy experts are summarized in **Table 3**, with additional detail provided in **Supplementary Appendix 2 Table A7**. Representative quotations from these comments are highlighted below.

*Group 2 - Performance measure group is a meaningful indicator in describing the at-risk population. Face validity ranking = High.*

Group 2 measures included engagement in OAT stratified by being “ever in OAT” (PWOU with at least 1 OAT dispensation record), “recently in OAT” (PWOU with recent OAT discontinuation  $\leq 30$  days before the end of follow up) or “on OAT” (currently engaged in OAT at the end of follow up). The 49 respondents deemed this measure to have high face validity emphasizing the relevancy and importance of this measure.

- “This is an easy measurement and describes the exposure of at-risk patients to evidence-based treatment”

**Table 1**  
Respondent sampling scheme.

	Round One*		Round Two†	
	System	Clinical	System	Clinical
<i>Clinical and systems experts</i>				
Total responses received	26	23	24	20
Stakeholder location				
British Columbia	15	9	13	8
Rest of Canada	3	7	3	6
International	8	7	8	6
Stakeholder position				
Researcher	11	0	11	0
Health Administrator or Policy Advisor	15	0	13	0
Clinician	0	23	0	19
<i>People with lived experience</i>	11			

\*Face validity and unintended consequences.

†Measure endorsement.

- “Great idea to report on this. [It] should be a key indicator”

Group 7 - Parameters of guidelines are outdated and are no longer applicable. Face validity ranking = Mid

Only clinicians ( $n = 23$ ) ranked this group of measures, focusing on dose titration for buprenorphine/naloxone, methadone, and slow-release oral morphine. There was consensus regarding problems with the parameters of the measure, such as the maximum amount of titration and follow-up assessment, resulting in lower face validity ratings. The guidelines referenced were from 2017 and were referred to as “outdated”.

- “It is difficult to rate some as I have problems with the guidelines- in this case, the slow methadone increases. It is nearly a month before patients can get to an adequate dose”
- “The problem is this measures adherence to guidelines, but the guidelines are outdated and the measures you are capturing are inappropriate.”

Group 4 – Tapers are not considered best clinical practice. Face validity ranking = Low

Respondents expressed concern regarding the terminology used for describing OAT tapering. The measure group quantified completed OAT tapers in the last OAT episode and if there was record of relapse within the last one, six, or 12 months. Respondents highlighted key considerations around measuring completed tapers and personalized goals for treatment.

- “The term “successful taper” is inappropriate. I think this is just your definition of a taper, which may or may not be appropriate for the individual patient.”
- “I think ‘successful’ tapers could be monitored to understand system performance and health care utilization, but it should not be treated as a positive performance indicator. Also, I would not use the language ‘successful’ here. Consider ‘completed’ tapers instead, if keeping this measure.”

Subsequently, the label supporting these measures was changed to ‘completed’ tapers before the measure endorsement stage.

3.2.2.2. *Potential for unintended consequences.* Key themes on the potential for unintended consequences of each measure group that emerged from qualitative responses are summarized in [Table 3](#), with additional detail provided in [Supplementary Appendix 2 Table A8](#). Representative quotations are highlighted below.

Group 1 - Performance measure has the risk of tunnel vision by not capturing all PWOUD Unintended consequences rating = High

This group of measures explored the estimated prevalence of PWOUD in BC, and numbers of OUD diagnoses and residents receiving OUD care. Respondents highlighted the limitations of these estimates and how the numbers may not truly reflect the population’s size with OUD in the province.

- “There is a risk of tunnel vision by not capturing PWOUD that are accessing community or peer services who are not captured in the hospital or health services codes.”

Group 10 – Adjusting dosage may affect retention = High

This group of measures aimed to capture compliance with the missed dose guideline for buprenorphine/naloxone, methadone, and slow-release oral morphine. The loss of tolerance due to missed doses was deemed to be different for each of the three OAT medications, and the guideline for reassessing and reducing dosage elicited a negative reaction from the clinicians who responded.

- “These sets of measures are of less clinical importance than other measures geared at getting patients into OUD treatment with medication. I worry that a focus on these measures may lead systems/providers to focus less on more important measures”
- “Concerned about adjusting methadone doses like this as it may effect [sic] retention as people decide it is not worth the effort to come back for low doses and may be [sic] the recommendation needs to change to prevent unintended consequences.”

3.2.2.3. *Responses from people with lived experience (PWLE).* A summary of responses from community members with lived experiences of substance use is included in [Table 4](#), with full responses provided in [Supplementary Appendix 2 Table A9](#). Respondents highlighted their distinct treatment goals, as one respondent stated.

- “Some people are in there [in treatment] to get their kids back ... I think if you are not doing it for your own reasons, you’re not going to have success” while another respondent expanded by stating, “I think everybody’s goal is different. Yes to get off of [drugs], but maybe to get on another one [substitution medication or OAT]”.

Furthermore, the need to reduce wait times to receive treatment for participants and how this issue contributes to their lack of power in their own life was another focal point of the conversation. Other issues highlighted were around gender dynamics, such as lack of treatment centres for partners and the need for a full continuum of care, which involves advocacy from clinicians on behalf of clients for more housing options. Overall, respondents showed their gratitude for community support services such as VANDU and emphatically called for similar supports across the province.

