INNOVATION FOUNDATIONS

Patient-reported outcome measures

Methods for analysis and reporting

DECEMBER 2023

NSW GOVERNMENT
AGENCY FOR CLINICAL INNOVATION

The information in this document should not replace a clinician's professional judgement.

Agency for Clinical Innovation

1 Reserve Road St Leonards NSW 2065
Locked Bag 2030, St Leonards NSW 1590

Phone: +61 2 9464 4666 | Email: aci-info@health.nsw.gov.au | Web: aci.health.nsw.gov.au

Further copies of this publication can be obtained from the Agency for Clinical Innovation website at aci.health.nsw.gov.au

Title	Patient-reported outcome measures: Methods for analysis and reporting		
Published	December 2023		
Next review	2028		
Produced by	Evidence		
Preferred citation	NSW Agency for Clinical Innovation. Patient-reported outcome measures: Methods for analysis and reporting. Sydney: ACI; 2023		
TRIM ACI/D23/3530	SHPN (ACI) 230705	ISBN 978-1-76023-614-4	ACI_8703 [12/23]

Contents

Preface

Patient-reported outcome measures

Patient-reported measures (PRMs) give patients the opportunity to provide feedback on their health-related experiences and outcomes. There are patient-reported experience measures (PREMs) and patient-reported outcome measures (PROMs). The focus of this report is PROMs. PROMs provide a way to quantify:

- patients' health-related quality of life
- functional status
- symptoms and symptom burden
- treatment burden
- health behaviours
- health status.

The Patient Reported Measures Data Governance and Management Framework defines the primary and secondary use of data collected with PRMs as:

- primary use in real-time shared decision making between the patient and clinician
- secondary use by the clinician, health service or organisation in reflective practice, evaluation, policy and predictive modelling applications.

PROMs data can be used to reflect on outcomes at individual, service, and system levels as seen in Table 1.

Table 1: PROMs data use at the individual, service, and system level

Level	Use
Individual	• Understand and enhance interactions between patients and their care providers • Support shared decision-making regarding care, treatment and/or interventions
Service	• Compare outcomes across patients, cohorts, providers, teams, wards, or services • Assess use patterns, efficacy of interventions or treatments, and links between processes and outcomes • Evaluate and improve quality
System	• Inform policy • Reveal trends in outcomes • Identify factors associated with value-based healthcare • Inform quality improvement decisions at a state level • Assess adherence to clinical guidelines • Measure performance across healthcare organisations and services

Health Outcomes and Patient Experience Platform

NSW Health has developed a digital platform called Health Outcomes and Patient Experience (HOPE) to support statewide collection and use of PRMs. Implementation of the HOPE platform began in 2021.

Future health and value-based healthcare

The NSW Health Future Health Strategy has an objective for patients and carers to have positive experiences and outcomes that matter. This objective is supported through the NSW Health Value Based Healthcare program that strives to continually improve:

- health outcomes that matter to patients
- experiences of receiving care
- experiences of providing care
- the effectiveness and efficiency of care.

PRMs collection, use and analysis, facilitated through the HOPE platform, is a key component in achieving these objectives and delivering high value care.

Purpose of this report

In 2021, the Agency for Clinical Innovation (ACI) published a report on analytic principles for patient-reported outcome measures.[1] The report outlined key principles for the analysis of PROMs data and described a phased approach to PROMs data analysis in NSW.

This report builds on the 2021 report to:

- demonstrate methods that can be used to address some of the PROMs analysis issues raised
- provide guiding principles for analysing and reporting on PROMs data at the service and system level.

PROMs data collected in the HOPE platform for Leading Better Value Care (LBVC) tranche one cohorts and linked to other health data in the Register of Outcomes, Value and Experience (ROVE), have been used to develop and demonstrate these methods and principles. At the time of analysis, data was available for patients who completed their first PROMs survey in 2021 with an eight-month follow-up period.

In recognition of the importance of PROMs to inform and deliver value-based healthcare, HOPE is being progressively expanded to other patient cohorts. Collections for LBVC cohorts will continue as these programs become integrated into business-as-usual operations, to provide ongoing measurement to inform decision making. The methods and principles outlined in this report can be applied to all cohorts in HOPE and to PROMs data collected by other means.

Guiding principles for PROMs analysis and reporting

This report is intended for data analysts, statisticians, and researchers who, in partnership with clinicians, consumers and health service managers, want to analyse and report on PROMs data collected in HOPE at a service and system level for evaluation and quality improvement.

It provides a starting point for principles and methods that can be applied to PROMs data for robust analysis in the context of NSW Health PROMs data collection in HOPE. These methods will be further developed and refined over time by the ACI and by a growing community of NSW Health staff that is analysing and reporting on PROMs data for value-based healthcare.

Ten guiding principles for analysing and reporting on PROMs data have been drawn from:

- the literature
- best practice in health performance reporting
- analysis of the LBVC PROMs data collected in the HOPE platform.

They form the structure for this report. While these principles were developed based on the analysis of PROMs data for LBVC cohorts, they can be applied to PROMs data for other cohorts.

Principle 1: Clearly define the cohorts of interest

The definition may include a health condition, a diagnosis, healthcare service use and treatment, and a time period.[2] For example, it could be people enrolled in a chronic heart failure outpatient program between 1 January 2021 and 31 December 2021; specifying the follow-up period for PROMs data, for example, one year following program commencement.

Principle 2: Determine how the analysis is going to attribute patients to a service provider

Some patients may attend multiple services for a given condition. A decision may be required on whether to attribute patients to the service first visited, last visited, or most frequently visited. Seek clinical advice on the most appropriate attribution rule for the given cohort and conduct sensitivity analyses to assess the impact of different attribution rules.[2]

Principle 3: Conduct data integrity checks

PROMs data is complex and, given it is longitudinal and related to healthcare use, requires data linkage. Check and correct data linkage errors where possible, e.g. data duplication.[3]

Principle 4: Measure sample size and response rate

Calculate the number and proportion of patients who were allocated PROMs, both at baseline and at subsequent collection points. Check how surveys were allocated, for example, was it at random or were patients with more frequent healthcare attendance more likely to be allocated surveys. Consider any potential biases that the allocation of surveys may introduce.[4, 5] Calculate the number and proportion of patients who completed PROMs, both at baseline and at subsequent collection points. At least 30 respondents are recommended to report PRMs results for a service, for statistical reliability and patient confidentiality.[6, 7]

Principle 5: Assess the level of missing data

Missing data can refer to missing items on individual PROMs surveys or missing PROMs surveys at designated collection points. With the advent of digital PROMs collection systems, missing items on PROMs surveys are rare but missing surveys are still common. Imputation of missing data can be considered.[8, 9] Alternatively, reporting on fewer time points where there is less missing data may be preferred.

Principle 6: Evaluate respondent representation

Compare the characteristics of the respondents to the overall cohort, e.g. age, sex, comorbidities, culture and language, socioeconomic status. If the characteristics of the respondents are not representative of the cohort, consider weighting the data to reduce potential bias in the results. Weighting is recommended when there is at least a 5% difference between characteristics in the sample and population.[10] Weighting methods include raking, matching, and propensity weighting.[11-14]

Principle 7: Assess floor and ceiling effects

These effects occur in surveys when a substantial proportion of respondents (at least 15%) score the best or worst score possible, limiting the ability to discriminate between respondents and detect change in patient-reported health status over time.[15, 16] If these effects are present, use appropriate analysis methods to reduce the impact on results.[17-19]

Principle 8: Produce descriptive statistics on baseline PROMs

These results can be used to understand the health status of the cohort when starting a healthcare program.[20]

Principle 9: Apply robust statistical methods when measuring change in PROMs results over time

This includes determining minimally important difference, and adjusting for case-mix to make fair comparisons across service providers.[1, 6, 21-24]

Principle 10: Tailor analysis and reporting methods for individual, service, and system level reporting

Determine reporting cadence at each level. Present data to patients, clinicians, managers, and decision makers in an easily interpretable manner through use of clear data tables, informative visualisations, and provision of interactive reports or dashboards. Maintain a feedback loop between analysts and clinicians to continuously improve reporting. Use the information arising from the analyses to improve uptake and data quality of PROMs across cohorts, services, and districts.[25]

Summary

This report presents principles and methods for analysing and reporting on PROMs data at a service and system level, using PROMs data collected for LBVC tranche one cohorts.

Key findings

The key findings of the PROMs analysis presented in this report are:

- Many clinics collecting PROMs data had fewer than 30 completed surveys, the minimum number recommended for public reporting on aggregate PROMs results.
- The higher the PROMs sample size, the more likely the sample characteristics will closely match the cohort characteristics, potentially negating the need for the weighting of survey responses.
- While some patients attend multiple clinics for a given condition, it is usually within a single local health district, simplifying attribution decisions.
- Floor and ceiling effects are common on individual survey items but rare on domain or total survey scores.
- Reviewing baseline PROMs at the service and system level, in addition to the individual level, can provide meaningful information on case-mix.
- Given current PROMs completion rates, it is better to initially focus on only two surveys to assess change in health outcomes over time, to maximise sample size.
- It is important to consider minimally important difference when assessing change in patient-reported health status over time. However, some PROM surveys do not have established thresholds, and some have reported thresholds that are suitable for very specific cohorts only.

Next steps

In partnership with clinicians, consumers, managers, decision makers, analysts, and researchers, further work is required in:

- Developing methods to assess change in patient-reported health status over more than two time points, including consideration of patient trajectories and response shift.
- Case-mix adjustment of PROMs results to ensure fair comparisons across services and districts.
- Investigation of the association between PROMs participation and other health outcomes, such as hospital readmissions.
- Investigation of the association between interventions and PROMs results.
- Development of routine PROMs benchmarking reports.
- Incorporation of PROMs into the measurement of value across the quadruple aim of improving health outcomes that matter to patients, experiences of receiving care, experiences of providing care, and the effectiveness and efficiency of care.

Introduction

NSW Health's Future Health Strategy has a clear objective for patients and carers to have positive experiences and outcomes that matter. This objective is supported by the NSW Health Value Based Healthcare program that strives to continually improve health outcomes that matter to patients, experiences of receiving care, experiences of providing care, and the effectiveness and efficiency of care.

To help achieve these objectives and deliver high-value care, NSW Health has established a statewide system for the collection and use of PRMs. As outlined in the NSW Health Guide to Measuring Value[26], PRMs data is essential to measure and improve the value of care. There are PREMs and PROMs. The focus of this report is the use and analysis of PROMs in value-based healthcare.

Patient-reported outcome measures

PROMs capture the patient's perspectives about how illness or care impacts on their health and wellbeing. PROMs are used to support clinician decision-making and shared care planning. PROMs can also be used to evaluate and improve the quality of care.

Health Outcomes and Patient Experience

NSW Health has developed the HOPE platform to collect PRMs. HOPE enables the collection and use of PRMs data at the point of care via personal computers, tablet devices, or smartphones. The data collected is reported in real time for clinicians to support shared-decision making about care, treatment, and health interventions.

Implementation of the HOPE platform began in February 2021 and initially focused on LBVC and integrated care cohorts.

This scope has now broadened to include cohorts such as:

- stroke
- rehabilitation
- menopause
- cataracts
- transitional aged care
- paediatrics.

Further broadening of cohorts and populations is planned. Admitted and non-admitted patient (NAP) locations delivering care to patients in these cohorts are being progressively added to the HOPE platform.

Purpose and structure of this report

In 2021, ACI published a report on analytic principles for patient-reported outcome measures.[1] The report outlined key principles for the analysis of PROMs data and described a phased approach to PROMs data analysis in NSW. The report also highlighted challenges in analysing PROMs data such as attribution uncertainty, non-response bias, loss to follow up, floor and ceiling effects, response shifts, and confounders.

The purpose of this report is to demonstrate methods that can be used to address some of the PROMs analysis issues raised in the 2021 report. PROMs data collected in the HOPE platform for LBVC tranche one cohorts are used to develop and demonstrate these methods, however these methods are broadly applicable to current endorsed cohorts within HOPE.

This report includes information and methods on:

- cohort definition
- patient attribution
- completion rates
- missing data
- sample representativeness
- weighting
- floor and ceiling effects
- cross sectional analysis
- longitudinal analysis.

Also provided in this report are the overarching framework and guiding principles for the analysis and reporting of PROMs data at service and system levels, and the next steps for further development of robust PROMs analysis and reporting methods.

Data sources

For the purposes of this report, the data source is the Register of Outcomes, Value and Experience (ROVE), SAPHaRI, NSW Ministry of Health. ROVE includes linked emergency department, admitted patient, and NAP data for LBVC patients. It also includes all completed PRMs surveys from the HOPE platform for LBVC patients. Record linkage was carried out by the Centre for Health Record Linkage (CHeReL). At the time of data analysis for this report, healthcare activity data from July 2012 to June 2022 and PRMs survey data from February 2021 to August 2022 were available in ROVE.

All HOPE PRMs data, including for cohorts outside LBVC, is available for linkage via the CHeReL.

To maintain patient privacy and confidentiality, small numbers (<5) are suppressed in tables and figures throughout this report.

Methods for PROMs analysis and reporting

Clearly defining the cohorts of interest

PROMs analysis and reporting requires well-defined cohorts to ensure fair comparisons of PROMs results across service providers can be made (comparing like with like).

The definition may include a health condition, a diagnosis, healthcare service use and treatment. The time period for identifying the cohort, as well as the follow-up period, should also be specified.[2]

This report uses LBVC tranche one cohorts to develop and demonstrate PROMs analysis and reporting methods.

Cohorts

- Chronic heart failure (CHF)
- Chronic obstructive pulmonary disease (COPD)
- High risk foot services (HRFS)
- Inpatient management of diabetes mellitus (IMDM)
- Osteoarthritis chronic care program (OACCP)
- Osteoporotic refracture prevention (ORP)
- Renal supportive care (RSC)

The definitions for these cohorts were developed when LBVC was established and are detailed in the ROVE Data Dictionary.[27] For CHF, COPD, and IMDM, the definitions are a hospital inpatient episode with a diagnosis for the condition of interest recorded. For HRFS, OACCP, and ORP, the definitions are a hospital inpatient episode with a diagnosis for the condition of interest recorded and/or attendance at a designated NAP clinic. For RSC, the definition is attendance at a designated NAP clinic.

The HOPE rollout started in February 2021 and implementation was prioritised by local health districts/specialty health networks (LHDs/SHNs), based on readiness assessments. As implementation began during the pandemic, a large proportion of it focused on NAP locations before admitted patient locations. To target patients where HOPE is available, the cohort definitions can be revised to LBVC patients (ROVE Data Dictionary)[27] attending LBVC clinics (Appendix 1). This definition will create more homogenous cohorts of patients receiving NAP care for their condition, enabling fairer comparisons across service providers.

Limitations of the definition

There are some limitations with this definition. For HRFS, OACCP, ORP, and RSC, there are specific LBVC NAP clinics with dedicated establishment types (Appendix 1). For CHF, COPD, and IMDM, specific clinics do not exist. Relevant clinics have been identified based on the establishment types of the clinics registered in HOPE for these cohorts. These establishment types were validated by analysing the clinics most frequently visited by CHF, COPD, and IMDM patients following discharge from an admitted patient episode (Appendix 1).

Also, for CHF, COPD, and IMDM, where the ROVE cohort definitions only include people who have been hospitalised with the condition of interest, we will not capture people referred to LBVC NAP clinics from primary care and may skew the analysis towards patients with more severe disease. However, it is not possible to identify all patients referred from primary care in the data currently available. If the bias is similar across districts, fair comparisons can still be made.

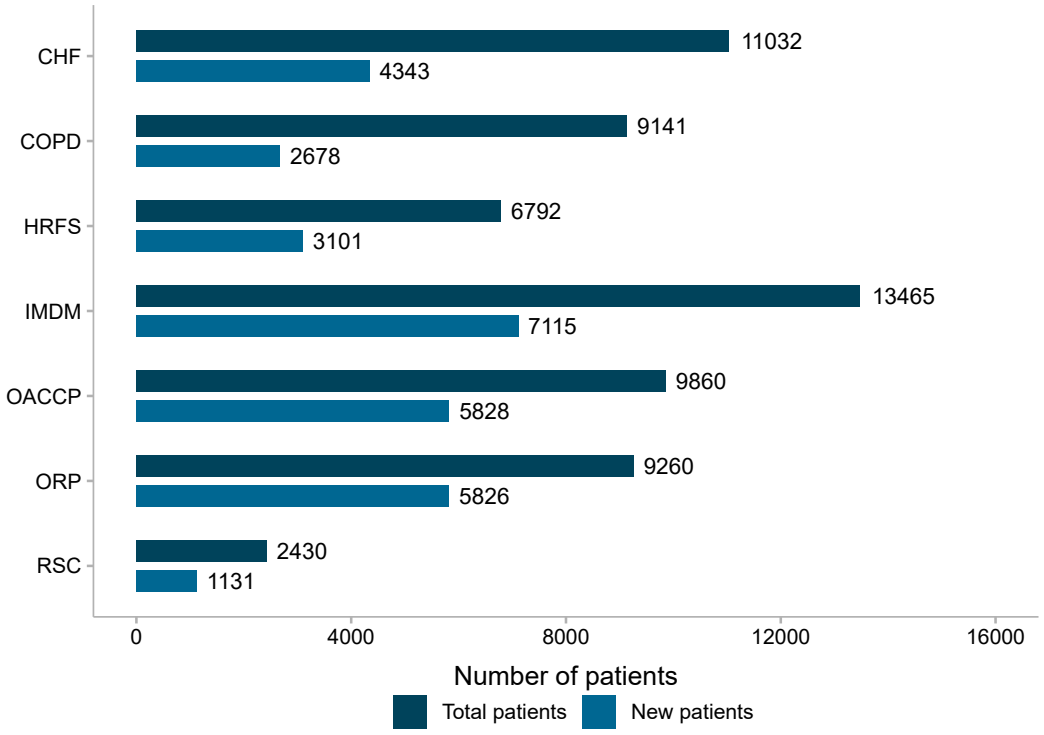
Time span of data analysis

At the time of data analysis for this report, PROMs data in ROVE was available from February 2021 to August 2022. Given the data available, we have identified LBVC patients in 2021 and followed them for eight months from the completion of their first survey for the analysis in this report.

Applying these cohort definitions at the NSW level, there are thousands of LBVC patients attending LBVC NAP clinics (Figure 1). Across LHDs, there is substantial variation in the number of LBVC patients attending LBVC NAP clinics (data not shown). This variation must be accounted for when making PROMs comparisons across LHDs.

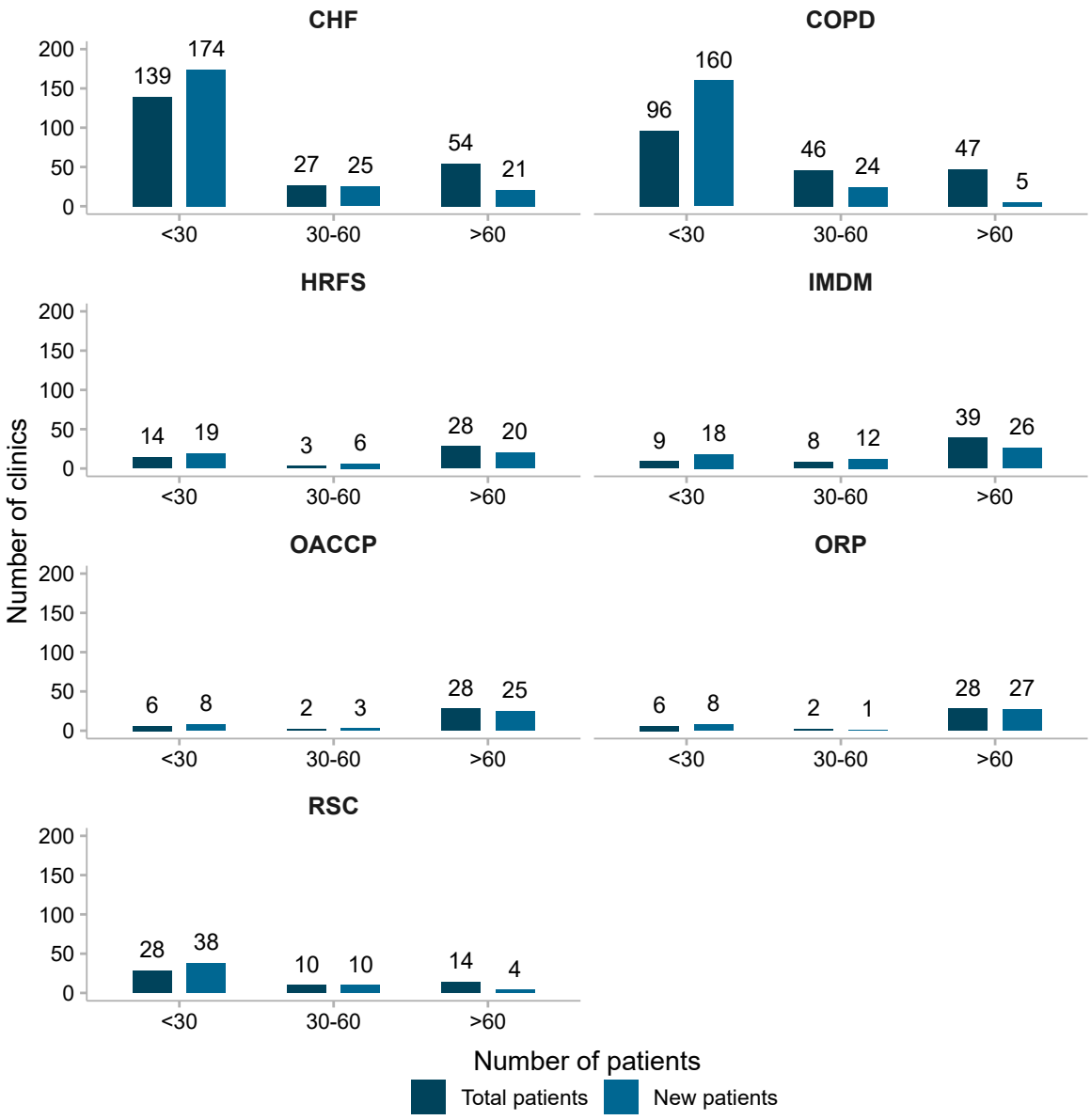
There is also wide variation in the number of patients seen in individual clinics and some clinics see very few patients (Figure 2). It may not be possible to benchmark PROMs for some clinics due to small numbers. At least 30 respondents are recommended to report PRMs results for a service, for statistical reliability and patient confidentiality.[6, 7] Out of the estimated 634 LBVC NAP clinics, 298 (47%) saw less than 30 total LBVC patients in 2021, and 425 (67%) saw less than 30 new LBVC patients in 2021.