### 3.3. Endorsement

#### 3.3.1. Endorsed measures

Overall, 37 measures received a strong endorsement, 33 received a moderate endorsement, and 32 were not endorsed ([Table 2](#)). Nine of 13 cascade measures, two of 27 guideline compliance measures, 17 of 44 integration measures (including 10 VCHA-specific measures, four

**Table 2**  
Validation results and endorsement ratings by measure group.

Performance measure group	No. Measures (PM No.)	Round one: Validation			Round two: Endorsement				
		No. Respondents	Ranking*		Total No. of Respondents	Ratings			
			Face Validity†	Unintended Consequences‡		Strong§N (%)	Moderate  N (%)	Not endorsed¶N (%)	
<i>OAT Engagement: Cascade of Care</i>									
1	OAT Engagement: Cascade of Care – Prevalence/Diagnosis	2 (1–2)	49	Mid	High	44	2 (100)	0 (0)	0 (0)
2	OAT Engagement: Cascade of Care – Ever on OAT	3 (3–5)	49	High	High	44	3 (100)	0 (0)	0 (0)
3	OAT Engagement: Cascade of Care – OAT Retention	4 (6–9)	49	High	High	44	4 (100)	0 (0)	0 (0)
4	OAT Engagement: Cascade of Care – Successful OAT Taper	4 (10–13)	49	Low	Low	43	0 (0)	0 (0)	4 (100)
<i>Clinical Guideline Compliance</i>									
5	Clinical Guideline Compliance: Baseline Assessment	3 (14–16)	23	Low	Low	20	0 (0)	2 (67)	1 (33)
6	Clinical Guideline Compliance: Initiation Dose	3 (17–19)	23	Mid	Mid	20	0 (0)	1 (33)	2 (67)
7	Clinical Guideline Compliance: Dose Titration	3 (20–22)	23	Mid	Low	20	0 (0)	2 (67)	1 (33)
8	Clinical Guideline Compliance: Maintenance Dose	3 (23–25)	23	Mid	Low	20	0 (0)	2 (67)	1 (33)
9	Clinical Guideline Compliance: Medication Switching	3 (26–28)	23	High	High	20	0 (0)	0 (0)	3 (100)
10	Clinical Guideline Compliance: Missed Doses	3 (29–31)	23	High	High	20	1 (33)	1 (33)	1 (33)
11	Clinical Guideline Compliance: Take-Home Dosing Eligibility	3 (32–34)	23	Low	Low	20	1 (33)	1 (33)	1 (33)
12	Clinical Guideline Compliance: Take-Home Dosing	3 (35–37)	23	Mid	High	20	0 (0)	2 (67)	1 (33)
13	Clinical Guideline Compliance: Dose Tapering	3 (38–40)	23	Low	Low	20	0 (0)	0 (0)	3 (100)
<i>Healthcare Integration – Concurrent Care &amp; Care Pathways</i>									
14	Integration: Concurrent Care – OAT	3 (41–43)	22	Low	Mid	20	0 (0)	0 (0)	3 (100)
15	Integration: Concurrent Care – Chronic Diseases	8 (44–51)	22	High	High	20	6 (75)	2 (25)	0 (0)
16	Integration: Care Pathways – Following ED Discharge	9 (52–60)	26	High	Mid	24	4 (44)	4 (44)	1 (11)
17	Integration: Care Pathways – Following Hospital Discharge	9 (61–69)	26	High	Mid	24	7 (78)	1 (11)	1 (11)
18	Integration: Care Pathways – Following Detox Discharge	9 (70–78)	26	Low	Mid	23	0 (0)	8 (89)	1 (11)
19	Integration: Care Pathways – Other community-based care	6 (79–84)	26	Mid	Mid	22	0 (0)	3 (50)	3 (50)
<i>Healthcare Utilization</i>									
20	Utilization: Service Utilization Among PWOU	11** (85–95)	26	Mid	Mid	24	1 (11)	3 (33)	5 (56)
21	Utilization: Supply of Health Services for OUD	9 (96–104)	26	High	High	24	8 (89)	1 (11)	0 (0)

**Abbreviations:** PM: performance measure; OAT: opioid agonist treatment, PWOU: people with opioid use disorder, OUD: opioid use disorder; ED: Emergency Department. \* Ranked according to group mean – categorized as high if mean group score falls in the top 1/3 of mean scores, mid if in the middle 1/3 of mean scores, low if in the bottom 1/3 of mean scores

† Face validity and potential for unintended consequences were rated on a 7-point Likert scale where 1=low face validity/high potential for unintended consequences and 7=high face validity/low potential for unintended consequences – high score is favorable

§ N: # of measures with ≥70% of experts giving a score of 5–7.

|| 50–69% of experts endorsing a performance measure (score of 5–7).

¶ <50% of experts endorsing a performance measure (score of 5–7); \*\* Measures 93 and 95 were set as sensitivity analyses for measures 92 and 94 after round one.

had no data available at initiation of the Delphi process), and nine of 18 healthcare utilization measures were strongly endorsed. Guideline compliance measures had the lowest level of endorsement, with 14 of 27 not endorsed. The list of strongly endorsed measures is presented in **Table 5**.

High rankings on face validity, unintended consequences, predictive validity, and the GRADE evidence quality scores were independently associated with higher levels of endorsement (**Table 6**).

All 44 respondents provided at least one comment for each measure they were asked to endorse (**Supplementary Appendix 2 Table A8**).

Most comments were within the clinical guideline compliance domain. Regular visits for missing doses were found impractical by some respondents, while suggestions were made for OAT initiation to begin while in hospital and to avoid wait times for treatment centres.

#### 4. Discussion

This study presents an endorsed set of health system performance measures for OUD and provides a comprehensive summary of expert responses regarding the factors influencing measure endorsement. The Delphi process resulted in 37 of 102 (36.3%) measures strongly endorsed

**Table 3**  
Key themes on face validity and unintended consequences collected from qualitative responses by measure group.