Figure 1: Number of LBVC patients attending LBVC NAP clinics, 2021



1. Total patients are all patients who attended an LBVC NAP clinic in 2021.
2. A new patient is a patient who attended an LBVC NAP clinic in 2021 for the first time in at least two years.

Figure 2: Number of LBVC patients in LBVC NAP clinics, 2021



Determining how the analysis is going to attribute patients to a service provider

To report and compare PROMs results across service providers, rules need to be established on patient attribution.

Some patients may attend multiple services for the condition of interest. In consultation with clinicians and the consumer, a decision may need to be made on which service the patient should be attributed. For example, is it the service first visited, last visited or most frequently visited that provided the most significant intervention?

Analysing the number and proportion of patients that attend multiple services, and the various patterns of attendance, may help guide attribution decisions. Sensitivity analyses can be conducted to assess the impact of attribution decisions.[2]

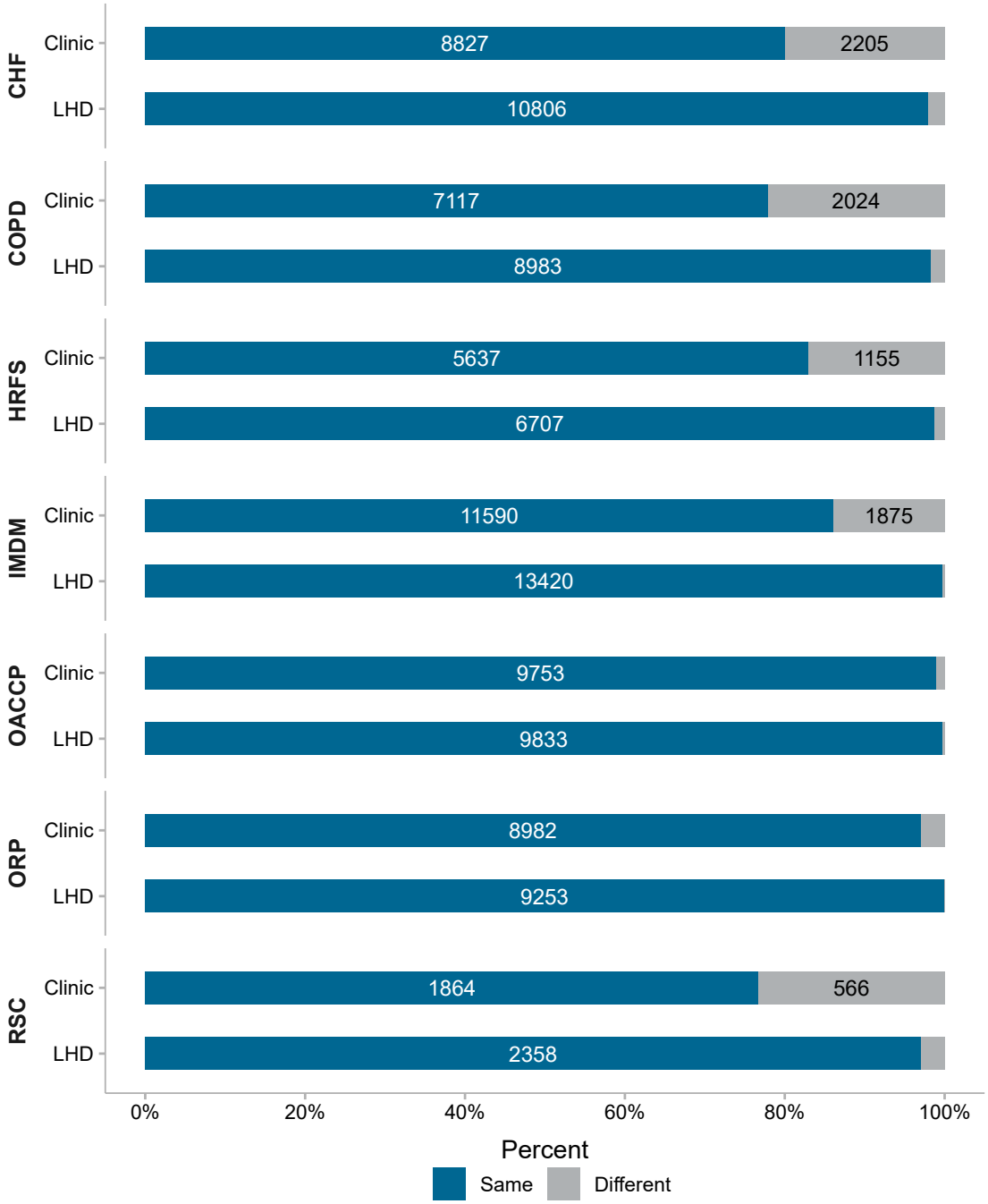
For some LBVC initiatives (CHF, COPD, HRFS, IMDM, RSC), a substantial proportion of patients (10-20%) attend multiple LBVC NAP clinics, but usually within the same LHD (Figure 3). This makes attribution simple at an LHD level but more complicated at a clinic level. While some patients attend multiple clinics, most (99%) completed their PROMs surveys within the one clinic (Table 2).

For the LBVC PROMs analysis presented in this report, we have attributed patients and their PROMs surveys to the first LBVC NAP clinic attended in 2021. This attribution rule will be reviewed with relevant stakeholders before routine reporting.

Table 2: Number and percentage of LBVC patients completing PROMs surveys in different clinics and different LHDs, 2021 for first survey and eight months follow up for subsequent surveys

LBVC cohort	Patients	Different clinic	%	Different LHD	%
CHF	232	<5	<2%	<5	<2%
COPD	211	20	9%	<5	<2%
HRFS	685	5	1%	<5	<1%
IMDM	102	0	0%	0	0%
OACCP	2974	19	1%	<5	<1%
ORP	873	10	1%	0	0%
RSC	506	12	2%	7	1%
Total	5583	68	1%	12	0%

Figure 3: Number and percentage of LBVC patients attending different clinics and different LHDs, 2021



Conducting data integrity checks

PROMs data is complex and can require extensive processing before it is ready for analysis.

Individual item responses may need to be aggregated into total scores. Multiple surveys completed by an individual need to be linked. Healthcare activity data may also need to be linked to compare survey completion with service use. There may be missing data and data linkage errors.[3]

The HOPE PRMs datasets are relatively new and not yet widely used. Unlike well-established administrative datasets, standard operating and analysis procedures are not well known and stringent data pre-processing and integrity checks are required.

Data linkage

Linkage was required between a dataset containing administrative information on completed surveys (location, datetime), a dataset containing descriptive information on the survey and survey items, and a dataset containing the survey responses. This data was also linked to NAP data to compare healthcare activity and PROMs completion.

Deduplication

There were duplicates in the HOPE PRMs administrative dataset in ROVE. There were 724 exact duplicate records out of 31,421 records (2%), that were removed from analysis. There were also 29 survey records that mapped to multiple people (0.1%), a linkage error within the standard error bounds for probabilistic linkage (0.5%).[3] This error was rectified by identifying the person with a matching NAP visit. There were no duplicates in the survey response dataset.

Data integrity

The collection of HOPE PRMs data is complex – people can attend several clinics and complete PRMs surveys for multiple conditions. We conducted several data checks to better understand PRMs collection patterns and to ensure the integrity of the linked PRMs and NAP data in ROVE.

PRM survey type and clinic type concordance

NAP clinics are assigned to one LBVC program based on their establishment type. In theory, patients can be assigned and complete PRMs surveys for multiple LBVC programs within the one NAP clinic. In 2021, most completed surveys matched the clinic type in ROVE (Table 3).

PRM survey completion and clinic visit concordance

Most people who have completed a PRM survey in an LBVC NAP clinic in HOPE data will have a visit recorded in the same LBVC NAP clinic in NAP data. Some people can be assigned and complete a PRM survey without a formal service event in a NAP clinic. In 2021, more than 80% of people who completed a PRM survey in an LBVC NAP clinic had a visit recorded in the same LBVC NAP clinic across LBVC programs (Table 4).

Table 3: Type of PRM survey completed by LBVC NAP clinic, 2021

	Number of PRMs surveys completed							
LBVC NAP clinic	CHF	COPD	HRFS	IMDM	OACCP	ORP	RSC	Other
CHF	568	0	0	0	0	0	0	19
COPD	0	730	0	0	0	0	0	10
HRFS	0	0	1135	0	0	0	0	0
IMDM	0	0	0	240	0	0	0	0
OACCP	0	0	0	0	7604	0	0	<5
ORP	0	0	0	0	0	1513	0	0
RSC	0	0	0	0	0	0	1170	0

Table 4: Number of people with PRM survey completion and visit in the same LBVC NAP clinic, 2021

	LBVC NAP clinic visit			
LBVC PRM survey	Yes	No	Total	Percent
CHF	304	61	365	83%
COPD	306	66	372	82%
HRFS	641	63	704	91%
IMDM	108	20	128	84%
OACCP	2808	224	3032	93%
ORP	835	38	873	96%
RSC	443	37	480	92%

Note: Restricted to cases where PRM survey type matches clinic type.

Among the people with PRM survey completion and visit in the same LBVC NAP clinic, not all of them were flagged as an LBVC patient in NAP data in ROVE for CHF, COPD, and IMDM (Table 5). This means that there are some people completing CHF, COPD, and IMDM PRM surveys who do not technically meet the cohort definitions for CHF, COPD, and IMDM.

Table 5: Number of people with PRM survey completion and visit in the same LBVC NAP clinic who were flagged as an LBVC patient in NAP data, 2021

	Flagged as LBVC patient			
LBVC program	**Yes**	**No**	**Total**	**Percent**
CHF	190	114	304	63%
COPD	192	114	306	63%
HRFS	641	0	641	100%
IMDM	87	21	108	81%
OACCP	2808	0	2808	100%
ORP	835	0	835	100%
RSC	443	0	443	100%

Measuring sample size and response rate

PROMs allocation and completion rates should be measured to assess uptake and respondent representation.

Allocation methods should be reviewed for any potential bias, e.g. patients with more severe disease less likely to be allocated PROMs.[4, 5]

If allocation and completion rates are low for baseline and follow up surveys, further work may be required to improve PROMs uptake before robust PROMs analysis at the service and system level can be performed. A comparison of completion rates across service providers and health conditions and an investigation of factors associated with completion may provide insights into where and how uptake can be improved.[28]

The completion rate can also provide an indication as to whether the respondents are likely to be representative of the cohort. Respondent representation can be further explored through an analysis of responder and cohort characteristics (see Evaluating respondent representation section). Completion rates will also affect the ability to detect statistically significant differences.[6]

Calculating PROMs completion rates

Implementation of the HOPE platform began in February 2021 and LBVC NAP clinics are progressively going live in the platform and collecting PROMs data from LBVC patients. For this report, HOPE PROMs data from February 2021 to August 2022 was available. PROMs completion rates are calculated for patients who attended an LBVC NAP clinic after the clinic went live in HOPE in 2021 ('eligible patients'). These patients were followed for eight months after the completion of their first survey in 2021.

For PROMIS-29 (see Appendix 2 for survey details), the percentage of eligible LBVC patients who completed at least one PROMIS-29 survey ranged from 4% to 37% across LBVC cohorts. The percentage who completed a follow-up survey within eight months ranged from <1% to 16% (Figure 4). The percentage of patients who died within eight months of completing their first survey ranged from <1% to 7.2% (CHF: 7.2%, COPD: 3.7%, HRFS: 5.4%, IMDM: 2.7%, OACCP: <1%, ORP: <1%).

For condition specific PROMs surveys for OACCP and RSC (see Appendix 2 for survey details), the percentage of eligible LBVC patients who completed at least one survey ranged from 1% to 35%. The percentage who completed a follow up survey within eight months ranged from <1% to 17% (Figure 5). For OACCP, less than 1% of patients died within eight months of completing their first Hip dysfunction and Osteoarthritis Outcome Score (HOOS), Knee injury and Osteoarthritis Outcome Score (KOOS), Oxford Hip Score (OHS) or Oxford Knee Score (OKS) survey. For RSC, 21.4% of patients died within eight months of completing their first Integrated Palliative Outcome Score (IPOS)-Renal survey.

Across districts with more than 30 eligible patients, there was substantial variation in the proportion who completed one or more PROMs surveys (Figure 6). For most surveys and cohorts, there was at least one district with zero completed surveys. The National Health Service (NHS) in England reports on PROMs results for organisations with at least 30 patients with PROMs data.[6] For services reaching that threshold, statistical significance testing can be used to determine whether observed differences are real or due to chance. For services that do not reach that threshold, results should not be publicly reported.

Figure 4: Number of eligible LBVC patients completing PROMIS-29 surveys, 2021 for first survey and eight months follow up for subsequent surveys

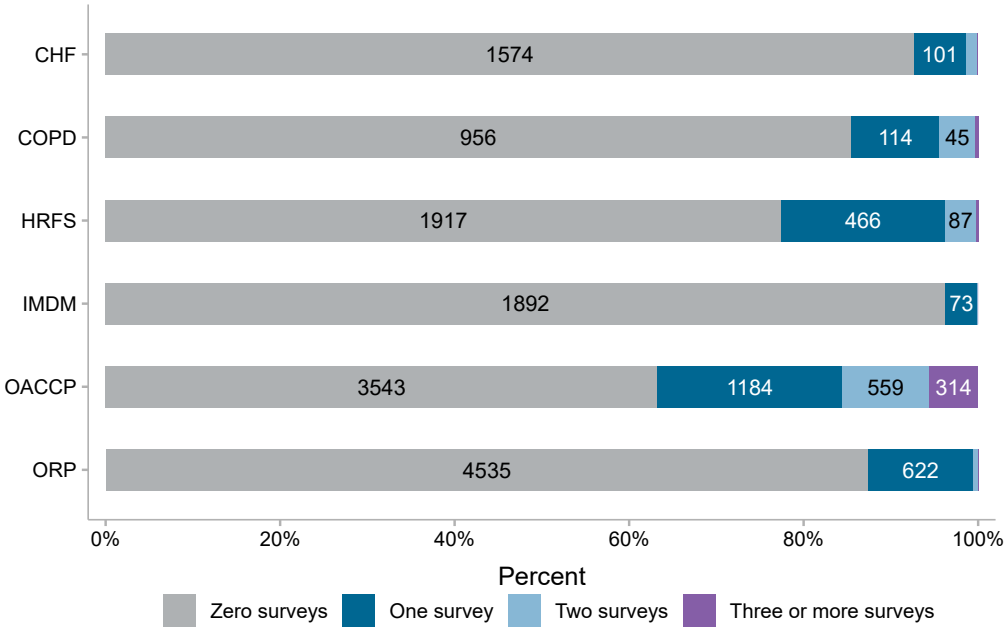


Figure 5: Number of eligible LBVC patients completing condition specific surveys, 2021 for first survey and eight months follow up for subsequent surveys

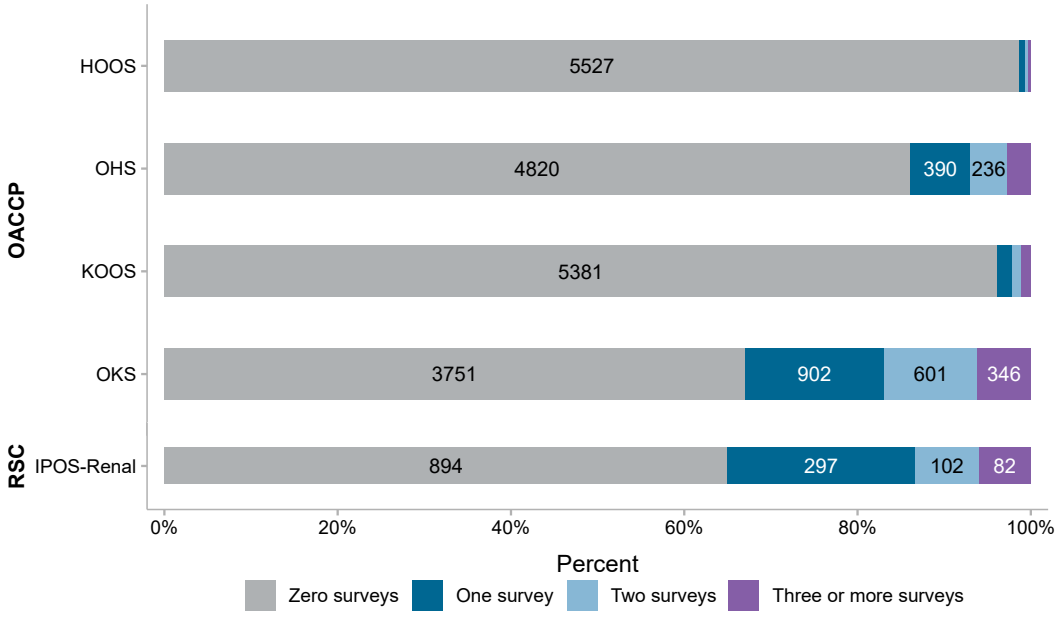
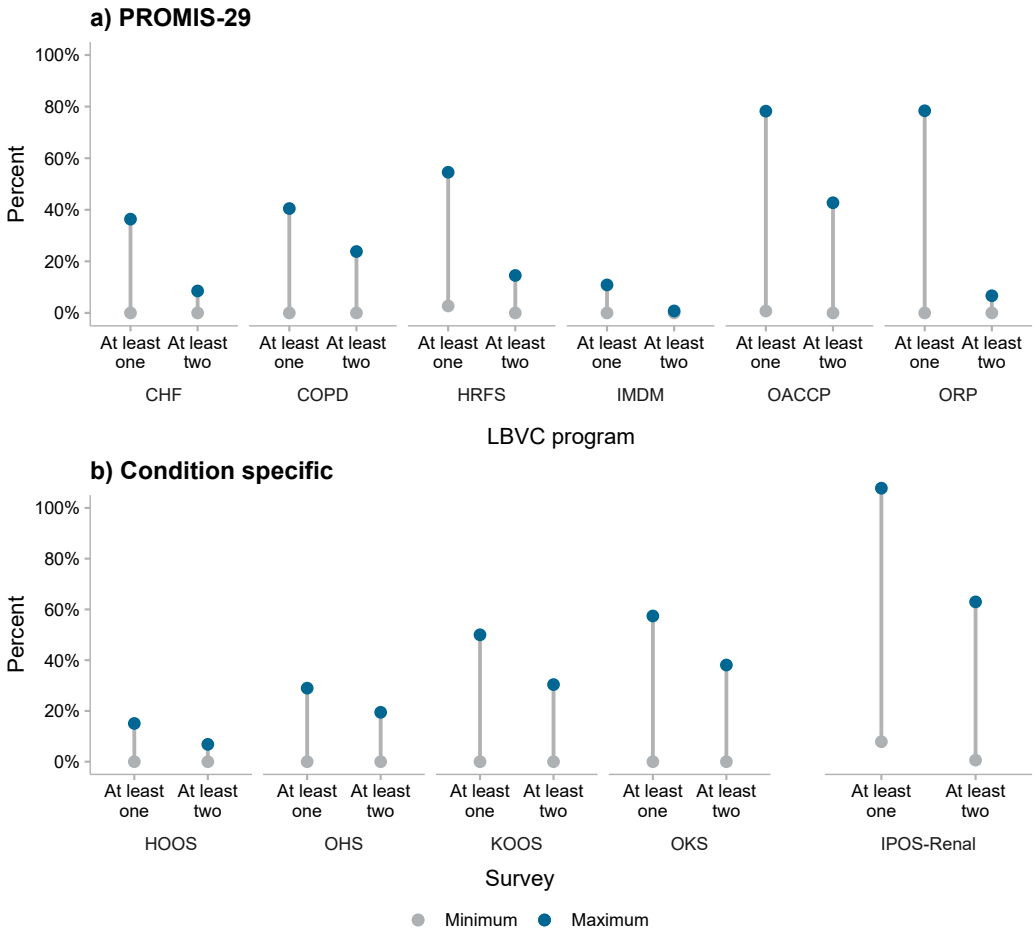


Figure 6: Variation in the percentage of eligible LBVC patients completing PROMs across LHDs, 2021 for first survey and eight months follow up for subsequent surveys



Note: LHDs with fewer than 30 eligible patients are not included.

Assessing the level of missing data

Missing data can refer to missing items on individual PROMs surveys or missing PROMs surveys at recommended collection points.

Missing data can bias the analysis of PROMs data if there is an association between health status and missing data.[8]

It can also limit the ability to assess change in health status over multiple collection points and to detect statistically significant differences between clinics.[6] It is important to assess the level of missing data and account for it in analysis.

Missing items on PROM surveys

In the HOPE system, all items on all surveys are compulsory to complete. Despite this feature, there was a small amount (~5%) of missing data in the HOPE PROMs data in ROVE (Table 6).

For PROMIS-29, we were able to recover some of the missing data by mapping the total raw score to the T-score, where one item was available, and the other item was not available. It was not possible to undertake a similar reconciliation for any of the condition specific surveys.

There were two possible causes for the missing item data:

1. Internet connectivity issues resulting in data not backing up to the servers after a survey was completed. These issues occurred early in the HOPE rollout and have since been rectified.
2. Errors during the data extraction, transfer and/or linkage process resulting in random missing item data in the final linked dataset available for analysis. This error may be rectified in subsequent updates of HOPE PRMs data in ROVE.

Missing data can be imputed[8, 9], but we have not attempted to impute the missing item data in this case because it is expected that this issue will be resolved.

Table 6: Percentage of missing item data on PROMs surveys for LBVC patients, January 2021 to August 2022

LBVC cohort	Survey	Percent missing across all individual items
CHF	PROMIS-29	4.6%
COPD	PROMIS-29	5.6%
HRFS	PROMIS-29	6.2%
IMDM	PROMIS-29	5.1%
OACCP	PROMIS-29	5.4%
	HOOS	4.6%
	KOOS	4.8%
	OHS	4.8%
	OKS	5.3%
ORP	PROMIS-29	5.4%
RSC	IPOS-Renal	6.0%

Missing surveys at recommended collection points

There are recommended collection points for PROMs surveys. For LBVC cohorts, it is generally at the start of the program, every three or six months depending on the cohort and survey, and at program completion. Clinicians are given the option to allocate surveys outside the recommended collection points where clinically indicated.

Surveys may not be allocated regularly if patients are not attending healthcare services regularly, if allocation is not clinically indicated, and/or to reduce survey burden.

For surveys where the recommended collection points are every three months (all OACCP surveys including PROMIS-29), we expect patients to have completed three of them within an eight-month period. We found that between 0.3% and 6.2% had completed at least three surveys. For surveys where the recommended collection points are every six months (PROMIS-29 for ORP, IPOS-Renal), we expect patients to have completed two of them within an eight-month period. We found between 0.6% and 13.4% had completed at least two surveys (Appendix 3).