Performance Measure Group	No. Measures (PM No.)	Key Theme		
		Face validity	Unintended consequences	
<i>OAT Engagement: Cascade of Care</i>				
1	OAT Engagement: Cascade of Care – Prevalence/Diagnosis	2 (1–2)	It is difficult to know how many people are diagnosed with OUD	Performance Measure has risk of tunnel vision by not capturing all PWOUD
2	OAT Engagement: Cascade of Care – Ever on OAT	3 (3–5)	Meaningful indicator and key indicator describing at-risk population	Measure could encourage over-prescribing
3	OAT Engagement: Cascade of Care – OAT Retention	4 (6–9)	Discrepancies in measuring retention for OAT	There is potential for over-treatment
4	OAT Engagement: Cascade of Care – Successful OAT Taper	4 (10–13)	Tapers are not considered best practice in treatment	High likelihood of seeing unintended consequences such as measurement fixation
<i>Clinical Guideline Compliance</i>				
5	Clinical Guideline Compliance: Baseline Assessment	3 (14–16)	Baseline data collection needs improvement	Barriers for patients in accessing treatment
6	Clinical Guideline Compliance: Initiation Dose	3 (17–19)	Limitation in PharmaNet databases to track home inductions and those done in detox	Patients might avoid treatment if initial dosing is too low
7	Clinical Guideline Compliance: Dose Titration	3 (20–22)	Guidelines are outdated and no longer applicable	Guidelines are outdated and overly conservative
8	Clinical Guideline Compliance: Maintenance Dose	3 (23–25)	Frequent visits for maintenance dose are unnecessary	Guidelines are geared towards inexperienced clinicians
9	Clinical Guideline Compliance: Medication Switching	3 (26–28)	Guidelines differ from current clinical practice as they are outdated	Data to capture measure is complex and may be inaccurate
10	Clinical Guideline Compliance: Missed Doses	3 (29–31)	Concerns about SROM missed dose	Adjusting dosage may affect retention
11	Clinical Guideline Compliance: Take-Home Dosing Eligibility	3 (32–34)	Guidelines are too restrictive for patients	Low numbers could be used to implement tighter regulations for take-home doses
12	Clinical Guideline Compliance: Take-Home Dosing	3 (35–37)	Restrictions may lead to discontinuation of OAT	Frequent physician visits will lead to patient dropouts
13	Clinical Guideline Compliance: Dose Tapering	3 (38–40)	Tapering should be patient-centered and flexible	Risk of not capturing the data properly
<i>Healthcare Integration – Concurrent Care &amp; Care Pathways</i>				
14	Integration: Concurrent Care – OAT	3 (41–43)	Inconsistency in OAT dispensing	Patients may prefer different pharmacy for OAT and other medications
15	Integration: Concurrent Care – Chronic Diseases	8 (44–51)	Measure does not seem inclusive towards all chronic diseases	Chronic diseases should not be treated in the same facility as OUD
16	Integration: Care Pathways – Following ED Discharge	9 (52–60)	Patient trajectories post ED discharge vary	Measure can be misinterpreted and may not reflect the true story
17	Integration: Care Pathways – Following Hospital Discharge	9 (61–69)	Discharge Planning should occur prior to discharge	Improved discharge planning needed
18	Integration: Care Pathways – Following Detox Discharge	9 (70–78)	Data capture issues stem from regional differences and type of services	Measure could be subject to myopia as it is not inclusive
19	Integration: Care Pathways – Other community-based care	6 (79–84)	Measure may have difficulty capturing referrals to community-based care	Detox Referrals should not be incentivized
<i>Healthcare Utilization</i>				
20	Utilization: Service Utilization Among PWOUD	11 (85–95)	Proportion of Days Covered with OAT index is the most relevant measure	Data needs to be more specific and focus on metrics related to OAT treatment
21	Utilization: Supply of Health Services for OUD	9 (96–104)	Measure of Detox referral from supervised injection facility too narrow	Data needs to be more specific as not all services are equipped with the same resources

*Abbreviations:* OAT: opioid agonist treatment, PWOUD: people with opioid use disorder; ED = Emergency Department; SROM = slow-release oral morphine.

by clinicians and policy experts and were independently associated with high ratings of face validity and low risk for unintended consequences. Performance measure groups that received highly ranked face validity ratings included cascade of care measures for OAT engagement and retention, clinical guideline compliance for medication switching and missed doses, integrated care for concurrent conditions and after inpatient discharge, and supply of health services for OUD. Furthermore, PWLE highlighted individual aspects of treatment, care gaps, and the

specific resources needed to further support PWOUD. Although these measures were assessed for a provincial setting, the findings could inform policy and health system evaluation in other provinces and internationally.

Strongly endorsed measures reported on prevalence/diagnosis of OUD, ever on OAT, and OAT retention (the core elements of the cascade of OUD care) supporting findings from our prior work. Measures of OUD prevalence and diagnosed prevalence are required to determine

**Table 4**

Key themes and representative quotations on performance measures collected from focus group discussion with people with lived experience.

PM Domain	Key Themes	Representative Quotes
Treatment Engagement	<i>Waiting times are too long and contribute to a lack of autonomy.</i>	<ul style="list-style-type: none"> <li>• “Less time waiting. Like I mean they could give me my shot and say “get on your way””</li> </ul>
	<i>Treatment of PWUD and drug-related stigma</i>	<ul style="list-style-type: none"> <li>• “Treat them [patients] as adults, don’t treat them like children. They have to be led around... they have to be watched...”</li> <li>• “Your postal code for being down here [Vancouver Downtown Eastside]. Living down here you’re treated different than anywhere else. If you have an apartment, say in Arbutus or something, they’re going to treat you different than if you have one down here. As soon as they know your postal code that means you are drug seeking.”</li> </ul>
	<i>PWUD may have distinct treatment goals</i>	<ul style="list-style-type: none"> <li>• Some people are in there [in treatment] to get their kids back. You know, they put themselves in a treatment center. I think if you’re not doing it for your own reasons you’re not going to have a success.”</li> <li>• “I think everybody’s goal is different. Yes to get off of [drugs], but maybe to get on another one [substitution medication or OAT]”</li> </ul>
	<i>Barriers to treatment access</i>	<ul style="list-style-type: none"> <li>• “You know people want to go into a treatment center and they don’t have it for the couples. There’s a small window sometimes when somebody wants that... within hours, they want to get in there. If they don’t get a place they’re going to start using again. And if there’s no place for them that’s it, they’re fucked. The only time that they really came to the point that they were going to go right in and there’s no place for them.”</li> </ul>
Clinical Guideline Compliance	<i>Addressing other SDOH</i>	<ul style="list-style-type: none"> <li>• “Medical professions should advocate for housing for sure, decriminalizing drugs, they have to get on board. Not just lip service and doing whatever they’re told. Stand up and be part of the solution, not the problem.”</li> </ul>
	<i>OAT compliance is burdensome for individuals</i>	<ul style="list-style-type: none"> <li>• “Methadone takes over your life, right? Like you have to take it and you have to drink it [under supervision] every single day until you have this many clean urine tests. You can’t leave town, you can’t go camping because you have to go [to the clinic]. Some doctors will let you change pharmacies but most of them don’t. It takes over your life”</li> <li>• “Nope. There’s not enough [resources]. There’s not enough.”</li> </ul>
Resource Utilization	<i>Lack of resources Concerning</i>	<ul style="list-style-type: none"> <li>• “I find there’s more resources for women”</li> </ul>

*Abbreviations:* PM: performance measure; OAT: opioid agonist treatment, PWUD: people with opioid use disorder; PWUD: people who use drugs.

local health service needs. We previously estimated the prevalence of PWUD above the age of 12 in BC to be 1.92% (95% C.I: 1.89 – 1.95) as of 2017, corresponding to a population size of 83,760 (Min et al., 2020). Despite demonstrated life-saving benefits of OAT (Pearce et al., 2020), we previously reported only 16% of PWUD who had received OAT were retained in OAT for at least one year (Piske et al., 2020). Sustained monitoring of population OUD prevalence, OAT engagement and long-term retention should therefore be prioritized in the health system response for OUD and related comorbidities common among PWUD as part of an overall continuum of care approach.

This study presents endorsed measures for future applications in settings with similar linkable data sources. These measures were disseminated to provincial health authorities and have been used by these agencies to evaluate the quality of care for PWUD via an online interactive dashboard with measures stratified by region and patient demographics. As an example of their application, metrics such as retention in OAT have been of particular interest to health authorities comparing treatment outcomes in regions offering rapid access to OAT and integrated care. Similarly, in other settings, performance measures for OUD and other chronic disease areas are used by organizations such as the National Committee for Quality Assurance to measure quality of care practices and for health plan accreditation and include OUD measures that assess 6-month OAT retention rates as well as high-risk opioid analgesic prescribing practices (high dosage, multiple providers and

pharmacies, and length of prescription) by insurer and provider organization (NCQA, 2022). Additionally, other US organizations such as Shatterproof have developed quality measurement systems for addiction treatment facilities which assess care quality and patient experience at residential and outpatient addiction treatment programs to assist with linking people to services based on treatment needs, but do not explicitly measure provider performance (NQF, 2019).

The cascade of care stages focusing on completed tapers and tapering clinical guidelines measures were not endorsed by respondents due to tapering being perceived as an ineffective and clinically irrelevant treatment measure. In a population-based retrospective cohort study, fewer than 5% of methadone tapers were successfully sustained without relapse or treatment re-entry within 18 months after tapering (4.4% among all episodes initiating a taper and 2.5% among all completed episodes) (Nosyk et al., 2012). Respondents in our study called for tapering to be patient-centered and flexible and emphasized that tapering should not be a recommended goal of OAT treatment. There are also concerns abrupt OAT discontinuation may cause withdrawal-associated injury site-pain (BCCSU, 2017, Rieb et al., 2016) and is associated with a greater likelihood of subsequent termination of care (Perez et al., 2020). Tapering remains an active option for many clinicians and remains in the OAT guidelines as an avenue for PWUD who request it.

Overall, clinical guideline compliance measures were noted to have a high potential for unintended consequences, with qualitative responses