For service level reporting, where individual patient results will be reviewed and used in clinical decision making, all available data should be used. For system level reporting, where comparisons will be made across districts and clinics on change in PROMs results, it may be preferable to report on two collection points only to maximise the sample size and improve the ability to detect statistically significant differences. The first and second survey could be selected or the first and last survey within the follow up period, depending on what is more clinically appropriate for the given cohort and survey. A rule can be applied for the minimum and maximum time required between the first and second survey, or the first and last survey, for them to be included in analysis, again based on clinical advice for each cohort and survey.

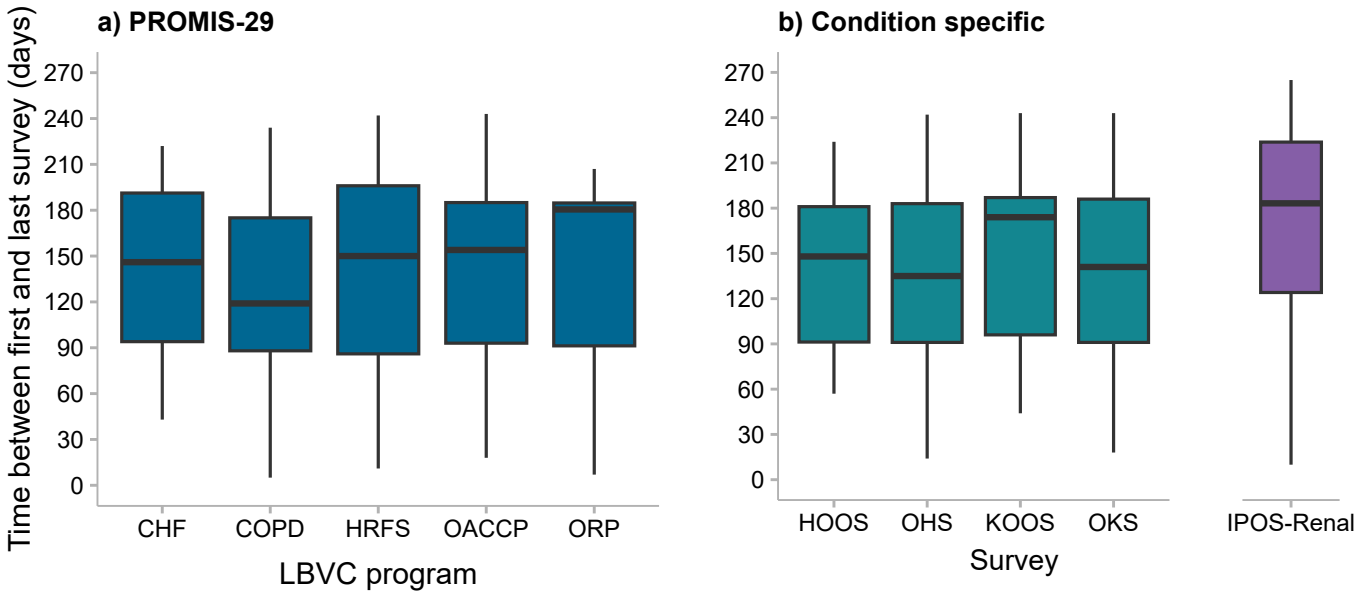
Measuring the time between surveys

The time between the first and second survey or the first and last survey should be measured across the cohort. If the time distribution is similar across the groups that are being compared, fair comparisons can be made. At the NSW level, across conditions and surveys, the time between the first completed survey and the last completed survey within an eight-month follow-up period, ranged from about one month to eight months (Figure 7).

For the primary purpose of PROMs point-of-care data capture and use in HOPE to support real-time health decisions by clinicians and patients, PROMs collection should occur when clinically indicated. However, for the secondary purpose, service and system level evaluation and improvement, it may be worthwhile having at least two compulsory collection points, such as program start and completion.

In England, the NHS does public performance reporting on PROMs data but it is for a well-defined cohort with a specific intervention and only two collection points: pre- and post-operative surveys for hip and knee replacement, groin hernia and varicose vein surgery.[6]

Figure 7: Time (days) between first and last survey for patients who have completed at least two surveys (PROMIS-29 and condition specific), 2021 for first survey and eight months follow up for subsequent surveys



Evaluating respondent representation

When reporting on PROMs data, it is important to assess whether the sample of patients completing PROMs is representative of the cohort.

If it is, the results can be generalised to the cohort. If it is not, there may be a bias in the results.

For example, it may be that patients who are doing well are more likely to complete follow-up surveys, suggesting that an intervention is working well when, in fact, patients who are deteriorating have not completed the survey.[29]

For the LBVC cohorts, the characteristics of the sample of patients completing PROMs were compared with the characteristics of the cohort eligible to complete PROMs, based on age, sex, indigenous status, socioeconomic status (Socio-Economic Indexes for Areas (SEIFA) Index), and comorbidities (Charlson Comorbidity Index – three-year look back at diagnoses recorded in hospital records).

Measures of representativeness were calculated to assess the generalisability of the sample of patients completing PROMs to the eligible cohort using the Histogram Intersection method.[30] This method measures the similarity between two histogram distributions based on the degree of intersection, or overlap, between them. A score of zero indicates the two distributions are completely different while a score of one indicates they are identical.

Respondent representation across varying sample sizes

For OACCP, where PROMIS-29 completion rates were relatively high, the sample characteristics were similar to the cohort characteristics (Figure 8). This was confirmed by the results of the measures of representativeness indicating over 95% intersection on all characteristics (Table 7).

For cohorts with smaller sample sizes, the sample characteristics may not adequately represent the cohort of interest. For example, for the IMDM sample, there were differences in some of the characteristics assessed when compared to the eligible cohort (Figure 9). The sample of patients completing at least one PROMIS-29 were more likely to be older, male and from relatively disadvantaged areas than the cohort eligible to complete PROMs. (See Appendix 4 for CHF, COPD, HRFS, and ORP PROMIS-29 figures.) Similarly, for the OACCP and RSC condition-specific PROMs comparisons, as the sample size increased, the similarity between the eligible cohort and the sample of patients completing PROMs also generally increased (Table 7).

For the surveys with larger sample sizes (IPOS-Renal, OHS, and OKS), there was over 90% intersection on all characteristics. For both the HOOS and KOOS surveys, there was a marked difference in the socioeconomic status of the sample of patients completing PROMs. Patients from more advantaged areas were more likely to complete surveys than those from disadvantaged areas. (See Appendix 4 for condition-specific PROMs figures.)

Table 7: Histogram intersection results for LBVC cohort (total patients after HOPE go-live) and sample (at least one survey), 2021

LBVC cohort	Survey	Total patients after HOPE go-live	Sample (at least one PROM)	Age	Sex	Indigenous status	SEIFA quintiles	Charlson comorbidity score
CHF	PROMIS-29	1699	125	0.880	0.974	0.996	0.812	0.927
COPD	PROMIS-29	1118	163	0.862	0.951	0.986	0.908	0.902
HRFS	PROMIS-29	2476	559	0.952	0.960	0.986	0.952	0.976
IMDM	PROMIS-29	1966	74	0.730	0.811	0.983	0.771	0.955
OACCP	PROMIS-29	5600	2057	0.981	0.999	0.997	0.963	0.991
	HOOS		73	0.830	0.935	0.965	0.552	0.965
	KOOS		219	0.891	0.983	0.987	0.575	0.982
	OHS		780	0.953	0.947	0.998	0.920	0.964
	OKS		1849	0.968	0.997	0.995	0.924	0.980
ORP	PROMIS-29	5187	652	0.941	0.992	0.992	0.910	0.952
RSC	IPOS-Renal	1375	481	0.900	0.997	0.993	0.925	0.977

Figure 8: Representative comparison for OACCP cohort (total patients after HOPE go-live) and sample (at least one PROMIS-29), 2021

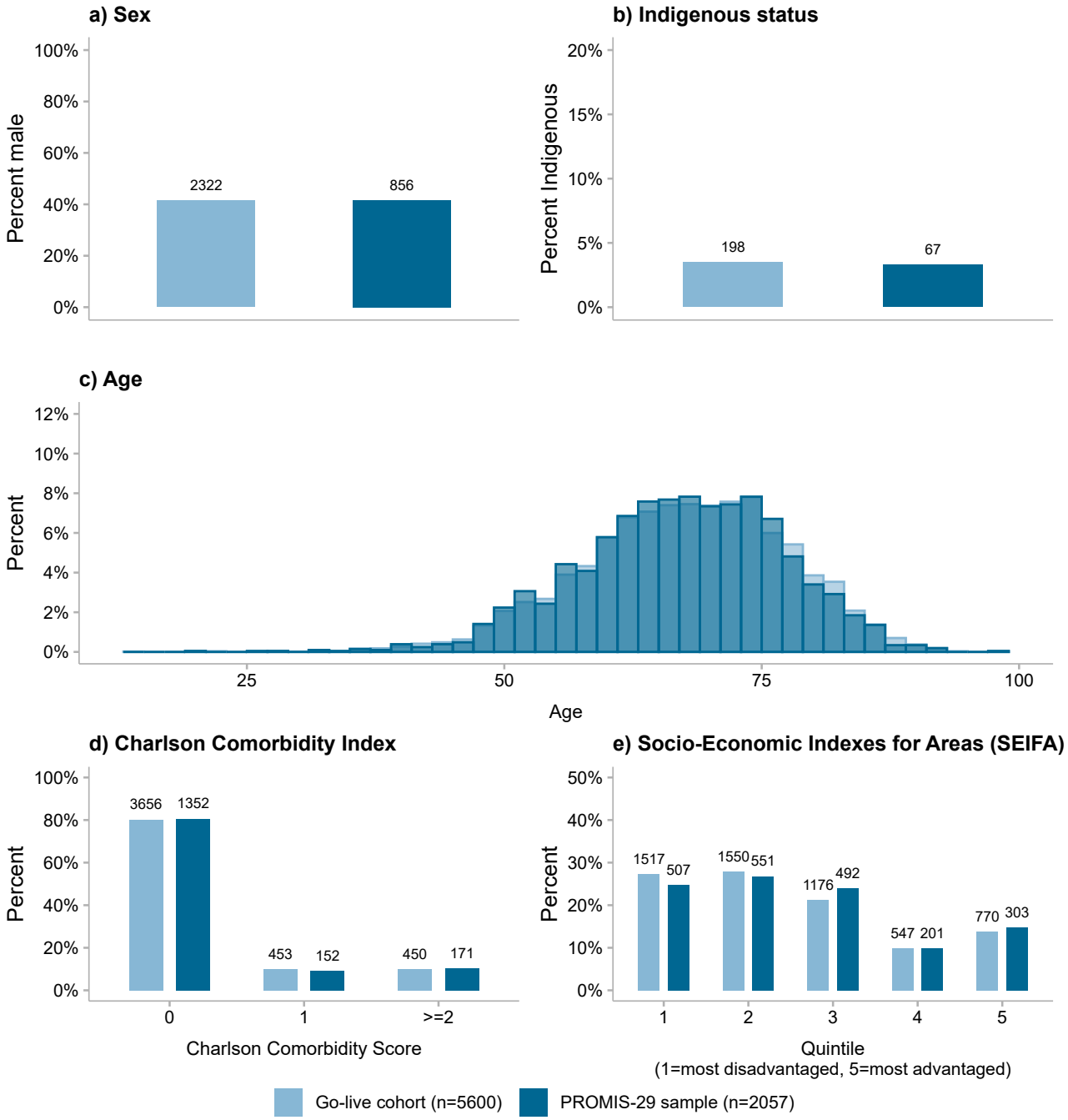
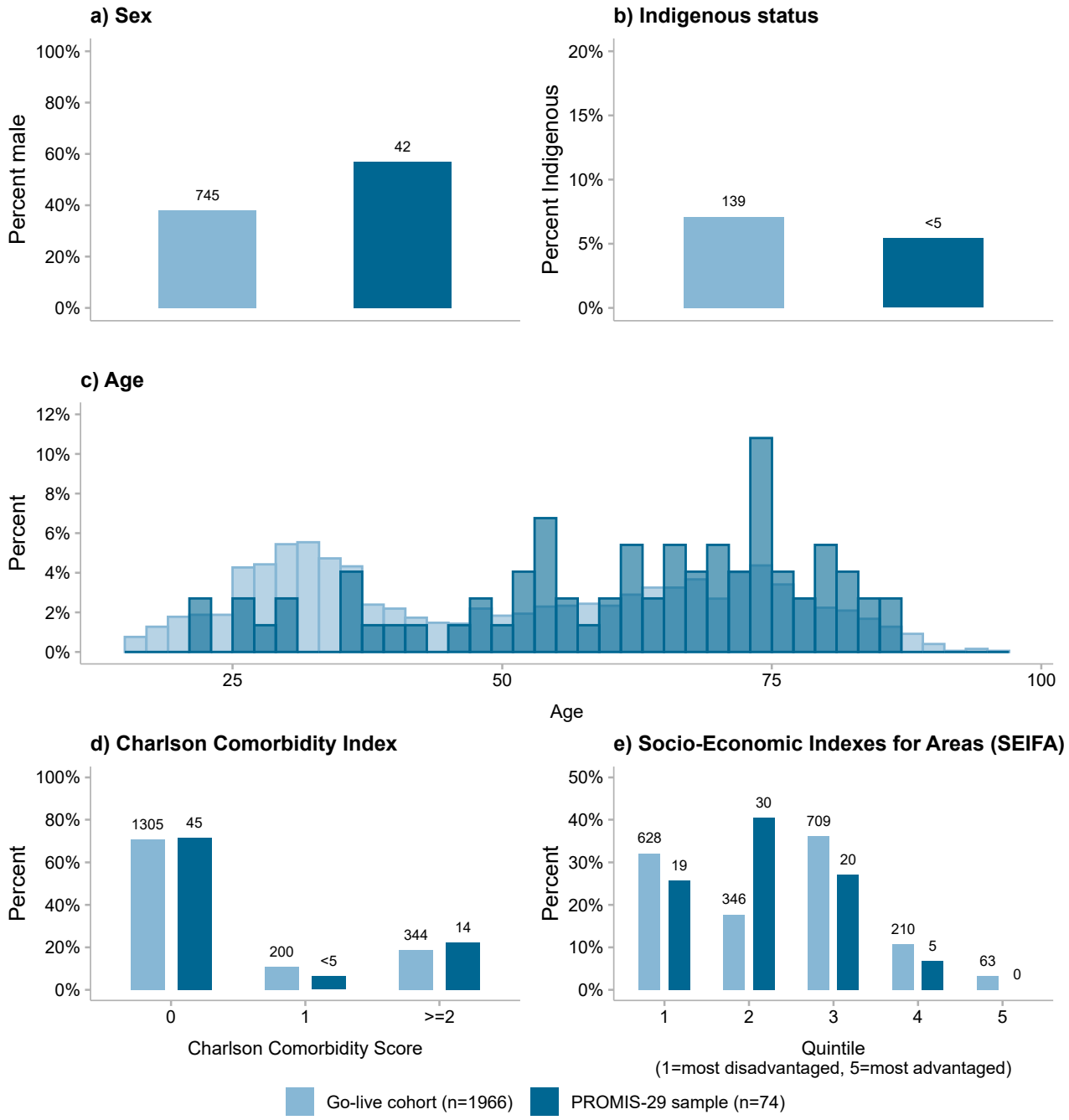


Figure 9: Representative comparison for IMDM cohort (total patients after HOPE go-live) and sample (at least one PROMIS-29), 2021



Survey weighting

To adjust for under or over representation of key characteristics in a sample, survey responses can be weighted to make the sample more representative of the population of interest. However, caution is required because weighting can also introduce additional biases and over-represent the responses of some people who may not be an accurate reflection of the entire group.

Weighting is recommended when there is at least a 5% difference in the distribution of characteristics between the sample and the population.[10]

There are several methods that can be used to weight survey responses including:

- Raking – for each variable to be used in weighting, the population proportions are divided by the sample proportions to get a weight for each category. The sample in each category is multiplied by these weights so the sample distribution matches the population distribution. This procedure is repeated for each variable using the previous weighted results. If subsequent weighting pushes the distribution of variables that were previously weighted out of alignment, adjustments are made (Figure 10).[11-13] Raking is used by government agencies, such as the NSW Bureau of Health Information and the Australian Bureau of Statistics, to weight public opinion surveys.[31, 32]
- Matching – a target sample of cases that are representative of the population is created. Each case in the target sample is paired with the most similar case from the actual sample. When all cases in the target sample have been matched, any unmatched cases from the actual sample are discarded.[33]
- Propensity weighting – each case is weighted by the inverse of their probability of selection. For probability-based surveys, the selection probabilities are known from the sample design. For opt-in surveys, the selection probabilities are unknown and are estimated.[12, 34, 35]
- A combination of methods, matching or propensity weighting followed by raking, can be more effective in reducing bias than a single method, although this is not common practise.[12]

Selecting variables for survey weighting

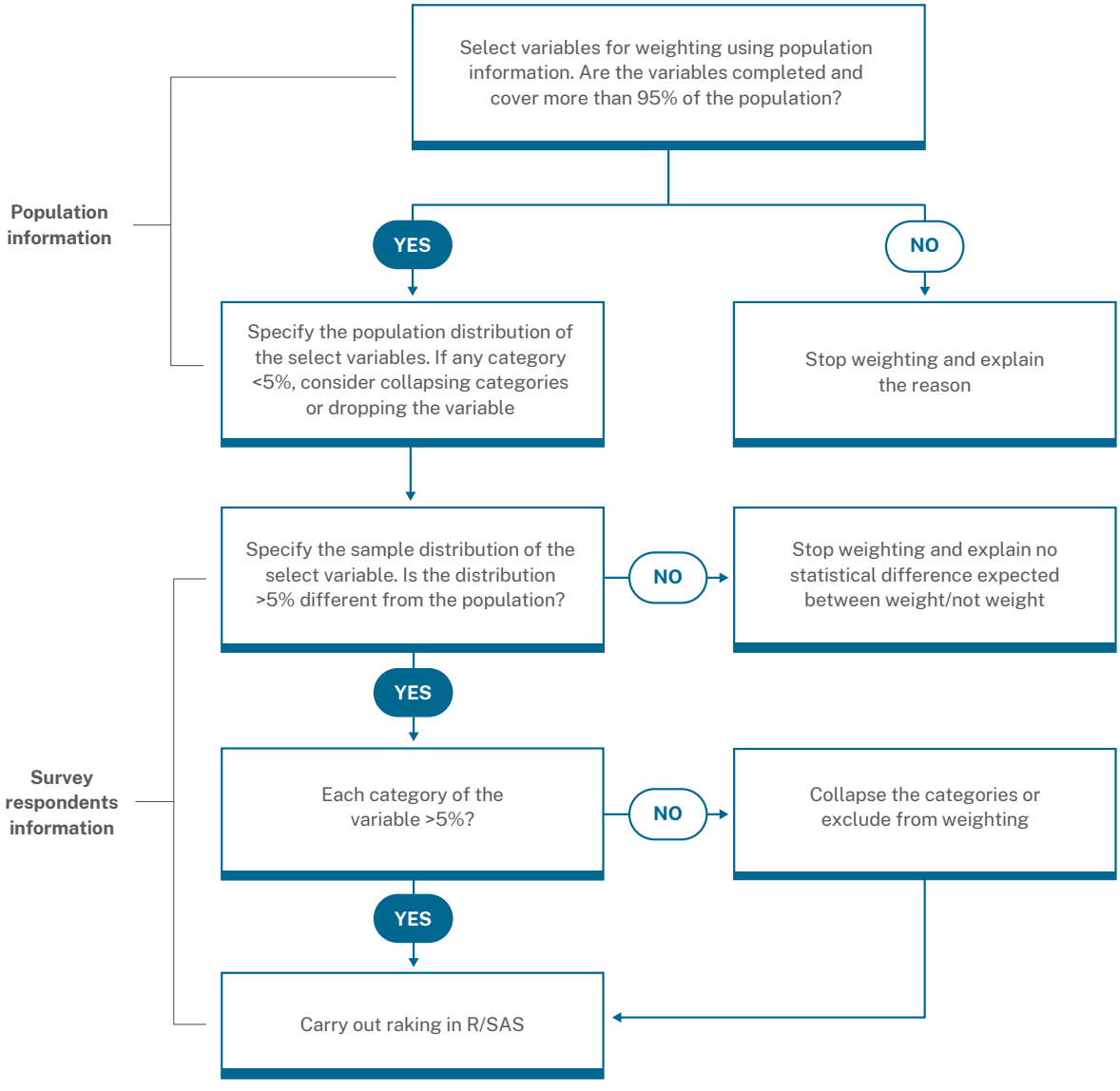
The effectiveness of survey weighting in reducing bias depends more on the variables selected for weighting than the weighting method used.[13, 36] Only variables with high accuracy and completeness (>95%) should be used.[10, 12, 13] Variables where some categories are less than 5% should not be used or the categories should be truncated.[10, 13] For example, Aboriginal status should not be used in weighting because less than 5% of the population is Aboriginal in NSW.

Various characteristics can be considered for weighting. For the PROMIS-29 survey, age, sex, education, occupation, income, household size, marital status, ethnicity, chronic disease, body mass index, self-rated general health, and survey administration mode have been used for weighting data.[37-40]

It is recommended to weight by as few variables as possible to avoid adverse interactions between the weighting variables.[10, 13] It is also advised to limit the number of categories in a variable used for weighting to no more than five.[10] It is ideal to obtain a highly representative sample at the outset as even the most effective weighting strategies can only reduce sample bias by about 30%.[41]

When HOPE implementation is further progressed and more data is available, sample representativeness will be reassessed, and consideration given to whether weighting is required. If it is required, thorough scoping of variables for weighting will be undertaken for each cohort.

Figure 10: Flowchart for variable selection for weighting surveys using the raking method



Assessing floor and ceiling effects

Floor and ceiling effects occur in measures or surveys when a substantial proportion of respondents score the best or worst score possible.

It means the measure is unable to discriminate between respondents at either end of the scale. It may also mean that it is not possible to detect change in the measure for some respondents over time. Floor or ceiling effects are considered present if 15% or more of respondents had the best or worst score possible.[15, 16]

Floor and ceiling effects were investigated in the PROMIS-29 and condition specific surveys using all completed surveys from January 2021 to August 2022.

For PROMIS-29 overall, there were many individual items where at least 15% of patients reported the best or worst score possible. For total domain scores, at least 15% of patients reported the best score on anxiety and depression (Figure 11).

For PROMIS-29 by LBVC cohort, there were many cases where at least 15% of patients reported the best total domain score. The domains in which this occurred varied by condition, although anxiety and depression were common across all cohorts (Figure 12).

For OACCP condition specific surveys, at least 15% of patients reported the worst score possible on the quality of life and function in the sports/recreation dimensions in both the HOOS and KOOS surveys (Figure 13).

For IPOS Renal, at least 15% of patients reported the best score possible on 22 out of 23 items (Figure 14). When the items were aggregated into a total IPOS Renal score, the best and worst score possible was reported by 0.14% and 0.14% of patients respectively (Figure 14).

Floor and ceiling effects should be assessed when measuring change over time on PROMs surveys. Survey results are commonly reported as the percentage of people who:

- improved
- deteriorated
- experienced no change.

Given some people may report the best or worst score possible on their first survey, two extra categories should be included when measuring change:

- ceiling (patient reported the best score on both surveys)
- floor (patient reported the worst score on both surveys).

When floor and ceiling effects have been identified, stratification of aggregate score results by subgroups, such as age or treatment level, may reveal floor and ceiling effects confined to a specific sub-group.[19] Advanced modelling methods, such as a Tobit model, can be used to analyse change in health status associated with a covariate of interest, and is suited to censored or truncated data.[17, 18]

Figure 11: Percent of responses with floor and ceiling effects for PROMIS-29 items and domains across all LBVC cohorts (n=5054), January 2021 to August 2022

Figure 12: Percent of responses with floor and ceiling effects for PROMIS-29 domains by LBVC cohort, January 2021 to August 2022

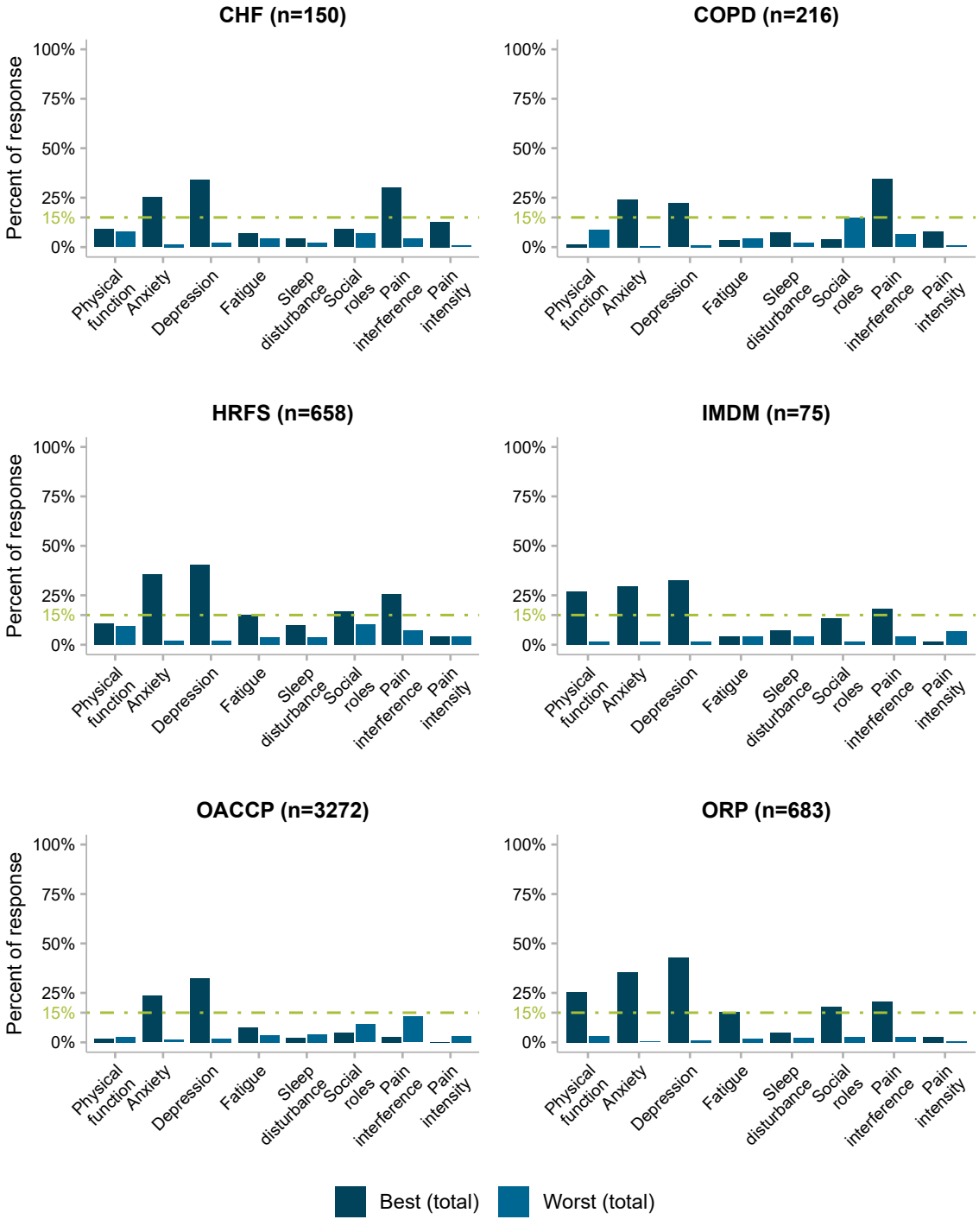


Figure 13: Floor and ceiling effects for condition specific surveys in OACCP, January 2021 to August 2022

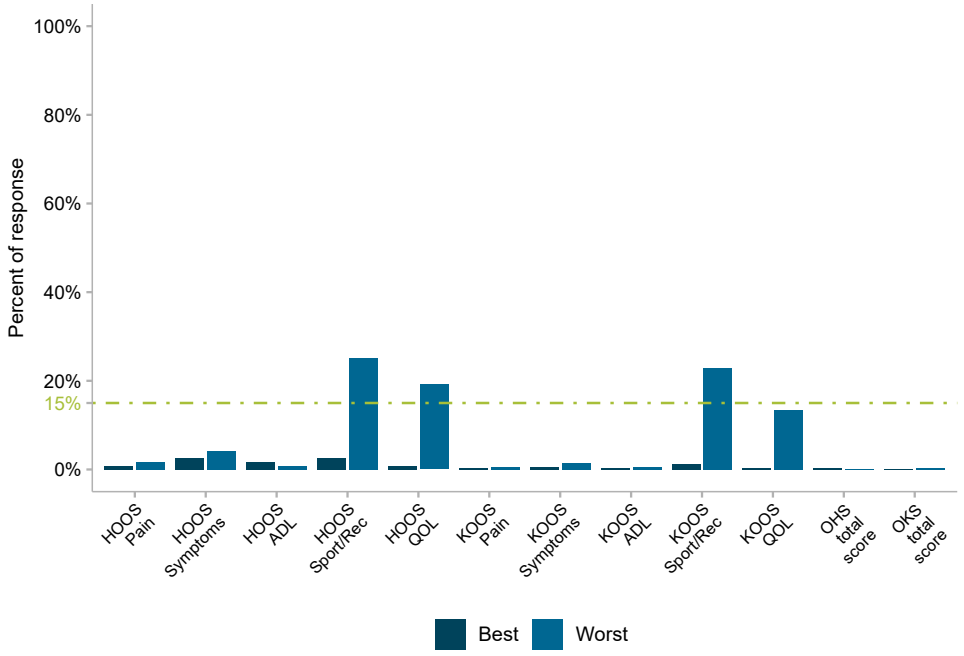
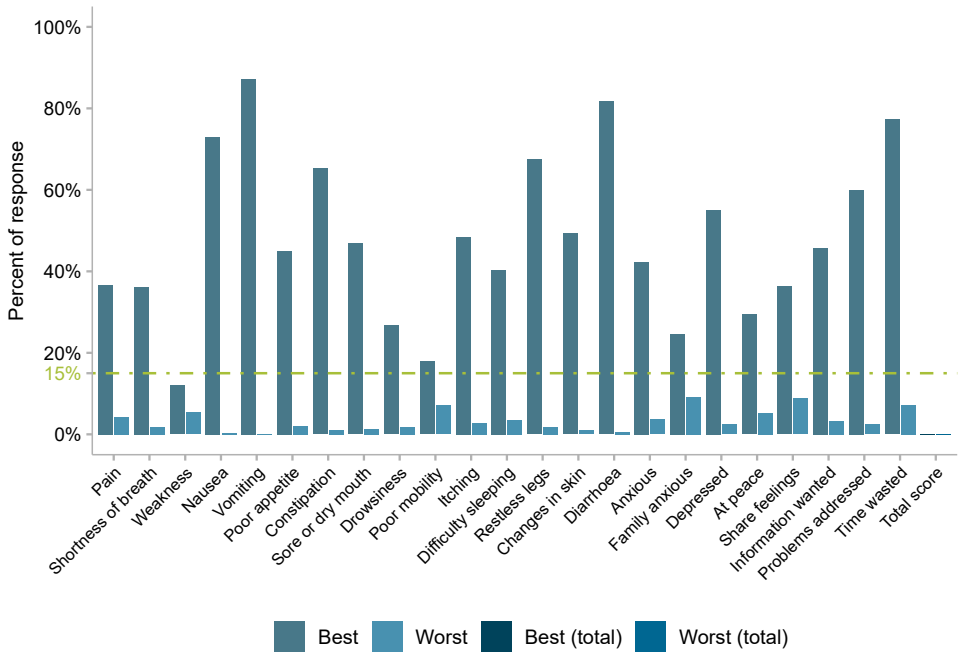


Figure 14: Floor and ceiling effects for IPOS Renal surveys (n=783), January 2021 to August 2022



Producing descriptive statistics on baseline PROMs

While outcomes can only be assessed when PROMs data are collected at multiple time points, it is still informative to review single PROMs results; especially when first collected at the start of a new healthcare program or intervention.

At the individual level, reviewing baseline PROMs results can guide shared treatment decisions between clinician and patient. At the service and system level, baseline PROMs results can be used to understand the health status of the cohort when starting a healthcare program. It may also highlight differences in case-mix and inequities in accessing care across services.[20] Some patient groups may be presenting to services at a later stage in disease progression and have:

- greater symptom burden
- lower functional status
- and/or lower quality of life as reported through the PROMs surveys.

Baseline or first survey

Implementation of the HOPE platform began in February 2021 and endorsed cohorts in admitted and non-admitted locations have been progressively going live in HOPE. As a result, the first PROMs survey completed by LBVC patients in the HOPE platform in 2021 may not be their baseline survey filled out when they started the program. In future, when the HOPE platform is embedded in clinical practice, the first survey completed by patients may be their baseline survey.

Among the LBVC patients in 2021 who completed their first PROMIS-29 survey, 2,555 out of 3,630 (70%) were new patients and their first survey could be their baseline survey. (They started attending the LBVC NAP clinic after it went live in HOPE.) Across LBVC conditions, between 23% and 83% were new patients. Among these new patients, 77% completed their first survey within four weeks of their first NAP visit. Across LBVC conditions, it ranged from 58% to 82% (Table 8).

Among the patients who completed their first OACCP condition specific survey in 2021, 77% were new patients. Among these new patients, 80% completed their first survey within four weeks of their first NAP visit. For IPOS Renal, 43% were new patients and, among these new patients, 64% completed their first survey within four weeks of their first NAP visit.

Table 8: Time between first LBVC NAP clinic visit after HOPE go live and completion of first PROMs survey for new patients, 2021

LBVC cohort	Survey	Total patients after HOPE go-live	Sample (at least one PROM)	New patients (n, %)	Time between first NAP service and first survey completion		
					Median (days)	Within ± 2 weeks (%)	Within ± 4 weeks (%)
CHF	PROMIS-29	1699	125	74 (59%)	13	50%	58%
COPD	PROMIS-29	1118	163	37 (23%)	5	68%	73%
HRFS	PROMIS-29	2476	559	271 (49%)	14	51%	68%
IMDM	PROMIS-29	1966	74	26 (35%)	0	58%	62%
OACCP	PROMIS-29	5600	2057	1603 (78%)	0	77%	82%
	HOOS		73	62 (85%)	0	79%	82%
	KOOS		219	159 (73%)	0	69%	75%
	OHS		780	624 (80%)	0	77%	80%
	OKS		1849	1415 (77%)	0	76%	80%
ORP	PROMIS-29	5187	652	544 (83%)	0	63%	72%
RSC	IPOS-Renal	1375	481	206 (43%)	8	57%	64%

Survey completion by proxy

In the HOPE platform, PROMs surveys can be completed by the patient or by a proxy (carer or legal guardian). This feature is important to make PROMs surveys more accessible, however it may introduce some error or bias in the responses.[42] The percentage of first surveys completed by a proxy ranged from 4% to 12% for PROMIS-29 across LBVC cohorts and from 7% to 14% for the OACCP condition-specific PROMs (Figure 15). A higher proportion of surveys were completed by a proxy for IPOS Renal compared with the other surveys (27%) (Figure 15).

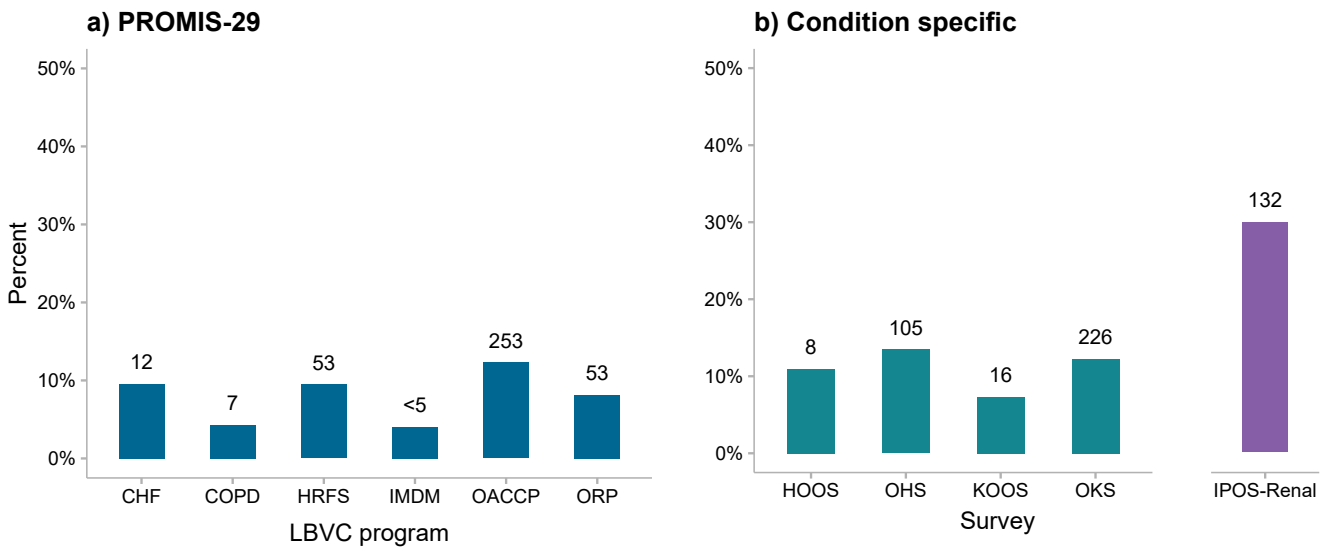
Response patterns

Response patterns were investigated for PROMs surveys, specifically the uniformity of responses (the proportion of patients selecting the same response for all items in a survey). Across LBVC programs and surveys, this pattern occurred in less than 1% of first surveys in 2021.

Descriptive analysis

Descriptive analysis of the first PROMs survey completed by LBVC patients in 2021 was conducted. The results are presented in Figures 16 to 25 for PROMIS-29, OACCP condition specific surveys and IPOS Renal. Note, there was a small amount (~5%) of missing data in the HOPE PROMs data in ROVE (see Missing Data Section for more information). As a result, there is variation in respondent numbers to individual items, subscale scores, total scores, and the survey in general.

Figure 15: Percentage of surveys completed by proxy (carer or guardian), first survey completed in 2021

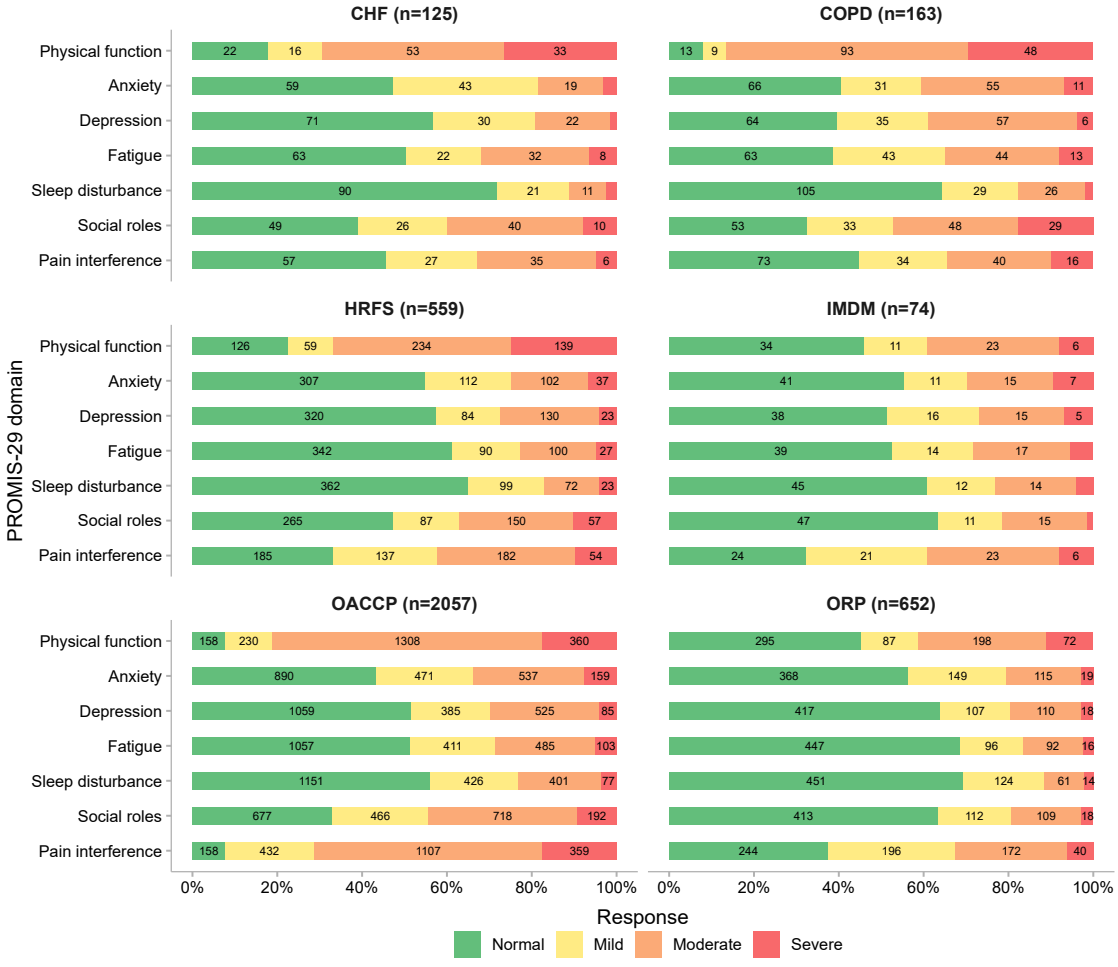


PROMIS-29 domain severity groups

Across the seven domains on the PROMIS-29 survey, physical function and pain interference tended to have the lowest proportion of LBVC patients reporting normal symptoms on their first survey. Sleep disturbance had the highest proportion reporting normal symptoms. COPD patients generally reported worse symptoms than CHF patients, and OACCP patients reported worse symptoms than ORP patients (Figure 16).

Information on the correlation between domains is included in Appendix 5.

Figure 16: PROMIS-29 domain severity groups, first survey completed in 2021

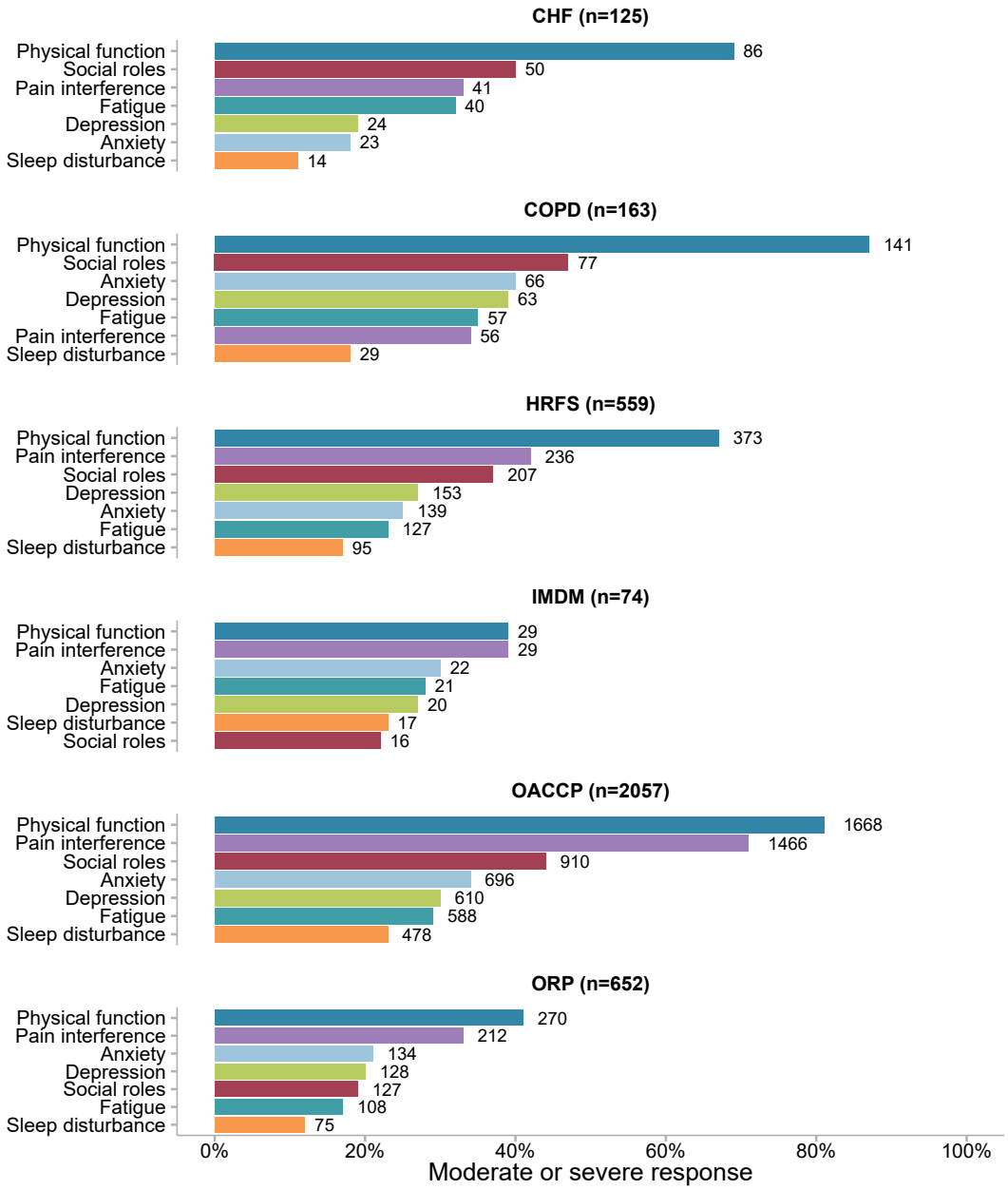


Note. Numbers for some domains may not add up to first survey sample size due to missing data.