**Table 5**  
Strongly endorsed performance measures.

Performance measure	Total Responses	
	N	% Endorsed
<i>OAT Engagement: Cascade of Care</i>		
No.1 Prevalent population of PWOU	44	82%
No.2 Diagnosed PWOU	44	89%
No.3 Ever engaged in OAT	43	81%
No.4 Recently engaged in OAT	44	82%
No.5 On OAT <sup>†</sup>	44	82%
No.6 Retained in OAT > 1m	44	91%
No.7 Retained in OAT > 3m	44	93%
No.8 Retained in OAT > 12m	43	93%
No.9 Retained in OAT ≥ 24m	43	88%
<i>Clinical Guideline Compliance</i>		
No.29 BNX missed dose guideline compliance	20	85%
No.33 MET take-home dose eligibility guideline compliance	19	74%
<i>Healthcare Integration – Concurrent Care &amp; Care Pathways</i>		
No.45 Receipt of medical care for multiple OUD-related chronic disease	20	70%
No.46 Receipt of medical care for mental health conditions	20	85%
No.47 Receipt of medical care for HIV	19	89%
No.48 Receipt of medical care for Hepatitis C	20	80%
No.49 Receipt of medical care for alcohol use disorder	20	79%
No.50 Receipt of medical care for substance use disorder	20	70%
No.52 Presentation to ED as a first contact for an OUD-related condition	24	75%
No.56 Receipt of OAT within 3 days after ED discharge	24	96%
No.57 Receipt of physician follow-up within 3 days after ED discharge	24	79%
No.58 Receipt of community-based service referral within 30 days of ED discharge	24	75%
No.61 Leaving hospital against medical advice	24	71%
No.62 All-cause hospital readmission within 30 days	24	75%
No.63 ED admission within 30 days of hospital discharge	24	75%
No.65 Receipt of discharge planning within 7 days after hospital discharge	24	83%
No.66 Receipt of physician follow-up within 7 days after hospital discharge	24	75%
No.67 Receipt of OAT within 3 days after hospital discharge	24	96%
No.68 Receipt of community-based service referral within 30 days after hospital discharge	24	71%
<i>Healthcare Utilization</i>		
No.86 Percentage of OAT episodes with a PDC index >80% <sup>‡</sup>	22	77%
No.96 Person-years on OAT	23	83%
No.97 Number of OAT prescribers <sup>‡</sup>	23	91%
No.98 Number of active OAT prescribers <sup>‡</sup>	23	96%
No.99 Number of OAT dispensing pharmacies	24	83%
No.100 SIF visit rate per 1000 VCHA population	24	78%
No.102 Naloxone kit distribution rate <sup>‡</sup>	24	92%
No.103 Percentage of naloxone kits refilled	24	79%
No.104 OD prevention site visit rate per 1000 VCHA population <sup>‡</sup>	23	78%

**Abbreviations:** PWOU: people with opioid use disorder, OAT: opioid agonist treatment, OUD: opioid use disorder, BNX: buprenorphine/naloxone, MET: Methadone, ED: emergency department, SIF: supervised injection facility, OD: overdose. Strength of endorsement based on an a-priori rule where 70% of experts endorsing a performance measure (score of 5–7) was considered a strong recommendation, 50–69% was considered a moderate recommendation, and measures receiving <50% endorsement were not recommended. <sup>‡</sup> per 100,000 population; <sup>§</sup> LACE index predicts 30-day readmission or death in patients on medicine and surgery wards; <sup>†</sup> PDC index is the proportion of days in the measurement period “covered” by prescription claims for OAT.

indicating enforcement of a more rigid dosing regimen (particularly related to managing missed doses) could result in care that is not well-aligned with the patient’s needs. While monitoring individuals’ missed doses and medication switching may be useful for measuring care trajectories, the possibility of greater scrutiny of dosing practices may lead to physicians turning away patients or refocusing their clinical efforts entirely if oversight through performance measures restricts flexibility in clinical practice. Guideline compliance measures may therefore be better suited for internal use and ongoing training initiatives (BC Centre on Substance Use, 2019). Additionally, dosing practices for each of the medications in question would benefit from further inferential analyses to determine safe and effective ranges that reduce the risk of OAT discontinuation and death (Piske et al., 2020). Furthermore, in many cases, dosing guidelines have been based on low quality evidence. To this end, linked health administrative data present a basis to execute a

range of comparative effectiveness analyses to refine and strengthen the clinical evidence base for OUD.

Our consultation with PWLE provided critical context and depth towards the evaluation of these measures through the lens of lived experience in substance use and personal navigations of the healthcare system. Specifically, in terms of take-home dosing, clinicians spoke about potential drawbacks of the guidelines being too restrictive and the fact that these guidelines may lead to treatment discontinuation and drop-outs. This feeling was shared by some PWLE, who from the outside looking in, felt that their physicians are choosing not to give take-home carries, although stability in treatment is a prerequisite for take away doses. PWLE also cited the lack of stability and the constant stress of travelling around town to receive their treatment, inability to leave the town, or inability to travel due to the requirement to attend the pharmacy daily for supervised consumption of their medications. A required prerequisite for

**Table 6**  
Regression results on performance measure endorsement\*.

Covariate	Coefficient (95% CI)	P-value
Intercept	33.8 (22.6,45.1)	
Face validity ranking**:	19.4 (9.4,29.5)	0.0002
High		
Moderate	8.3 (-1.1,17.7)	
Low	-	
Unintended consequences ranking**:	11.5 (-0.3,23.2)	0.055
High		
Moderate	10.7 (-0.9,22.2)	
Low	-	
VCHA specific	2.7 (-7.6,13.0)	0.609
F-test p-value		<0.001
Adjusted R <sup>2</sup>		0.52

\* Generated from an ordinary least squares regression with  $N = 102$  observations (one per measure); dependent variable: % endorsed (Likert scale response  $\geq 5$ ).

\*\* Based on group scores (21 groups, 2–9 unique measures per group).

obtaining OAT – Urine Drug Screening – was heavily criticized by PWLE, calling it “undignified” and a “punishment for using other drugs.” There is no evidence on UDS’s effectiveness on patient or community health outcomes (McEachern et al., 2019), which supports the perspective of PWLE who question its utility. Finally, PWLE highlighted the need for peer navigators in hospitals to support and act as a voice for PWOUD in decision-making processes. Peer navigators’ support is associated with behavioral and psychological benefits, such as increased empowerment and self-esteem (Marshall et al., 2018). Notably, measures of care integration revealed exceptionally high rates of emergency department and hospital readmissions within 30 days of discharge, highlighting key areas in health care where peer engagement and collaboration should be integrated (Nosyk et al., 2022).