PROMIS-29 domain moderate or severe response

For all LBVC conditions, the physical function domain had the highest proportion of patients reporting moderate or severe symptoms. For CHF and COPD, social roles was second highest. For HRFS, IMDM, OACCP, and ORP, pain interference was second highest (Figure 17).

Figure 17: PROMIS-29 domain moderate or severe response, first survey completed in 2021

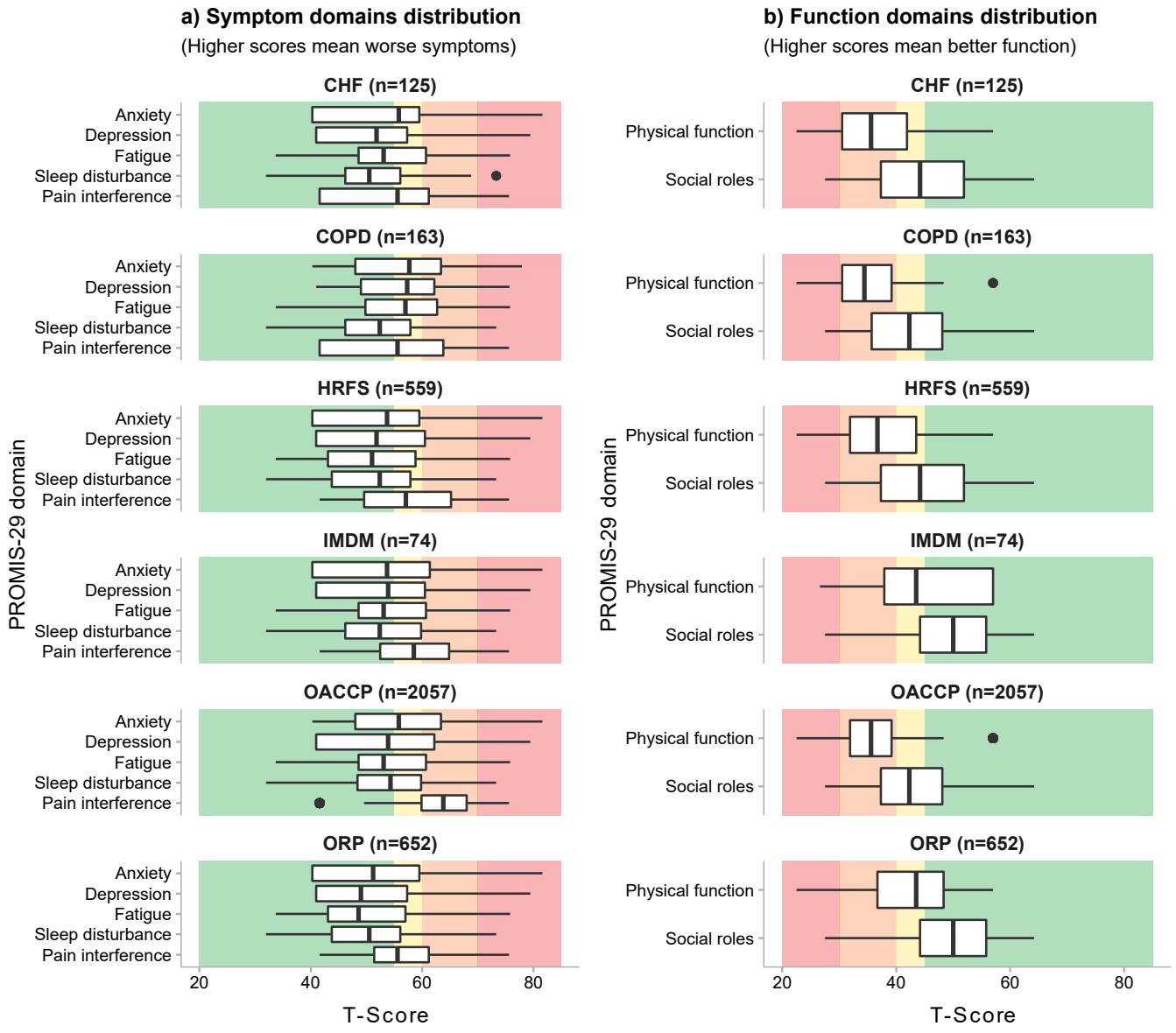


PROMIS-29 domain T-scores

PROMIS-29 domain scores can be converted to T-scores, where a T-score of 50 represents the average expected score based on references developed in the United States general population.

Across most LBVC conditions in the symptom domains, at least 50% of patients had normal symptoms based on the population comparison. For the physical function domain, most LBVC patients had mild, moderate, or severe function based on the population comparison (Figure 18).

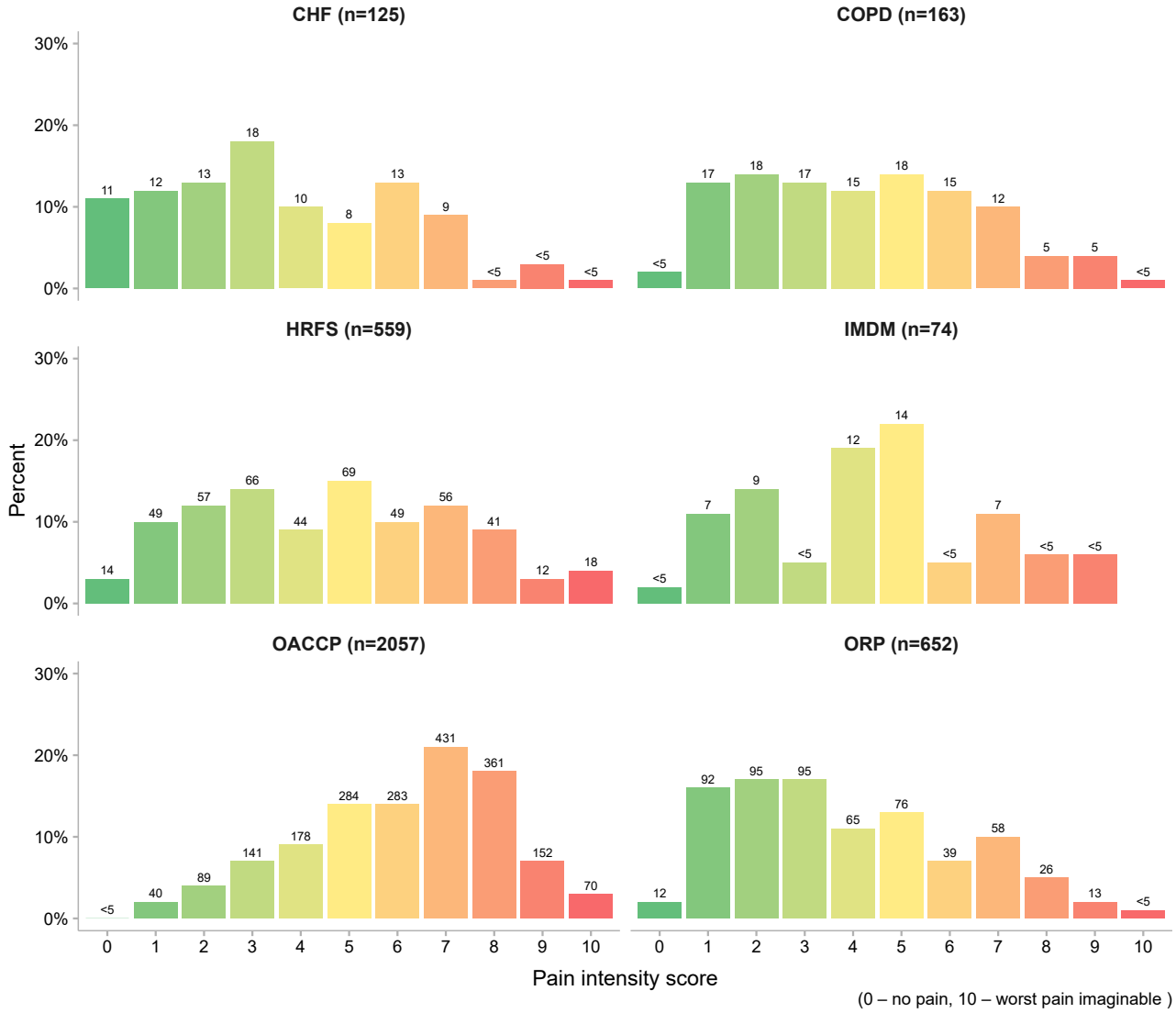
Figure 18: PROMIS-29 domain T-scores, first survey completed in 2021



PROMIS-29 pain intensity

As well as the seven domains, PROMIS-29 includes a pain intensity item with a 0-10 numeric rating scale (0 = no pain, 10 = worst pain possible). The median pain score reported was 4 for CHF, 5 for COPD, 6 for HRFS, 5.5 for IMDM, 7 for OACCP, and 4 for ORP. For all LBVC conditions, there were some patients reporting high pain intensity (7 or above) (Figure 19).

Figure 19: PROMIS-29 pain intensity scores, first survey completed in 2021



Note. Numbers in the figure may not add up to the first survey sample size due to missing data on the pain intensity item.

OACCP HOOS and KOOS

The developers of the HOOS and KOOS surveys recommend visualising the data by plotting the mean dimension scores in a HOOS/KOOS profile (0 = extreme symptoms, 100 = no symptoms) (Figures 20 and 21). From these graphs we can see that both hip and knee osteoarthritis patients report higher mean scores for activities in daily living, indicating least difficulty in this area. They report lower mean scores for sport and recreation and quality of life, indicating greatest difficulty in these areas.

Figure 20: OACCP HOOS mean profile scores (n=73), first survey completed in 2021

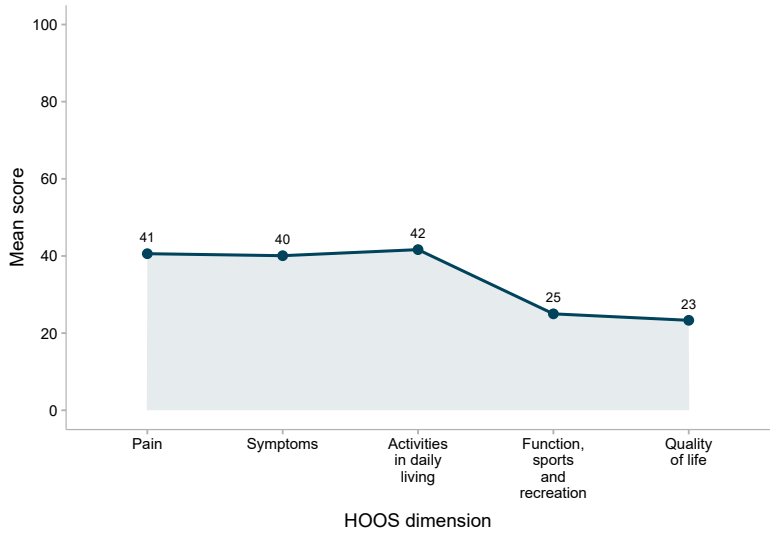
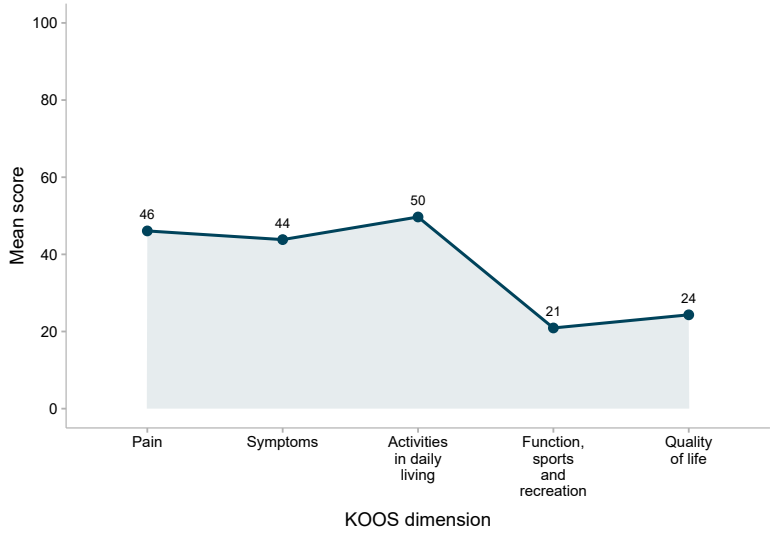


Figure 21: OACCP KOOS mean profile scores (n=219), first survey completed in 2021



OACCP Oxford hip score and Oxford knee score

On the OHS and OKS surveys, symptom scores ranging from 0 (most severe symptoms) to 48 (least symptoms) were reported by hip and knee osteoarthritis patients (Figures 22 and Figure 23). For OHS, the median score was 19 for both left hip and right hip. For OKS, the median score was 23 for left knee and 22 for right knee.

Figure 22: OACCP OHS total score distribution (n=780), first survey completed in 2021

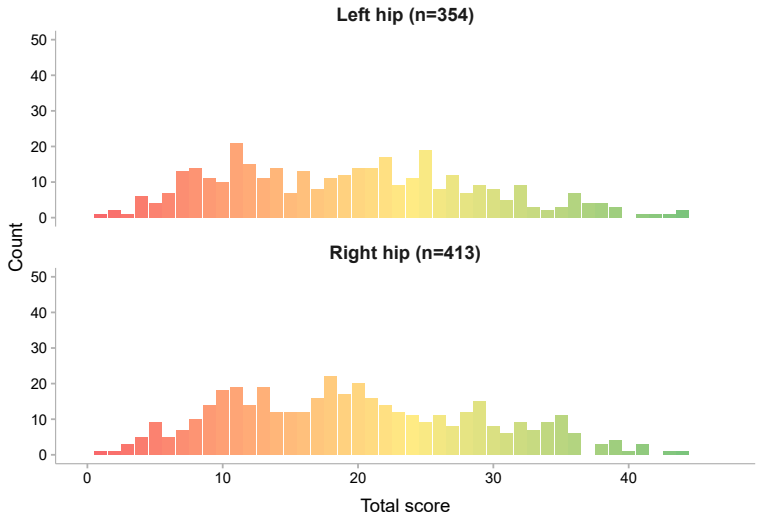
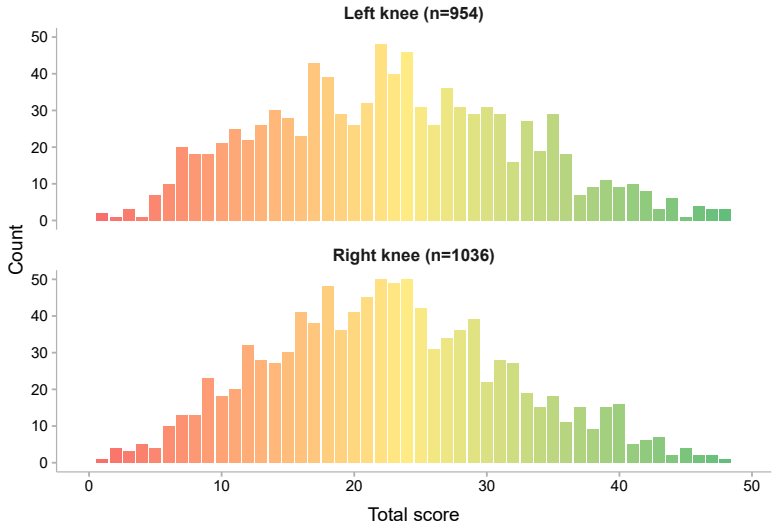


Figure 23: OACCP OKS total score distribution (n=1849), first survey completed in 2021

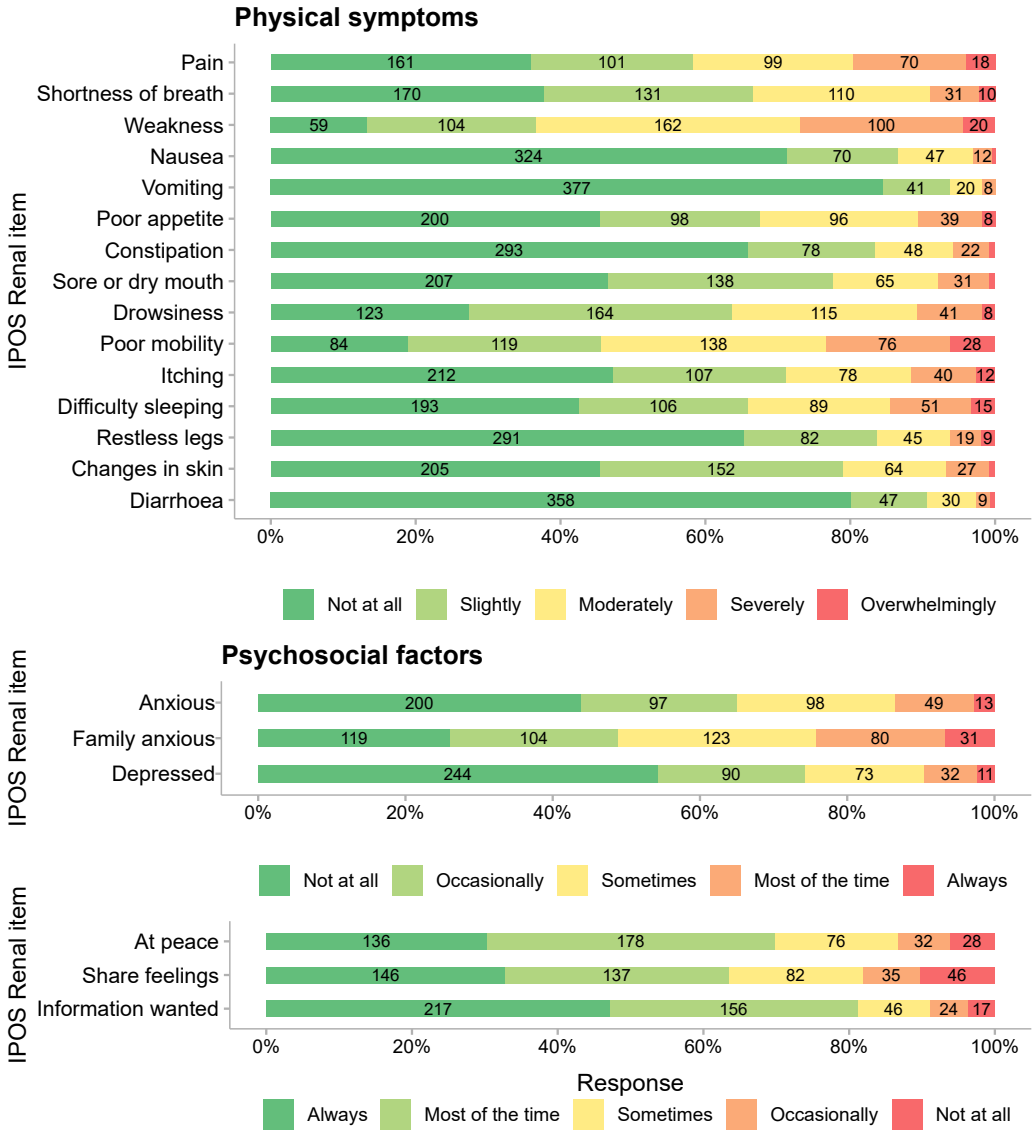


Note. Numbers in the figure may not add up to the first survey sample size due to missing data; patients may complete multiple surveys if both hips or knees are affected.

RSC IPOS Renal item response

The most common physical symptoms reported by renal supportive care patients as severe or overwhelming were weakness (27%) and poor mobility (23%). The most common psychosocial factors reported were family or friends being anxious or worried (24% most of the time or always) and being able to share how you are feeling (18% occasionally or not at all) (Figure 24).

Figure 24: IPOS Renal ratings (n=481), first survey completed in 2021

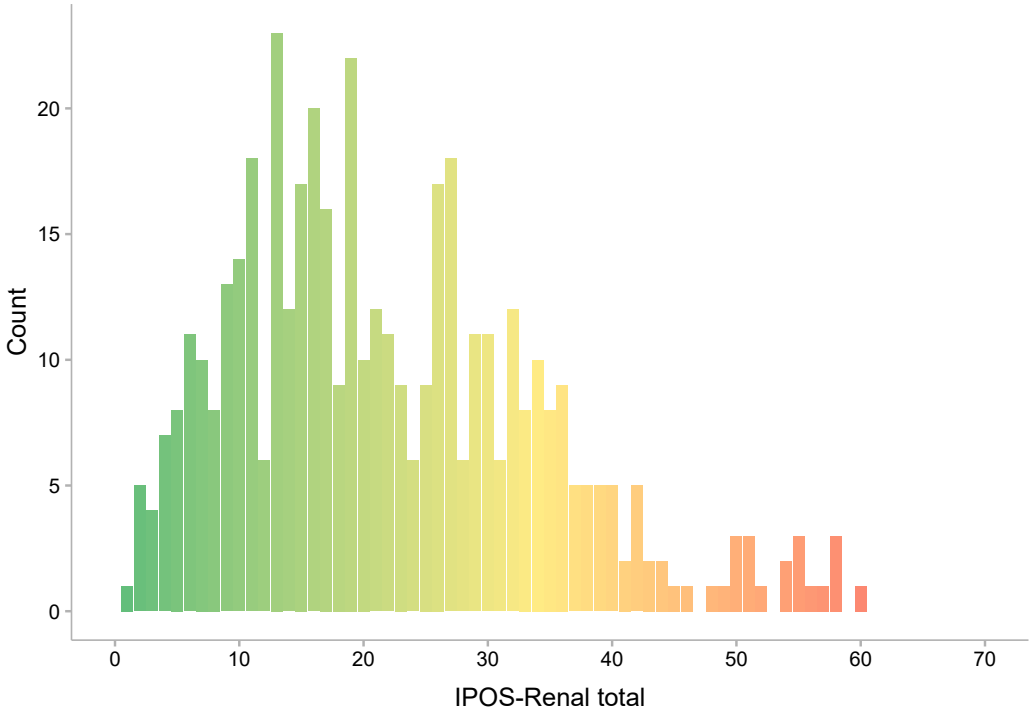


Note. Numbers for some items may not add up to the first survey sample size due to missing data.

RSC IPOS Renal overall score

The IPOS Renal total score indicates the overall symptoms, concerns, and status of the patient at a specific point in time and ranges from 0 (best outcomes) to 90 (worst possible outcomes). The median score was 20 for renal supportive care patients in 2021 (Figure 25).

Figure 25: IPOS Renal overall score distribution (n=481), first survey completed in 2021



Applying robust statistical methods when measuring change in PROMs results over time

When PROMs data are collected at multiple time points, before and after or throughout a healthcare program or intervention, outcomes can be assessed.

At the individual level, clinicians can use PROMs data collected at multiple time points to monitor patient outcomes and identify factors associated with improvement or deterioration in health. At the system level, they can be used to benchmark outcomes across districts, providing evidence for best practice, setting standards and expectations, and potentially driving improvement.[25]

When measuring change over time, and especially when comparing across districts or clinics, it is important to consider the case mix of patients completing the surveys,[6, 21] and potential confounders that may impact results. Minimally important difference (MID) should be considered to aid in clinically meaningful interpretation of PROMs. This is the smallest change perceived as important by the patient or clinician and which could lead to a change in treatment.[1, 22, 23] Thresholds for MID may differ at the individual and group level.[24]

Descriptive analysis of the first and last PROMs survey completed by LBVC patients was conducted, with the first survey completed in 2021 and an eight-month follow-up period applied for the last survey based on data availability. The results are presented below for PROMIS-29, OACCP condition specific surveys and IPOS Renal.

Care should be taken when interpreting outcomes

Data is included only for cohorts and surveys where the sample size was at least 30 respondents. When reviewing the results, it is important to note that this sample size is still considered low for statewide analysis and care should be taken when interpreting outcomes. The figures and tables are intended to provide examples of how change over time in PROMs can be reported and visualised. In coming years, when HOPE implementation is complete and the numbers of completed surveys is higher, appropriate meaning can be derived from the results.

Other considerations to keep in mind include:

- Given PROMs completion rates, variation in collection points and data availability, first and last survey within an eight-month period were selected to assess outcomes to maximise sample size.
  - These surveys may have been completed at any point in the patient's health trajectory and healthcare journey.
  - For system level assessment of patient outcomes, it may be appropriate to designate two compulsory collection points relative to the patient's healthcare pathway, e.g. pre- and post-surgery, or admission and discharge from care.
- For most cohorts and surveys, the number of PROMs available for analysis was low and may not be sufficient to assess change at the group level.
- Some surveys may not be sensitive in detecting change in patient outcomes during routine care.[43]

- Several factors may affect the outcomes reported by patients, such as disease severity and comorbidities. These factors will be accounted for in future analyses.
- Most patients completing these surveys have chronic conditions and their symptoms may get worse over time despite good care.
- Some of the domains assessed in the PROMs may not have been targeted for improvement in the healthcare program.

PROMIS-29 domain severity groups

Follow up PROMIS-29 surveys were included if they occurred at least two weeks after the first survey. Change in scores was calculated as the raw score difference between the first and last survey. For COPD, sleep disturbance scores had the highest percentage improvement, for HRFS it was physical function, and for OACCP it was pain interference. Minimal floor effects were identified (patients maintaining the worst possible score on both first and last surveys). Anxiety (HRFS), depression (HRFS, OACCP), and pain interference (HRFS, COPD) domain scores had at least 15% of patients displaying no change due to ceiling effects for some LBVC cohorts (patients maintaining the best possible score on first and last surveys) (Figure 26).

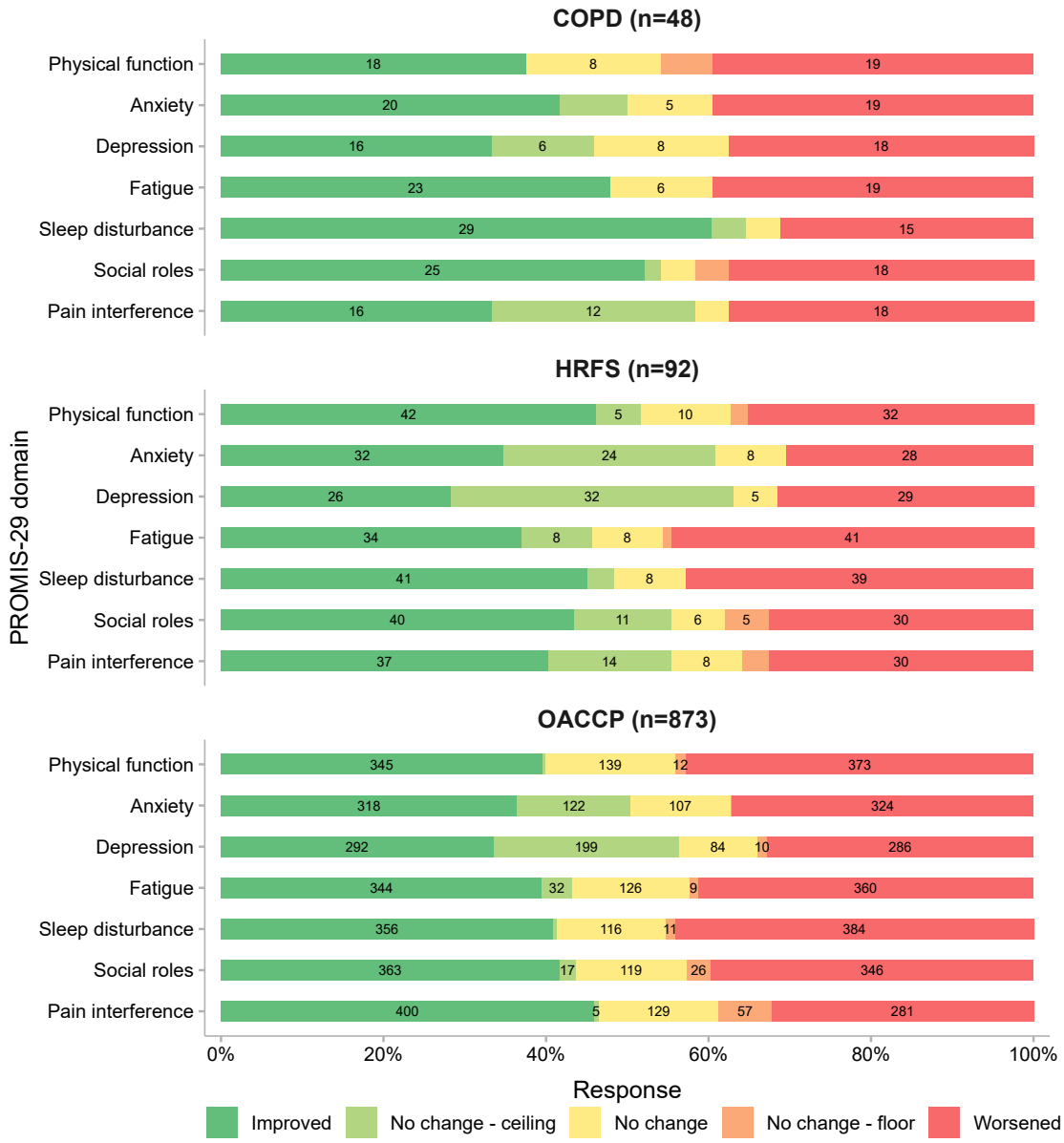
PROMIS-29 pain intensity

Nearly a third of patients reported reduced pain from first to last survey (COPD: 33%; HRFS: 29%; OACCP: 39%; Figure 27). About a quarter of patients reported the same level of pain and about a third reported worse pain. Few patients had the worst pain intensity (score of 10) at both the first and last survey (0-2% of patients across all cohorts). Also, few had the lowest pain intensity (score of 0) at both the first and last survey (0-7% of patients in each cohort). The median pain intensity scores were similar at both first and last surveys for all three cohorts (COPD: 3.9 to 3.7; HRFS: 4.4 to 4.2; OACCP 6.1 to 5.9).

PROMIS-29 minimally important difference

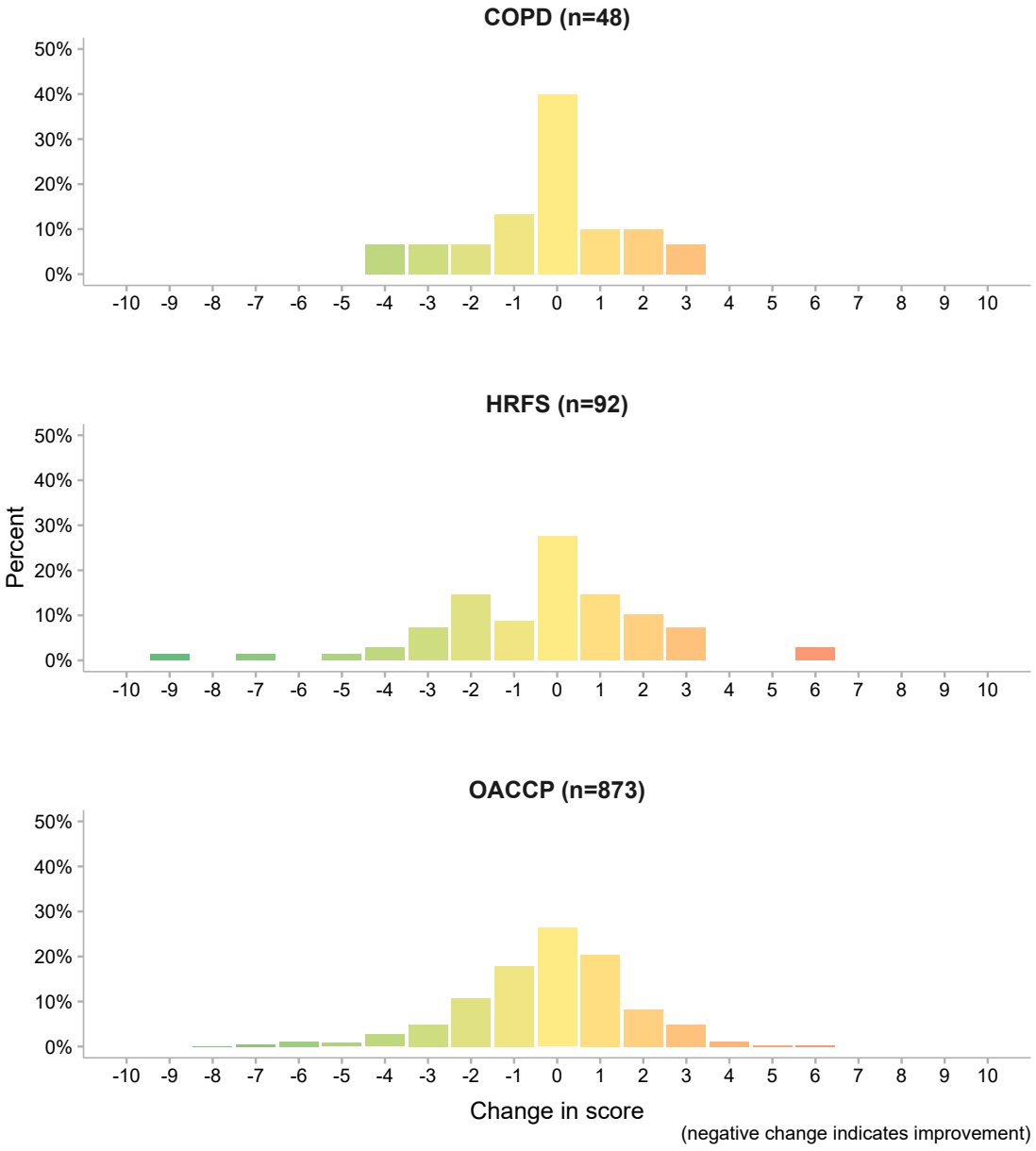
For PROMIS-29, a change of half a standard deviation (or five points as measured on the T score distribution) can be used as a threshold to signify MID.[22, 43, 44] The percentage of patients who improved by at least five points using the T scores is shown for all domains in Figure 28. For COPD, between 25% and 38% of patients passed the threshold for MID on sleep disturbance, anxiety, social roles and depression domains. For HRFS, at least 25% of patients improved by five points or more on five domains. For OACCP, over 20% of patients improved by five points or more on all domains except physical function.

Figure 26: Change in PROMIS-29 scores from first to last survey, first survey completed in 2021 and last survey completed within eight months



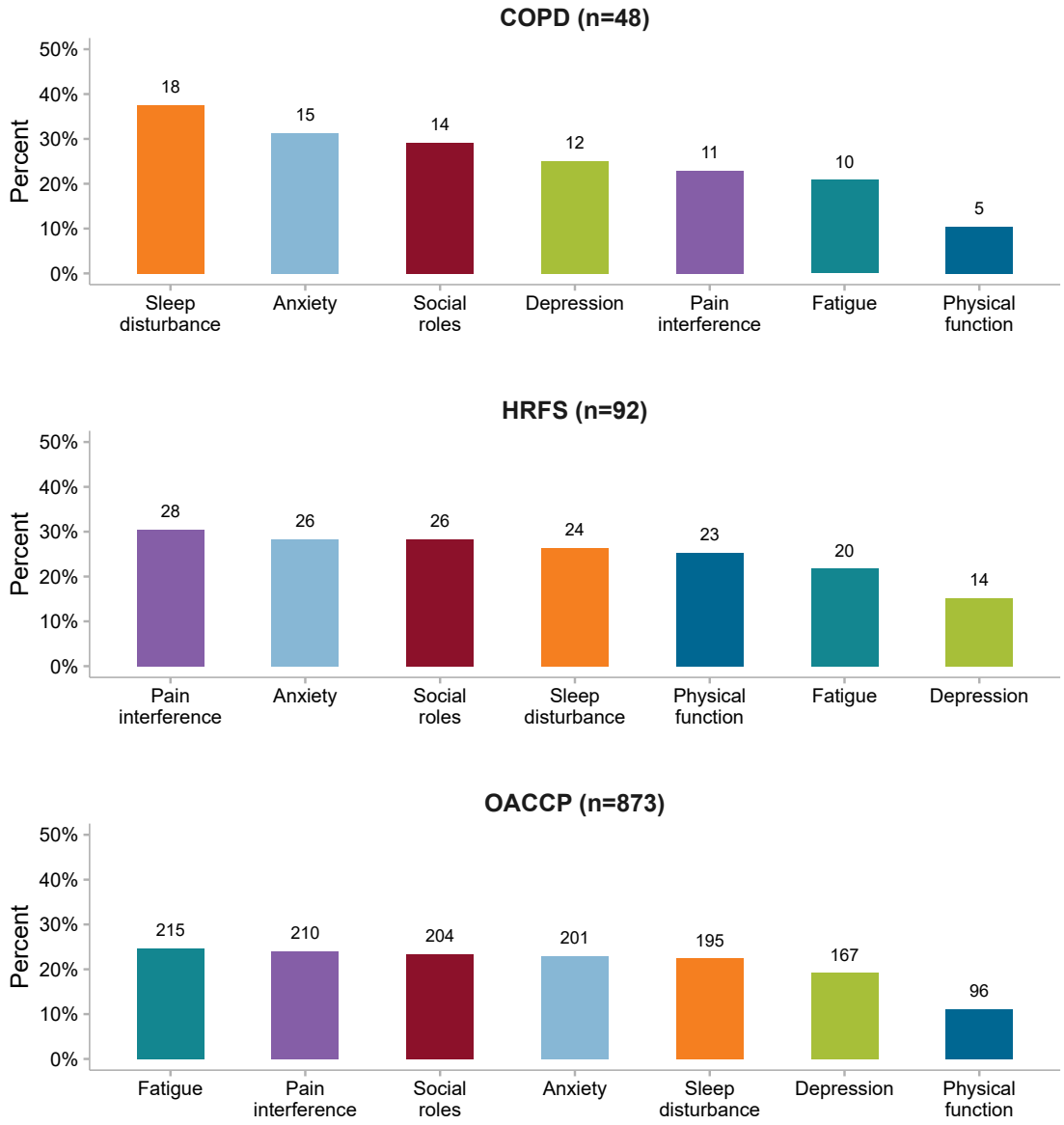
Notes: CHF, IMDM, ORP results are not presented due to insufficient numbers (less than 30 eligible patients completed first and last surveys). Numbers in the figure for some domains may not add up to the first and last survey sample size due to missing data.

Figure 27: Change in PROMIS-29 pain intensity scores from first to last survey, first survey completed in 2021 and last survey completed within eight months



Note: Numbers in the figure may not add up to the first and last survey sample size due to missing data on the pain intensity item.

Figure 28: Proportion of patients with minimally important difference on PROMIS-29 domain T scores (≥ 5 point improvement) from first to last survey; first survey completed in 2021 and last survey completed within eight months



OACCP HOOS and KOOS survey results

A one-week minimum follow-up time was set for the HOOS and KOOS as these surveys can be administered weekly (no surveys removed). For OACCP, scores on all dimensions improved on average from the first to last survey for both the HOOS and KOOS (Figures 29 and 30). The MID has not been determined for the HOOS and KOOS. Guidelines suggest that MID for the HOOS and KOOS should be calculated per clinical cohort and consider various factors such as type of intervention performed and time to follow-up.[45]

Figure 29: Change in HOOS profile scores between first and last survey (n=38), first survey completed in 2021 and last survey completed within eight months

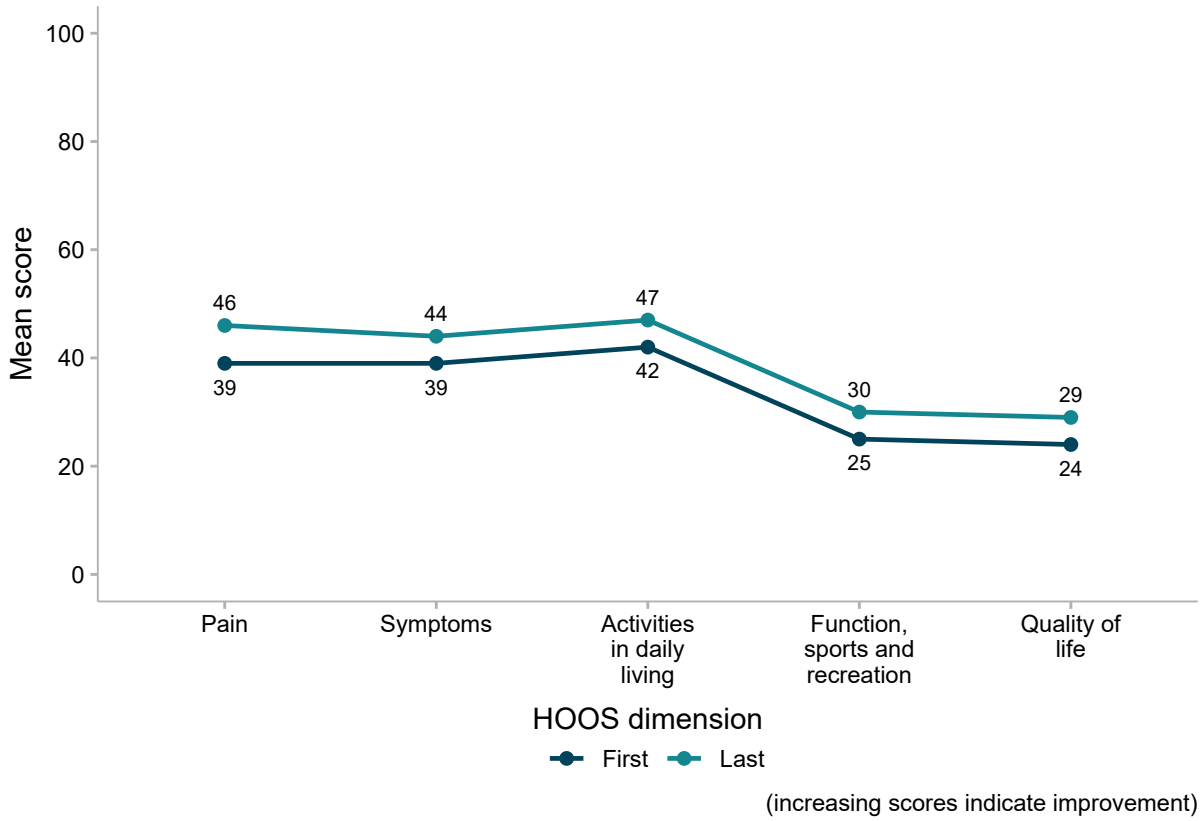
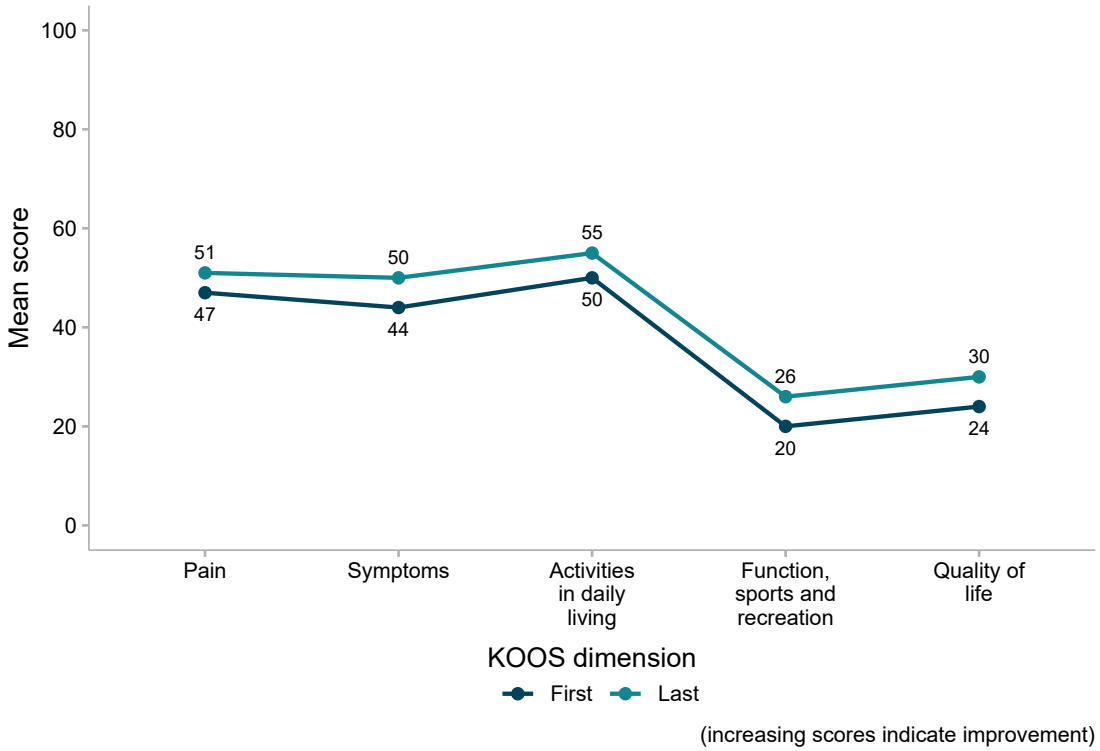


Figure 30: Change in KOOS profile scores between first and last survey (n=121), first survey completed in 2021 and last survey completed within eight months



OACCP Oxford hip score and Oxford knee score

A one-week minimum follow-up time was set for the OHS and OKS as these surveys can be administered weekly (no surveys removed). There were no floor or ceiling effects present for the OHS or OKS total scores – no patients retained the best or worst possible score from first to last survey. Most patients (≥ 65%) experienced a change of between ± 6 points on both the OKS and OHS (Figure 31 and 32). The median total OKS scores increased from the first to last survey (left knee: 22 to 24; right knee: 22 to 25) indicating improvement. Minimal changes over time were observed for left (first: 21; last: 20) and right hips (first: 19.5; last: 20) on the OHS.