#### 4.1. Limitations

This study is reported with several limitations. We note a relatively small number of respondents, the panel’s potential selectivity, a relatively high respondent time burden, and desynchronized recruitment. Nonetheless, respondents included representatives from various organizations and health authorities across the province, including the BC Ministries of Health and Mental Health and Addictions. A streamlined process with broader engagement, particularly amongst local clinicians and a broader representation of PWLE should be prioritized in future applications. Furthermore, given that qualitative data was collected only through optional comment fields, there was a lack of depth in the comments for some performance measures. In addition, some respondents did not provide full context to their responses leaving them challenging to interpret and analyze. Further insight could be developed through in-depth qualitative interviews or focus groups, and this should be pursued through future research. Despite these limitations, this work presents the final results of a multi-stage process providing measures with practical considerations to comprehensively and appropriately assess health system performance in OUD.

#### 5. Conclusions

Our study provides a finalized, endorsed set of 37 health system performance measures for OUD in addition to expert perspectives on the benefits and potential pitfalls of these measures to monitor public health efforts across different domains of care to improve health outcomes for PWOUD.

#### Contributors

HK conducted the analysis, performed literature reviews, and wrote the first draft of the article. MP facilitated study activities and assisted with survey development and with writing the first draft of the article. JEM, ZH, FH and LW contributed to performance measure development and manuscript development. LAP facilitated study activities and assisted with survey and manuscript development. WS provided consultation on the analysis and contributed to manuscript development. BN conceptualized and secured funding for the study. All authors aided in the interpretation of results, provided critical revisions to the manuscript, and approved the final draft.

#### Role of funding source

The funding source was independent of the design of this study and did not have any role during its execution, analyses, interpretation of the data, writing, or decision to submit results. All authors had full access to the results in the study and take responsibility for the integrity of the data and accuracy of analysis.

#### Declaration of Competing Interest

BN received a Health Canada Substance Use and Addictions Program award (1819-HQ-000,036) to his institution. All other authors have no conflicts to declare.

#### Acknowledgments

This work was sponsored by the Health Canada Substance Use and Addictions Program award no. 1819-HQ-000036. BN is a Michael Smith Foundation for Health Research Scholar. We thank the Vancouver Area Network of Drug Users (VANDU) and the BC Association for People on Methadone (BCAPOM) for their support with the community engagement process and our community members with lived and living experience for their participation and valuable input. We acknowledge the support of colleagues at the BC Ministry of Health for the acquisition of provincial datasets (Liz Keay, Adele Green, Cindy Mei, Ruixia Mu, Jan Paolo Chua, Paul Lam, Tony Wang, David Godfrey, Katrina Barber, Meghan Will, and Kenny Der [PHSA]) and colleagues at Vancouver Coastal Health Authority for VCHA datasets (Rachael McKendry, Michael Li, Joleen Wright and Kent Hawkins). We thank staff at the BC Ministry of Mental Health and Addictions for their support for this project (Mikhail Torban, Darlene Seto, and Carolyn Davison) and also acknowledge the editorial services of Dr. Sarah Jefferies.

We acknowledge and extend our thanks to our expert panelists and survey respondents: Aaron Fox (MD, Albert Einstein College of Medicine); Alexis Crabtree (MD, BC center for Disease Control); Ambros Uchtenhagen (MD, University of Zurich); Andrew Kestler (MD, University of British Columbia); Annabel Mead (MD, Vancouver Coastal Health); Ben Ridout (BC Ministry of Health); Brooke Kinniburgh (MPH, Fraser Health Authority); Bruce Schackman (PhD, Weill Cornell Medical College); Carol Strike (PhD, University of Toronto); Carolyn Davison (BC Ministry of Health); Christian Schütz (MD, University of British Columbia, Provincial Health Services Authority); Christine Grella (PhD, University of California Los Angeles); Christopher Lukonis (MD, Gifford Medical Center); Darlene Seto (BC Ministry of Mental Health and Addictions); David Marsh (MD, Northern Ontario School of Medicine); Elaine Hyska (PhD, University of Alberta); Ellie Grossman (MD, Harvard Medical School); Eugenia Socias (MD, BC center on Substance Use, University of British Columbia); Frank Scheuermeyer (MD, University of British Columbia); Greg Whalen (MSW, Fraser Health Authority); Holly Clow (BC Ministry of Health); Jessica Moe (MD, University of British Columbia); John Brooklyn (MD, University of Vermont); Julie Bruneau (MD, University of Montreal); Karen Urbanoski (PhD, University of Victoria); Katie Hughes (MA, First Nations Health Authority);

Larissa Mooney (MD, University of California Los Angeles); Launette Rieb (MD, University of British Columbia); Lucas Wiessing (PhD, European Monitoring center for Drugs and Drug Addiction); Margaret Emslie (BC Ministry of Health); Mark Gilbert (MD, BC center for Disease Control); Michelle Wong (BC Ministry of Health); Mikhail Torban (BC Ministry of Mental Health and Addictions); Miranda Compton (BC Ministry of Health); Naushaba Degani (PhD, Health Quality Ontario); Patrick Day (BC Ministry of Health); Peter Selby (MD, The center for Addiction and Mental Health); Phillip Coffin (MD, University of San Francisco); Richard Rawson (PhD, University of Vermont); Samantha Arsenault (MA, Shatterproof); Selene Etches (MD, Dalhousie University); Sharon Stancliff (MD, New York State Department of Health); Sheena Taha (PhD, Canadian center on Substance Use and Addiction); Suzanne Brissette (MD, University of Montreal); Tami Mark (PhD, RTI International); Todd Sakakibara (MD, Vancouver Coastal Health); Tony George (MD, The center for Addiction and Mental Health).

## Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:[10.1016/j.dadr.2022.100095](https://doi.org/10.1016/j.dadr.2022.100095).

## References

- Aarons, G.A., Sklar, M., Mustanski, B., Benbow, N., Brown, C.H., 2017. Scaling-out" evidence-based interventions to new populations or new health care delivery systems. *Implement. Sci.* (1) 12. doi:[10.1186/s13012-017-0640-6](https://doi.org/10.1186/s13012-017-0640-6).
- ASAM Releases Report on Performance Measures for Addiction Specialist Physicians. (2014). <https://www.asam.org/Quality-Science/publications/magazine/read/article/2015/03/26/asam-releases-report-on-performance-measures-for-addiction-specialist-physicians>
- BC Centre on Substance Use. (2019). *Provincial opioid addiction treatment support program online course*. <https://www.bccsu.ca/provincial-opioid-addiction-treatment-support-program/>
- Bradley, N.M.E., Robinson, P.D., Greenberg, M.L., Barr, R.D., Klassen, A.F., Chan, Y.L., Greenberg, C.M., 2013. Measuring the quality of a childhood cancer care delivery system: assessing stakeholder agreement. *Value Health* 16 (4), 639–646. doi:[10.1016/j.jval.2013.02.016](https://doi.org/10.1016/j.jval.2013.02.016).
- Braithwaite, J., Hibbert, P., Blakely, B., Plumb, J., Hannaford, N., Long, J.C., Marks, D., 2017. Health system frameworks and performance indicators in eight countries: a comparative international analysis. *SAGE Open Med.* 5. doi:[10.1177/2050312116686516](https://doi.org/10.1177/2050312116686516), 2050312116686516.
- Braithwaite, J., Hibbert, P.D., Jaffe, A., White, L., Cowell, C.T., Harris, M.F., Runciman, W.B., Hallahan, A.R., Wheaton, G., Williams, H.M., Murphy, E., Molloy, C.J., Wiles, L.K., Ramanathan, S., Arnold, G., Ting, H.P., Hooper, T.D., Szabo, N., Wakefield, J.G., Muething, S., 2018. Quality of health care for children in Australia, 2012–2013. In: *JAMA - Journal of the American Medical Association*, 319. American Medical Association, pp. 1113–1124. doi:[10.1001/jama.2018.0162](https://doi.org/10.1001/jama.2018.0162).
- Corporation, T.R. (2016). *Delphi Method | RAND*. <https://www.rand.org/topics/delphi-method.html>
- Donabedian, A., 1966. Evaluating the quality of medical care. *The Milbank Memorial Fund quarterly*, 44 doi:[10.2307/3348969](https://doi.org/10.2307/3348969).
- Fortney, J.C., Pyne, J.M., Burgess, J.F., 2014. Population-level cost-effectiveness of implementing evidence-based practices into routine care. *Health Serv. Res.* 49 (6), 1832–1851. doi:[10.1111/1475-6773.12247](https://doi.org/10.1111/1475-6773.12247).
- Giovannetti, E.R., Dy, S., Leff, B., Weston, C., Adams, K., Valuck, T.B., Pittman, A.T., Blaum, C.S., McCann, B.A., Boyd, C.M., 2013. Performance measurement for people with multiple chronic conditions: conceptual model. *Am. J. Manag. Care* 19 (10), e359 [/pmc/articles/PMC4459786/?report=abstract](https://doi.org/10.1016/j.amj.2013.08.001).
- Grunfeld, E., Urquhart, R., Mykhalovskiy, E., Folkes, A., Johnston, G., Burge, F.I., Earle, C.C., Dent, S., 2008. Toward population-based indicators of quality end-of-life care: testing stakeholder agreement. *Cancer* 112 (10), 2301–2308. doi:[10.1002/cncr.23428](https://doi.org/10.1002/cncr.23428).
- Guyatt, G.H., Oxman, A.D., Vist, G.E., Kunz, R., Falck-Ytter, Y., Alonso-Coello, P., Schünemann, H.J., 2008. GRADE: an emerging consensus on rating quality of evidence and strength of recommendations. *BMJ* 336 (7650), 924–926. doi:[10.1136/bmj.39489.470347.ad](https://doi.org/10.1136/bmj.39489.470347.ad).
- Hsu, C.C., Sandford, B.A., 2007. The Delphi technique: making sense of consensus. *Pract. Assess. Res. Eval.* 12 (10), 1–8. doi:[10.7275/pdz9-th90](https://doi.org/10.7275/pdz9-th90).
- Mannion, R., Braithwaite, J., 2012. Unintended consequences of performance measurement in healthcare: 20 salutary lessons from the English National Health Service. *Intern. Med. J.* 42 (5), 569–574. doi:[10.1111/j.1445-5994.2012.02766.x](https://doi.org/10.1111/j.1445-5994.2012.02766.x).
- Marshall, C., Piat, M., Perreault, M., 2018. Exploring the psychological benefits and challenges experienced by peer-helpers participating in take-home naloxone programmes: a rapid review. In: *Drugs Educ. Prev. Policy* (Vol. 25, Issue 3), 280–291. doi:[10.1080/09687637.2016.1269724](https://doi.org/10.1080/09687637.2016.1269724).
- McEachern, J., Adaye-White, L., Priest, K.C., Moss, E., Gorfinkel, L., Wood, E., Cullen, W., Klimas, J., 2019. Lacking evidence for the association between frequent urine drug screening and health outcomes of persons on opioid agonist therapy. *Int. J. Drug Policy* 64, 30–33. doi:[10.1016/j.drugpo.2018.08.006](https://doi.org/10.1016/j.drugpo.2018.08.006).