Figure 31: Change in OHS total scores between first and last survey (n=390), first survey completed in 2021 and last survey completed within eight months

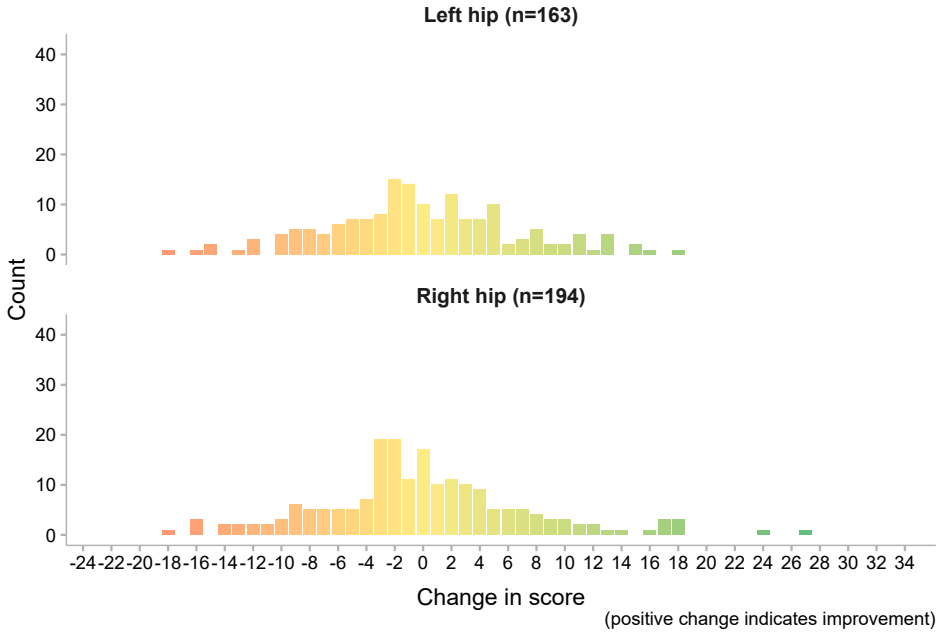
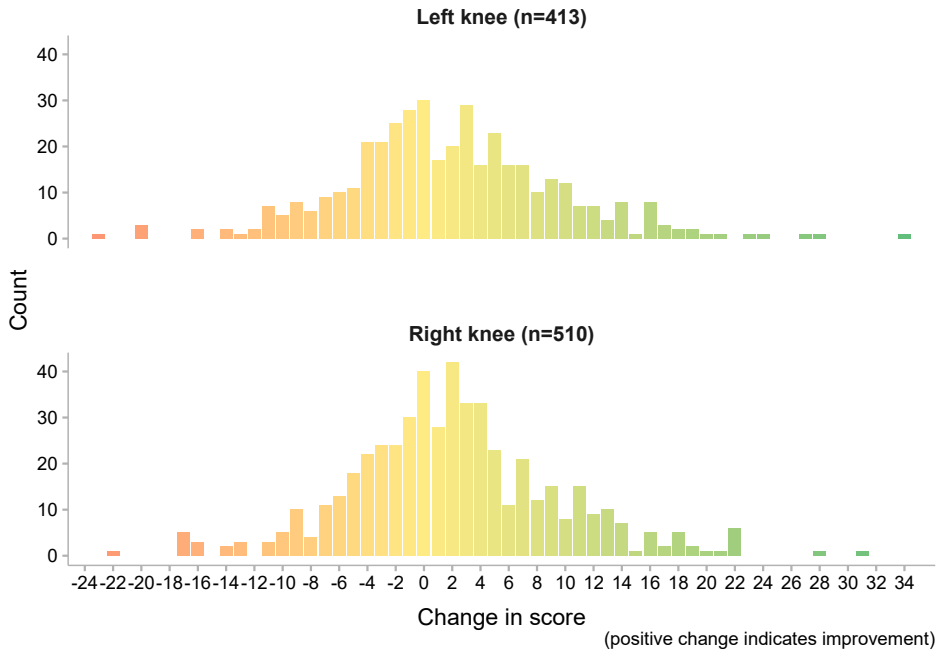


Figure 32: Change in OKS total scores between first and last survey (n=947), first survey completed in 2021 and last survey completed within eight months

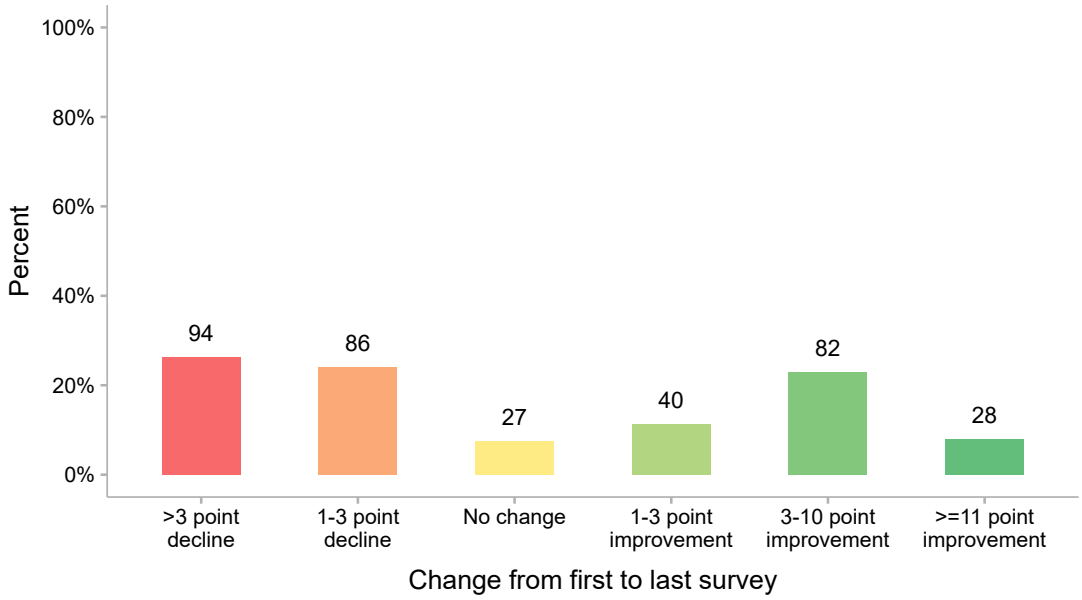


Note. Numbers in the figures may not add up to the first and last survey sample size due to missing data.

OACCP Oxford hip score and Oxford knee score minimally important difference

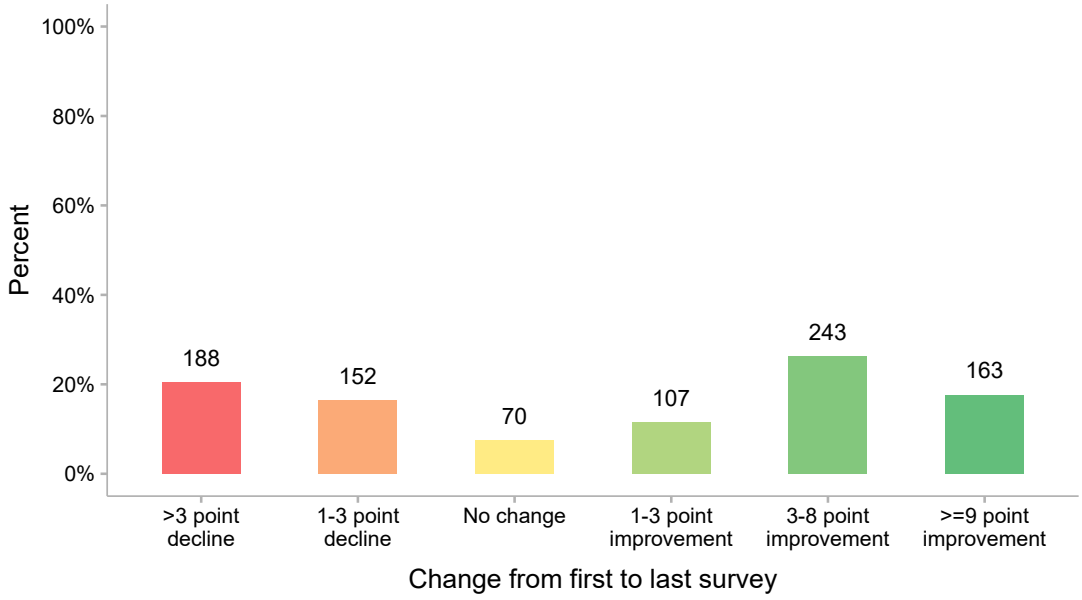
The NHS England has reported an MID at the group level of 11 points for the OHS and nine points for the OKS.[24] Other researchers have suggested an MID of 3-5 points.[46] For the OHS, 110 (30.8%) patients achieved clinically relevant three-point change and 28 (7.8%) patients achieved clinically meaningful change at the 11-point threshold (Figure 33). About 50% of patients declined and 7.6% experienced no change. Of those that declined, just under 50% had less than or equal to a three-point decrease, which may not reflect real change. For the OKS, 406 (44.0%) patients achieved clinically relevant three-point change and 163 (17.7%) patients achieved clinically meaningful change at the nine-point threshold (Figure 34). About 37% of patients declined and 7.6% experienced no change. Of those that declined, just under 45% had less than or equal to a three-point decrease.

Figure 33: Minimally important difference in OHS total scores from first to last survey (n=390), first survey completed in 2021 and last survey completed within eight months



Note. Numbers in the figure may not add up to the first and last survey sample size due to missing data.

Figure 34: Minimally important difference in OKS total scores from first to last survey (n=947), first survey completed in 2021 and last survey completed within eight months

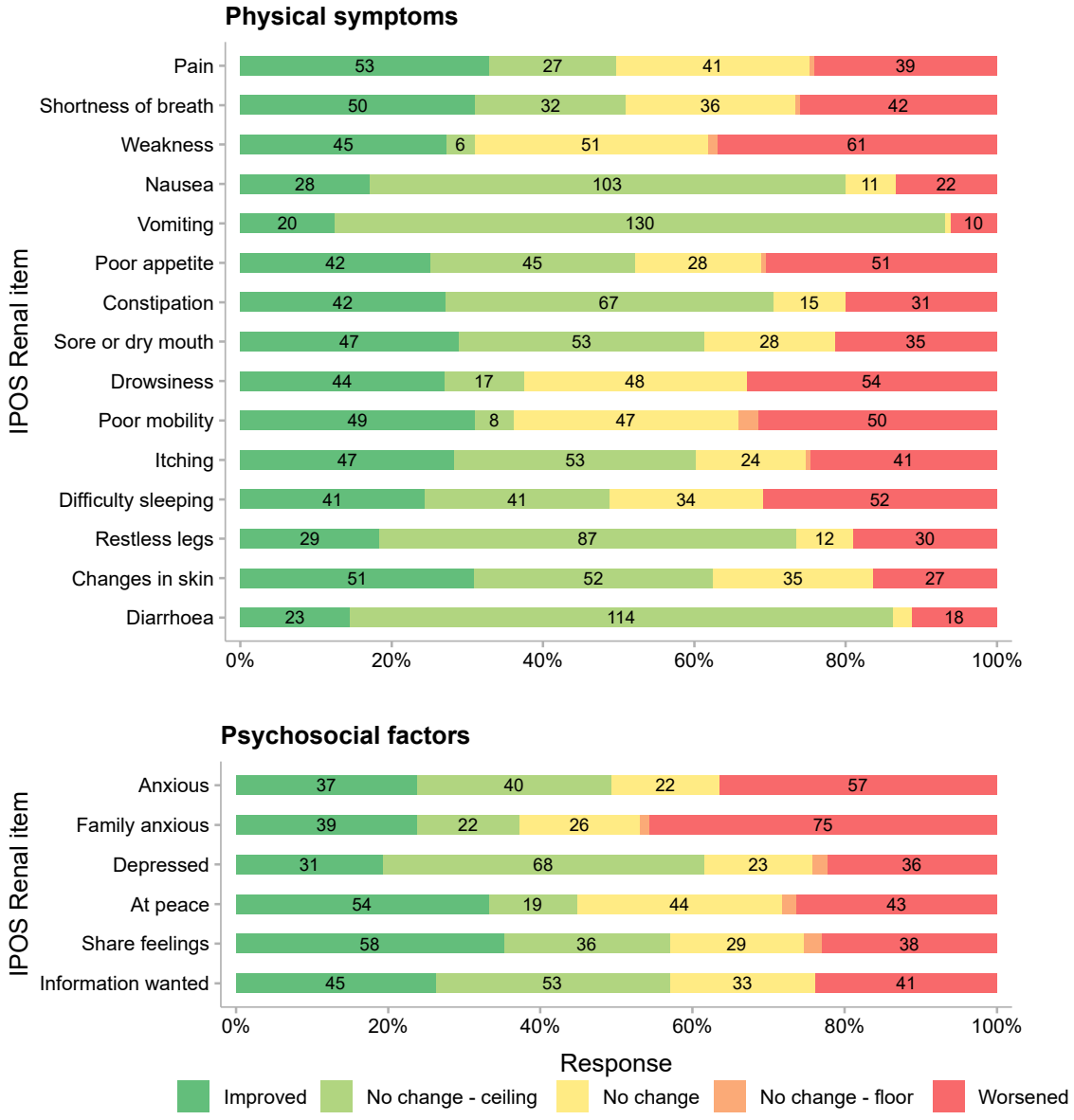


Note. Numbers in the figure may not add up to the first and last survey sample size due to missing data.

RSC IPOS Renal

A one-week minimum follow-up time was set for the IPOS-Renal as the survey can be administered weekly (no surveys removed). At the item level, the percentage of patients improving (by at least one point) ranged from 12% (vomiting) to 35% (sharing feelings) while the percentage of patients who worsened ranged from 6% (vomiting) to 46% (family anxiety). There was a subset of patients who had the best possible score on both the first and last survey (range 4-81% per item; Figure 35) with the highest percentages on vomiting (81%), diarrhoea (72%), nausea (63%) and restless legs (55%). Very few patients retained the worst possible score at the item level from first to last survey (range 1-3% per item). Outside of floor and ceiling change categories, the percentages of patients experiencing no change on items ranged from 1% (vomiting) to 31% (weakness).

Figure 35: Change in IPOS Renal item scores between first and last survey (n=184), first survey completed in 2021 and last survey completed within eight months follow up

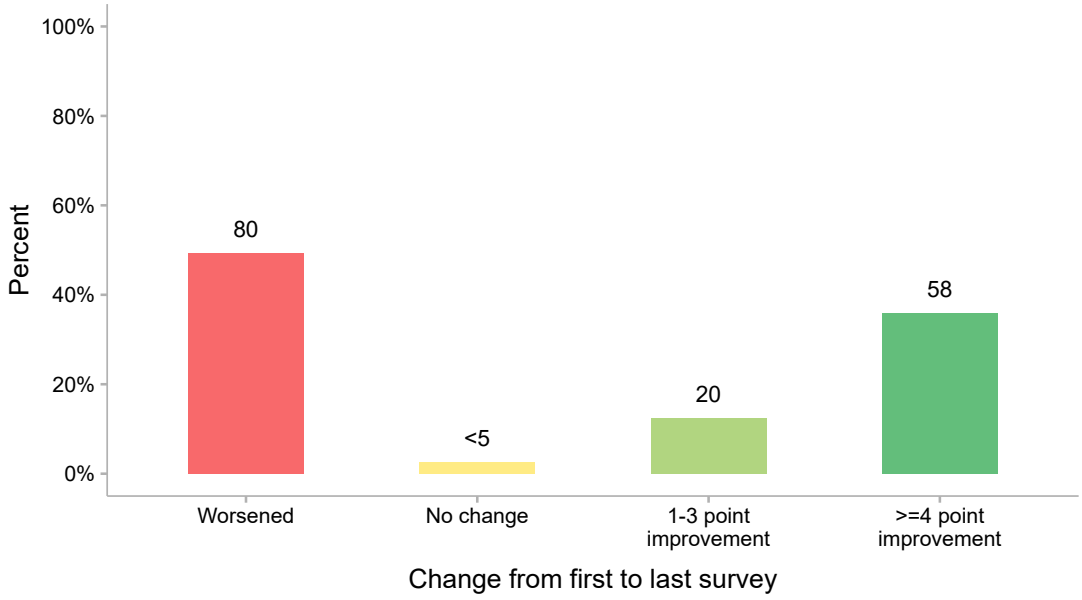


Note. Numbers in the figure may not add up to the first and last survey sample size due to missing data.

RSC IPOS Renal minimally important difference

MID has not been determined for the IPOS Renal. Guidelines for the IPOS, from which the IPOS Renal was derived, can be used as estimates in the interim. At the item level, a one-point change in scores can be interpreted as clinically meaningful.[47] For the overall outcome score, a four-point change is considered clinically meaningful.[48] Using a change of four points to signify MID, 35.8% of patients achieved clinically relevant change on the overall outcome score (Figure 36). There were no floor or ceiling effects on the overall outcome score (no patients maintained the best or worst possible score at the first and last survey).

Figure 36: Minimally important difference in IPOS Renal overall outcome score between first and last survey (n=184), first survey completed in 2021 and last survey completed within eight months follow up



Note. Numbers in the figure may not add up to the first and last survey sample size due to missing data.

Implications and next steps

This report addresses some of the analysis issues raised in the ACI analytic principles for patient-reported outcome measures report using LBVC tranche one cohorts.[1] When more data is available, we will need to reassess analysis methods for appropriateness and draw further on clinical expertise to guide reporting principles.

The next step for PROMs analysis is the development of robust methods to detect change over time and measure outcomes effectively.

- To measure outcomes, PROMs may need to be more closely linked to meaningful clinical program timepoints. This will allow us to attribute any change to clinical treatment or intervention. Future analyses may focus on surveys completed at admission and discharge from care, or those completed pre- and post-clinical procedures.
- For cohorts where PROMs are collected at regular intervals, we can look at measuring change over more than two timepoints. This analysis may require advanced modelling of patient trajectories (fluctuations) and response shift for better assessment of change in health status over multiple timepoints.
- Floor and ceiling effects will need to be accounted for using appropriate modelling techniques to enable detection of change and reduce unwanted bias.
- Future analyses should focus on determining the minimally important difference specific to each cohort and survey at both the individual and group level.
- For the subset of patients completing multiple types of PROMs, investigate concordance between survey results, and between the PROMIS-29 and condition specific surveys.
- Investigate the association between PROMs participation and other health outcomes, such as:
  - unplanned emergency department presentations
  - hospital admissions
  - hospital length of stay
  - hospital readmission
  - integration of care and referrals.
- Investigate factors associated with an improvement in PROMs results, e.g. treatments and interventions.
- Map PROMs results to quality-adjusted life year and disability-adjusted life year to assess the value of care.

We will need to develop methods for fair comparisons across clinics and districts. This could include case-mix adjustment or stratification with established risk adjustment variables, such as comorbidities, severity of underlying disease, frailty scores and socioeconomic status.

We will continue to work collaboratively with the Ministry of Health, LHDs and SHNs and the clinical working group to develop an analysis and reporting methodology that ensures appropriate and useful PROMs reporting at the individual, service, and system level.