- Min, J.E., Pearce, L.A., Homayra, F., Dale, L.M., Barocas, J.A., Irvine, M.A., Slaunwhite, A.K., McGowan, G., Torban, M., Nosyk, B., 2020. Estimates of opioid use disorder prevalence from a regression-based multi-sample stratified capture-recapture analysis. *Drug Alcohol Depend.* 217, 108337. doi:[10.1016/j.drugalcdep.2020.108337](https://doi.org/10.1016/j.drugalcdep.2020.108337).
- NCQA, 2022. National committee for quality assurance. | About Us. <https://www.ncqa.org/about-ncqa/>.
- Nosyk, B., Min, J.E., Pearce, L.A., Zhou, H., Homayra, F., Wang, L., Piske, M., McCarty, D., Gardner, G., O'Briain, W., Wood, E., Daly, P., Walsh, T., Henry, B., 2022. Development and validation of health system performance measures for opioid use disorder in British Columbia, Canada. *Drug Alcohol Depend.* 233, 109375. doi:[10.1016/j.drugalcdep.2022.109375](https://doi.org/10.1016/j.drugalcdep.2022.109375).
- Nosyk, B., Sun, H., Evans, E., Marsh, D.C., Anglin, M.D., Hser, Y.I., Anis, A.H., 2012. Defining dosing pattern characteristics of successful tapers following methadone maintenance treatment: results from a population-based retrospective cohort study. *Addiction* 107 (9), 1621–1629. doi:[10.1111/j.1360-0443.2012.03870.x](https://doi.org/10.1111/j.1360-0443.2012.03870.x).
- NQF, 2019. : shatterproof rating system for addiction treatment programs measure set. National Quality Forum. [https://www.qualityforum.org/Projects/s-z/Shatterproof/Shatterproof\\_Rating\\_System\\_for\\_Addiction\\_Treatment\\_Programs\\_Measure\\_Set.aspx](https://www.qualityforum.org/Projects/s-z/Shatterproof/Shatterproof_Rating_System_for_Addiction_Treatment_Programs_Measure_Set.aspx).
- Pearce, L.A., Homayra, F., Dale, L.M., Moallem, S., Barker, B., Norton, A., Hayashi, K., Nosyk, B., 2020. Non-disclosure of drug use in outpatient health care settings: findings from a prospective cohort study in Vancouver, Canada. *Int. J. Drug Policy* 84, 102873. doi:[10.1016/j.drugpo.2020.102873](https://doi.org/10.1016/j.drugpo.2020.102873).
- Perez, H.R., Buonora, M., Cunningham, C.O., Heo, M., Starrels, J.L., 2020. Opioid taper is associated with subsequent termination of care: a retrospective cohort study. *J. Gen. Intern. Med.* 35 (1), 36–42. doi:[10.1007/s11606-019-05227-9](https://doi.org/10.1007/s11606-019-05227-9).
- Piske, M., Thomson, T., Krebs, E., Hongdilokkul, N., Bruneau, J., Greenland, S., Gustafson, P., Karim, M.E., McCandless, L.C., MacLure, M., Platt, R.W., Siebert, U., Eugenia Socías, M., Tsui, J.I., Wood, E., Nosyk, B., 2020. Comparative effectiveness of buprenorphine-naloxone versus methadone for treatment of opioid use disorder: a population-based observational study protocol in British Columbia, Canada. *BMJ Open* 10 (9), 17. doi:[10.1136/bmjopen-2019-036102](https://doi.org/10.1136/bmjopen-2019-036102).
- Rieb, L.M., Norman, W.V., Martin, R.E., et al., 2016. Withdrawal-associated injury site pain (WISP): a descriptive case series of an opioid cessation phenomenon. *Pain* 157 (12), 2865–2874.
- Rewa, O.G., Eurich, D.T., Noel Gibney, R.T., Bagshaw, S.M., 2018. A modified Delphi process to identify, rank and prioritize quality indicators for continuous renal replacement therapy (CRRT) care in critically ill patients. *J. Crit. Care* 47, 145–152. doi:[10.1016/j.jcrc.2018.06.023](https://doi.org/10.1016/j.jcrc.2018.06.023).
- Salavati, M., Waninge, A., Rameckers, E.A.A., van der Steen, J., Krijnen, W.P., van der Schans, C.P., Steenberg, B., 2017. Development and face validity of a cerebral visual impairment motor questionnaire for children with cerebral palsy. *Child. Care Health Dev.* 43 (1), 37–47. doi:[10.1111/cch.12377](https://doi.org/10.1111/cch.12377).
- Scudder, A.T., Taber-Thomas, S.M., Schaffner, K., Pemberton, J.R., Hunter, L., Herschell, A.D., 2017. A mixed-methods study of system-level sustainability of evidence-based practices in 12 large-scale implementation initiatives. *Health Res. Policy Syst.* 15 (1), 102. doi:[10.1186/s12961-017-0230-8](https://doi.org/10.1186/s12961-017-0230-8).
- Wu, L.T., Zhu, H., Swartz, M.S., 2016. Treatment utilization among persons with opioid use disorder in the United States. *Drug Alcohol Depend.* 169, 117–127. doi:[10.1016/j.drugalcdep.2016.10.015](https://doi.org/10.1016/j.drugalcdep.2016.10.015).
- Government of Canada. Special Advisory Committee on the Epidemic of Opioid Overdoses. Opioid- and stimulant-related harms in Canada. Ottawa: public health agency of Canada; June 2022. <https://health-infobase.canada.ca/substance-related-harms/opioids-stimulants>