References

1. NSW Agency for Clinical Innovation (ACI). Analytic principles for patient-reported outcome measures. Sydney, NSW: ACI; 2021 [cited 26 September 2023]. Available from: https://aci.health.nsw.gov.au/__data/assets/pdf_file/0007/633454/Analytic-Principles-for-Patient-Reported-Outcome-Measures.pdf
2. NSW Bureau of Health Information (BHI). Data Matters. Linking data to unlock information: The use of linked data in healthcare performance assessment. Sydney, NSW: BHI; 2015 [cited 26 September 2023]. Available from: https://www.bhi.nsw.gov.au/__data/assets/pdf_file/0003/290712/DataMatters.pdf
3. Centre for Health Record Linkage (CHeReL). Quality Assurance. Sydney, NSW: CHeReL; 2023 [cited 26 September 2023]. Available from: https://www.cherel.org.au/quality-assurance
4. Williams K SJ, Morris D, Grootemaat P, Thompson C. Patient-reported outcome measures: Literature review. Sydney, NSW: Australian Commission on Safety and Quality in Health Care; 2016. Available from: https://www.safetyandquality.gov.au/sites/default/files/migrated/PROMs-Literature-Review-December-2016.pdf
5. Rathod KS, Wragg A. Do patient-reported outcome measures speak for all patient subgroups: is everyone included? EHJ-QCCO. 2018;4(2):79-80. DOI: 10.1093/ehjqcco/qcy004
6. Health and Social Care Information Centre. Patient Reported Outcome Measures (PROMs) in England - A guide to PROMs methodology. London, UK: NHS Digital; 2017 [cited 26 September 2023]. Available from: https://digital.nhs.uk/binaries/content/assets/website-assets/data-and-information/data-tools-and-services/data-services/proms/proms_guide_v12.pdf
7. NSW Bureau of Health Information (BHI). Adult Admitted Patient Survey 2021: Technical Supplement. Sydney, NSW: BHI; 2022 [cited 26 September 2023]. Available from: https://www.bhi.nsw.gov.au/__data/assets/pdf_file/0005/802625/BHI_AAPS_2021_TECH_SUPP.pdf
8. Gomes M, Gutacker N, Bojke C, et al. Addressing missing data in patient‐reported outcome measures (PROMS): Implications for the use of PROMS for comparing provider performance. Health Econ. 2016;25(5):515-28. DOI: 10.1002/hec.3173
9. Bell ML, Fairclough DL. Practical and statistical issues in missing data for longitudinal patient-reported outcomes. Stat Methods Med Res. 2014;23(5):440-59. DOI: 10.1177/0962280213476378
10. DeBell M, Krosnick JA. Computing weights for American national election study survey data. ANES Technical Report Series, no nes012427. Ann Arbor, MI, and Palo Alto, CA: American National Election Studies; 2009 [cited 26 September 2023]. Available from: https://electionstudies.org/wp-content/uploads/2018/04/nes012427.pdf
11. Battaglia MP, Izrael D, Hoaglin DC, et al. Tips and tricks for raking survey data (aka sample balancing). Abt Associates. 2004;1:4740-4.
12. Kalton G, Flores-Cervantes I. Weighting methods. J Off Stat. 2003;19(2):81-97.
13. Battaglia MP, Hoaglin DC, Frankel MR. Practical considerations in raking survey data. Surv Pract. 2009;2(5). DOI: 10.29115/SP-2009-0019
14. Valliant R, Dever JA, Kreuter F. Practical tools for designing and weighting survey samples: Springer; 2013.
15. McHorney CA, Tarlov AR. Individual-patient monitoring in clinical practice: are available health status surveys adequate? Qual Life Res. 1995 Aug;4(4):293-307. DOI: 10.1007/BF01593882
16. Terwee CB, Bot SD, de Boer MR, et al. Quality criteria were proposed for measurement properties of health status questionnaires. J Clin Epidemiol. 2007 Jan;60(1):34-42. DOI: 10.1016/j.jclinepi.2006.03.012
17. Sayers A, Whitehouse MR, Judge A, et al. Analysis of change in patient-reported outcome measures with floor and ceiling effects using the multilevel Tobit model: a simulation study and an example from a National Joint Register using body mass index and the Oxford Hip Score. BMJ Open. 2020 Aug 27;10(8):e033646. DOI: 10.1136/bmjopen-2019-033646
18. Twisk J, Rijmen F. Longitudinal tobit regression: a new approach to analyze outcome variables with floor or ceiling effects. J Clin Epidemiol. 2009 Sep;62(9):953-8. DOI: 10.1016/j.jclinepi.2008.10.003

19. Lim CR, Harris K, Dawson J, et al. Floor and ceiling effects in the OHS: an analysis of the NHS PROMs data set. BMJ Open. 2015 Jul 27;5(7):e007765. DOI: 10.1136/bmjopen-2015-007765
20. Raine R, Fitzpatrick R, Barratt H, et al. Challenges, solutions and future directions in the evaluation of service innovations in health care and public health. HSDR. 2016;4(16). DOI: 10.3310/hsdr04160
21. Burgess R, Bishop A, Lewis M, et al. Models used for case-mix adjustment of patient reported outcome measures (PROMs) in musculoskeletal healthcare: A systematic review of the literature. Physiotherapy. 2019;105(2):137-46. DOI: 10.1016/j.physio.2018.10.002
22. Revicki DA, Cella D, Hays RD, et al. Responsiveness and minimal important differences for patient reported outcomes. Health Qual Life Outcomes. 2006;4(70):1-5. DOI: 10.1186/1477-7525-4-70
23. Schünemann HJ, Akl EA, Guyatt GH. Interpreting the results of patient reported outcome measures in clinical trials: the clinician's perspective. Health Qual Life Outcomes. 2006;4(1):1-8. DOI: 10.1186/1477-7525-4-62
24. Beard DJ, Harris K, Dawson J, et al. Meaningful changes for the Oxford hip and knee scores after joint replacement surgery. J Clin Epidemiol. 2015;68(1):73-9. DOI: 10.1016/j.jclinepi.2014.08.009
25. Al Sayah F, Jin X, Johnson JA. Selection of patient-reported outcome measures (PROMs) for use in health systems. J Patient-Rep Outcomes. 2021;5(Suppl 2):1-5. DOI: 10.1186/s41687-021-00374-2
26. NSW Ministry of Health. NSW Health Guide to Measuring Value. Sydney, NSW: NSW Ministry of Health; 2023.
27. NSW Ministry of Health. Register of Outcomes, Value and Experience Data Dictionary. Sydney: NSW MoH; 2022.
28. Ho A, Purdie C, Tirosh O, et al. Improving the response rate of patient-reported outcome measures in an Australian tertiary metropolitan hospital. Patient Relat Outcome Meas. 2019;10:217-26. DOI: 10.2147/PROM.S162476
29. KÖhn S, Schlumbohm A, Marquardt M, et al. Predicting non-response in patient-reported outcome measures: results from the Swiss quality assurance programme in cardiac inpatient rehabilitation. Int J Qual Health Care. 2022;34(4):1-7. DOI: 10.1093/intqhc/mzac093
30. Swain MJ, Ballard DH. Color indexing. Int J Comput Vision. 1991;7(1):11-32. DOI: 10.1007/BF00130487
31. NSW Bureau of Health Information (BHI). Weighting and statistical analysis emergency department patient survey (April 2013-March 2014) Sydney, NSW: BHI; 2015 [cited 26 September 2023]. Available from: https://bhi.nsw.gov.au/__data/assets/pdf_file/0008/269954/EDS_Technical_supplement_Weighting_and_Statistical_Analysis.pdf
32. Australian Bureau of Statistics (ABS). Basic survey design - analysis. Canberra, ACT: ABS; 2023 [cited 26 September 2023]. Available from: https://www.abs.gov.au/websitedbs/d3310114.nsf/home/Basic+Survey+Design+-+Analysis
33. Rosenbaum PR. Model-based direct adjustment. J Am Stat Assoc. 1987;82(398):387-94. DOI: 10.1080/01621459.1987.10478441
34. Austin PC. An introduction to propensity score methods for reducing the effects of confounding in observational studies. Multivariate Behav Res. 2011;46(3):399-424. DOI: 10.1080/00273171.201 1.568786
35. Harder VS, Stuart EA, Anthony JC. Propensity score techniques and the assessment of measured covariate balance to test causal associations in psychological research. Psychol Methods. 2010;15(3):234-49. DOI: 10.1037/a0019623
36. Dever JA, Rafferty A, Valliant R. Internet surveys: Can statistical adjustments eliminate coverage bias? Surv Res Methods. 2008;2(2):47-62. DOI: 10.18148/srm/2008.v2i2.128
37. Fischer F, Gibbons C, Coste J, et al. Measurement invariance and general population reference values of the PROMIS Profile 29 in the UK, France, and Germany. Qual Life Res. 2018;27(4):999-1014. DOI: 10.1007/s11136-018-1785-8

38. Liu H, Cella D, Gershon R, et al. Representativeness of the patient-reported outcomes measurement information system internet panel. J Clin Epidemiol. 2010;63(11):1169-78. DOI: 10.1016/j.jclinepi.2009.11.021

39. Pilkonis PA, Choi SW, Reise SP, et al. Item banks for measuring emotional distress from the Patient-Reported Outcomes Measurement Information System (PROMIS®): depression, anxiety, and anger. Assessment. 2011;18(3):263-83. DOI: 10.1177/1073191111411667

40. Elsman EB, Roorda LD, Smidt N, et al. Measurement properties of the Dutch PROMIS-29 v2. 1 profile in people with and without chronic conditions. Qual Life Res. 2022;31(12):3447-58. DOI: 10.1007/s11136-022-03171-6

41. Mercer A, Lau A, Kennedy C. For weighting online opt-in samples, what matters most? USA: Pew Research Center Methods; 2018 [cited 26 September 2023]. Available from: https://www.pewresearch.org/methods/2018/01/26/for-weighting-online-opt-in-samples-what-matters-most/

42. Zini MLL, Banfi G. A narrative literature review of bias in collecting patient reported outcomes measures (PROMs). Int J Environ Res Public Health. 2021;18(23):12445. DOI: 10.3390/ijerph182312445

43. Homco J, Rodriguez K, Bardach DR, et al. Variation and change over time in PROMIS-29 survey results among primary care patients with type 2 diabetes. J Patient Cent Res Rev. 2019;6(2):135-47. DOI: 10.17294/2330-0698.1694

44. Norman GR, Sloan JA, Wyrwich KWJMc. Interpretation of changes in health-related quality of life: the remarkable universality of half a standard deviation. Med Care. 2003;41(5):582-92. DOI: 10.1097/01.MLR.0000062554.74615.4C

45. Roos EM, Lohmander LS. The Knee injury and Osteoarthritis Outcome Score (KOOS): from joint injury to osteoarthritis. Health Qual Life Outcomes. 2003;1(64):1-8. DOI: https://doi.org/10.1186/1477-7525-1-64

46. Murray D, Fitzpatrick R, Rogers K, et al. The use of the Oxford hip and knee scores. J Bone Joint Surg Br. 2007;89(8):1010-4. DOI: 10.1302/0301-620X.89B8.19424

47. Bausewein C, Daveson B, Benalia A, et al. Outcome measurement in palliative care: the essentials London, UK: PRISMA; 2014 [cited 26 September 2023]. Available from: https://kclpure.kcl.ac.uk/portal/en/publications/outcome-measurement-in-palliative-care-the-essentials

48. Scherer JS, Rau ME, Krieger A, et al. A pilot randomized controlled trial of integrated palliative care and nephrology care. Palliat Med. 2022;3(10):1720-9. DOI: 10.34067/KID.0000352022

49. HealthMeasures. PROMIS Adult Profile Instruments. California: Northwestern University; 2023 [cited 26 September 2023]. Available from: https://www.healthmeasures.net/score-and-interpret/interpret-scores/promis

Appendix 1: LBVC NAP clinics

Table A1.1: LBVC NAP clinics

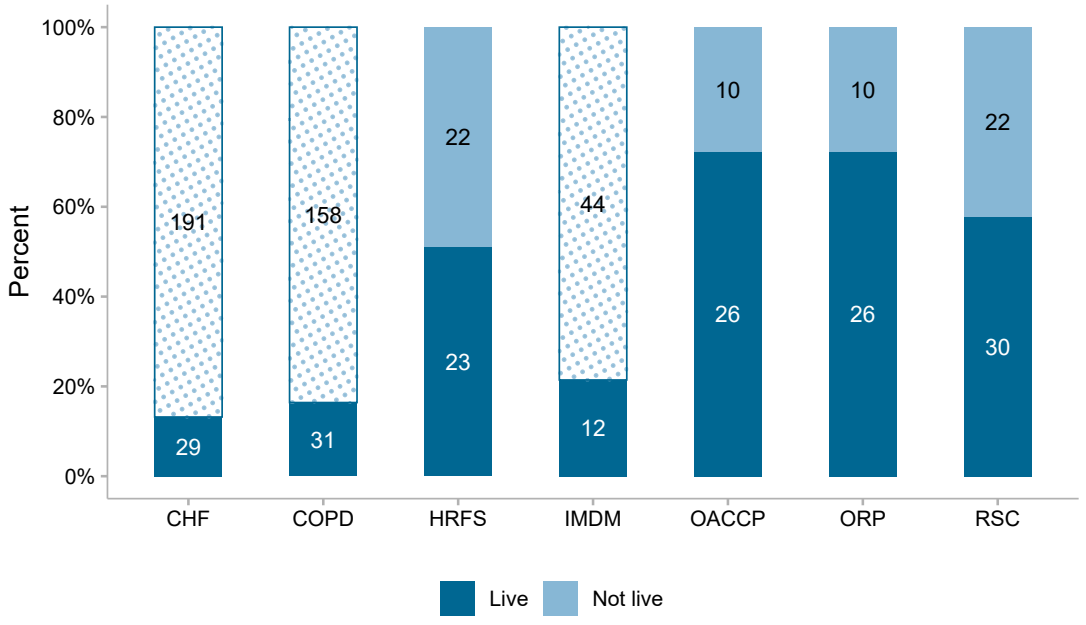
LBVC cohort	LBVC NAP clinics (establishment type)
CHF	16.01 Cardiology Medical Consultation Unit
	16.02 Cardiac Rehabilitation Allied Health / Nursing Unit
	16.11 Circulatory Allied Health / Nursing Unit
COPD	36.01 Respiratory Medical Consultation Unit (NHDD Code 20.19)
	36.05 Respiratory Pulmonary Rehabilitation Medical Consultation Unit (NHDD)
	36.11 Chronic Obstructive Pulmonary Disease Medical Consultation Unit
	36.13 Respiratory General Allied Health / Nursing Unit (NHDD Code 40.40)
	36.16 Respiratory Pulmonary Rehabilitation Allied Health / Nursing Unit
	36.22 Chronic Obstructive Pulmonary Disease Allied Health / Nursing Unit
HRFS	12.25 High Risk Foot Service Allied Health / Nursing Unit
	39.30 High Risk Foot Service Medical Consultation Unit
IMDM	19.05 Diabetes Allied Health / Nursing Unit
OACCP	29.09 Osteoarthritis Chronic Care Program Medical Consultation Unit
	29.10 Osteoarthritis Chronic Care Program Allied Health / Nursing Unit
ORP	29.11 Osteoporosis Refracture Prevention Program Medical Consultation Unit
	29.12 Osteoporosis Refracture Prevention Program Allied Health / Nursing Unit
RSC	34.12 Renal Supportive Care Medical Consultation Unit
	34.13 Renal Supportive Care Allied Health / Nursing Unit

HRFS, OACCP, ORP, and RSC have specific clinics with dedicated establishment types.

For CHF, COPD, and IMDM, we identified relevant clinics based on the establishment types registered in HOPE for these conditions. We validated these establishment types by analysing the clinics most frequently visited by CHF, COPD, and IMDM patients up to 90 days following discharge from an admitted patient episode between 2017-18 (the start of LBVC tranche one) and 2020-21 (latest available data with up to 90 days follow up). We used the LBVC admitted patient cohort definitions to identify CHF, COPD, and IMDM patients (ROVE Data Dictionary[1]). The establishment types registered in HOPE are among the most common clinics visited by these patients and are clinically related.

The HOPE rollout in LBCV NAP clinics started in February 2021. By December 2021, 177 out of an estimated 634 LBVC NAP clinics (28%) were live in HOPE. Across LBVC cohorts it ranged from 13% to 72% (CHF 13%, COPD 16%, HRFS 51%, IMDM 21%, OACCP 72%, ORP 72%, RSC 58%) (Figure A1.1). The HOPE rollout was higher in the LBVC cohorts with dedicated clinics. This report includes PROMs results for the LBVC NAP clinics that were live in HOPE by the end of 2021.

Figure A1.1: Number and percentage of LBVC NAP clinics live in HOPE at December 2021



Note: CHF, COPD, and IMDM not live clinics are presented as a dotted pattern in the figure to reflect uncertainty in the number of eligible clinics for these initiatives.

Appendix 2: PROMs surveys

Several PROMs surveys are used in LBVC, including both generic (can be applied across different populations) and condition-specific (used to assess outcomes associated with a particular disease). This report includes analysis and results for:

- Patient Reported Outcomes Information System 29 (PROMIS-29) - a generic PROMs survey that is used in most LBVC tranche one cohorts except for RSC.
- Hip dysfunction and Osteoarthritis Outcome Score (HOOS), Oxford Hip Score (OHS), Knee injury and Osteoarthritis Outcome Score (KOOS), and Oxford Knee Score (OKS) – condition specific PROMs for OACCP with relatively high volumes of completion.
- Integrated Palliative Outcome Score (IPOS) Renal – a condition specific PROM for RSC

Patient Reported Outcomes Information System 29 (PROMIS-29)

Survey link: https://www.healthmeasures.net/index.php?option=com_instruments&view=measure&id=849&Itemid=992

PROMIS-29 consists of 28 individual items scored on a 1-5 rating scale. Patients respond based on the last week. These items are grouped into seven domains (four items in each domain and a domain score ranging from 4-20). One additional item assesses pain intensity on a 0-10 numeric rating scale score (0=no pain, 10=worst pain possible).

For the five symptom domains (anxiety, depression, fatigue, pain interference, and sleep disturbance), higher scores mean worse symptoms. For the function domains (physical function, and social roles), higher scores mean better function.

All domain scores are converted to T-scores where a score of 50 represents the average expected score (based on United States general population), a score of 40 is one standard deviation (SD) lower than the mean and a score of 60 is one SD higher than the mean.[49] Domain level T-scores can be used to create severity cut points (normal, mild, moderate, and severe; Table A2.1).

Table A2.1: Score cut points (severity rating groups) for the PROMIS-29

LBVC cohort	SD range	T-Score range	Severity rating group
Symptom domains: anxiety, depression, fatigue, pain interference, sleep disturbance	-	<55	Normal
	0.5-1SD worse than average	55-59	Mild
	1-2 SD worse than average	60-69	Moderate
	Greater than 2 SD worse than average	≥70	Severe
Function domains: ability to participate in social roles and activities, physical function	-	>45	Normal
	0.5-1 SD worse than average	41-45	Mild
	1-2 SD worse than average	31-40	Moderate
	Greater than 2 SD worse than average	≤30	Severe

Hip dysfunction and Osteoarthritis Outcome Score (HOOS)

Survey link: https://orthotoolkit.com/hoos/static/media/HOOS.efb723c2.pdf

The HOOS asks patients with hip disability about their symptoms and problems. It can be used to assess outcomes over both short (minimum one week) and long-term intervals. The HOOS consists of 40 items scored on a 0-4 rating scale. The items are grouped into five subscales: pain (10 items), symptoms and stiffness (5 items), activities of daily living (17 items), function in sports and recreational activities (four items) and quality of life (four items). Standardised scores are computed for each subscale with a range of 0-100 (0 = extreme symptoms, 100 = no symptoms). No total score is computed.

Knee injury and Osteoarthritis Outcome Score (KOOS)

Survey link: https://orthotoolkit.com/koos/static/media/KOOS.5d16a12a.pdf

The KOOS asks patients with knee injury about their symptoms and problems. It can be used to assess outcomes over both short (minimum one week) and long-term intervals. The KOOS consists of 42 items scored on a 0-4 rating scale. The items are grouped into five subscales: pain (nine items), other symptoms (seven items), activities of daily living (17 items), sport and recreation function (five items) and knee-related quality of life (four items). Standardised scores are computed for each subscale with a range of 0-100 (0 = extreme problems, 100 = no problems). No total score is computed.

Oxford Hip Score (OHS)

Survey link: https://www.orthopaedicscore.com/scorepages/oxford_hip_score.html

The OHS asks for patients' perceptions on their hip disability problems. The OHS consists of 12 items scored on a 0-4 rating scale that are summed to produce a single total score (range 0-48; 0 = most severe symptoms, 48 = least symptoms). Patients are asked to indicate whether they are receiving treatment for their left or right hip. If both sides are affected, patients will complete the survey twice. Results have been analysed separately in this report to reflect this.

Oxford Knee Score (OKS)

Survey link: https://www.orthopaedicscore.com/scorepages/oxford_knee_score.html

The OKS asks for patients' perceptions on their knee disability problems. The OKS consists of 12 items scored on a 0-4 rating scale that are summed to produce a single total score (range 0-48; 0 = most severe symptoms, 48 = least symptoms). Patients are asked to indicate whether they are receiving treatment for their left or right knee. If both sides are affected, patients will complete the survey twice. Results have been analysed separately in this report to reflect this.

Integrated Palliative Outcome Score (IPOS) Renal

Survey link: https://pos-pal.org/maix/ipos-renal.php#renglish

IPOS Renal is a biopsychosocial assessment tool combining common symptoms renal patients may experience with other concerns, such as information needs and practical issues. Patients respond to each question based on how they have been affected over the past week. It has 21 items covering physical symptoms and psychosocial factors, each with a five-step rating scale (0-4). A higher score indicates a poorer outcome for the respondent. A summation of these items provides a total physical symptom subscale score, total psychosocial factor subscale score, and overall total score.

PROMs collection method

The LBVC PROMs surveys used in this analysis were collected through the HOPE platform. Implementation of the HOPE platform commenced in February 2021 and LBVC locations (admitted patient settings and NAP clinics) are progressively going live in the platform and collecting PROMs data from LBVC patients.

PROMs collection points

There are recommended collection points for PROMIS-29, HOOS, OHS, KOOS, OKS, and IPOS Renal for LBVC cohorts (Table A2.2). PROMs surveys can be allocated outside of these collection points where clinically indicated.

Table A2.2: LBVC PROMs collection points

LBVC cohort	PROMs surveys	Collection point
CHF	PROMIS-29	• As part of inpatient discharge planning process to identify required support/referrals (for example, to social worker, dietitian etc) • On commencement of outpatient service • Upon completion from outpatient service • Six monthly within a Primary Care setting
COPD	PROMIS-29	• As part of inpatient discharge planning process to identify required support/referrals (for example, to social worker, dietitian etc) • On commencement of outpatient service • Upon completion from outpatient service • Six monthly within a Primary Care setting
HRFS	PROMIS-29	• Upon initial presentation to a service • Upon completion from the service • Six monthly within a Primary Care setting
IMDM	PROMIS-29	• As part of inpatient discharge planning process to identify required support/referrals (for example, to social worker or psychologist) • On commencement of outpatient service • Upon completion from outpatient service • Six monthly within a Primary Care setting
OACCP	PROMIS-29 HOOS/OHS KOOS/OKS	• On commencement of the program • Every three months after commencement of the program • Upon completion from the program • Six monthly within a Primary Care setting
ORP	PROMIS-29	• On commencement of the program • Six monthly • Upon completion from the program • Six monthly within a Primary Care setting
RSC	IPOS Renal	• On commencement of the program • Six monthly

Appendix 3: PROMs completion rates – all surveys

For the LBVC tranche one conditions, there were 24,768 completed HOPE PROMs surveys between January 2021 and August 2022 in ROVE (Table A3.1). Some of these surveys were not included in analysis because they did not link to an LBVC patient who visited an LBVC NAP clinic in 2021 – the cohort definition for the PROMs analysis.

Table A3.1: Number of completed HOPE PROMs surveys in ROVE, January 2021 to August 2022

LBVC cohort	Survey type	Number of surveys
CHF	COPD (CAT)	<5
	KCQ-12	950
	EQ-5D-5L	<5
	PROMIS-29	1,024
COPD	COPD (CAT)	1,292
	KCQ-12	7
	DASS 21	11
	EQ-5D-5L	<5
	PROMIS-29	1,134
	SGRQ	184
HRFS	CWIS	309
	DASS 21	<5
	DDS Scale	<5
	PROMIS-29	1,122
IMDM	DDS Scale	30
	PAID	193
	PROMIS-29	154

LBVC cohort	Survey type	Number of surveys
OACCP	DASS 21	7
	FES-1	<5
	HOOS	274
	KOOS	950
	OHS	2,524
	OKS	5,718
	PROMIS-29	6,212
ORP	FES-1	1,405
	OHS	<5
	OKS	<5
	PROMIS-29	1,104
RSC	DASS 21	<5
	EQ-5D-5L	1,189
	IPOS-Renal	1,819
	PROMIS-29	23

Note. Total survey numbers include surveys completed by the same patient within the same LBVC program and on the same date. These numbers also include multiple surveys completed per person. There are also condition specific surveys completed by patients assigned to LBVC programs not collecting those surveys.

Table A3.2: Number of patients completing surveys, first survey completed in 2021 and eight months follow up for subsequent surveys

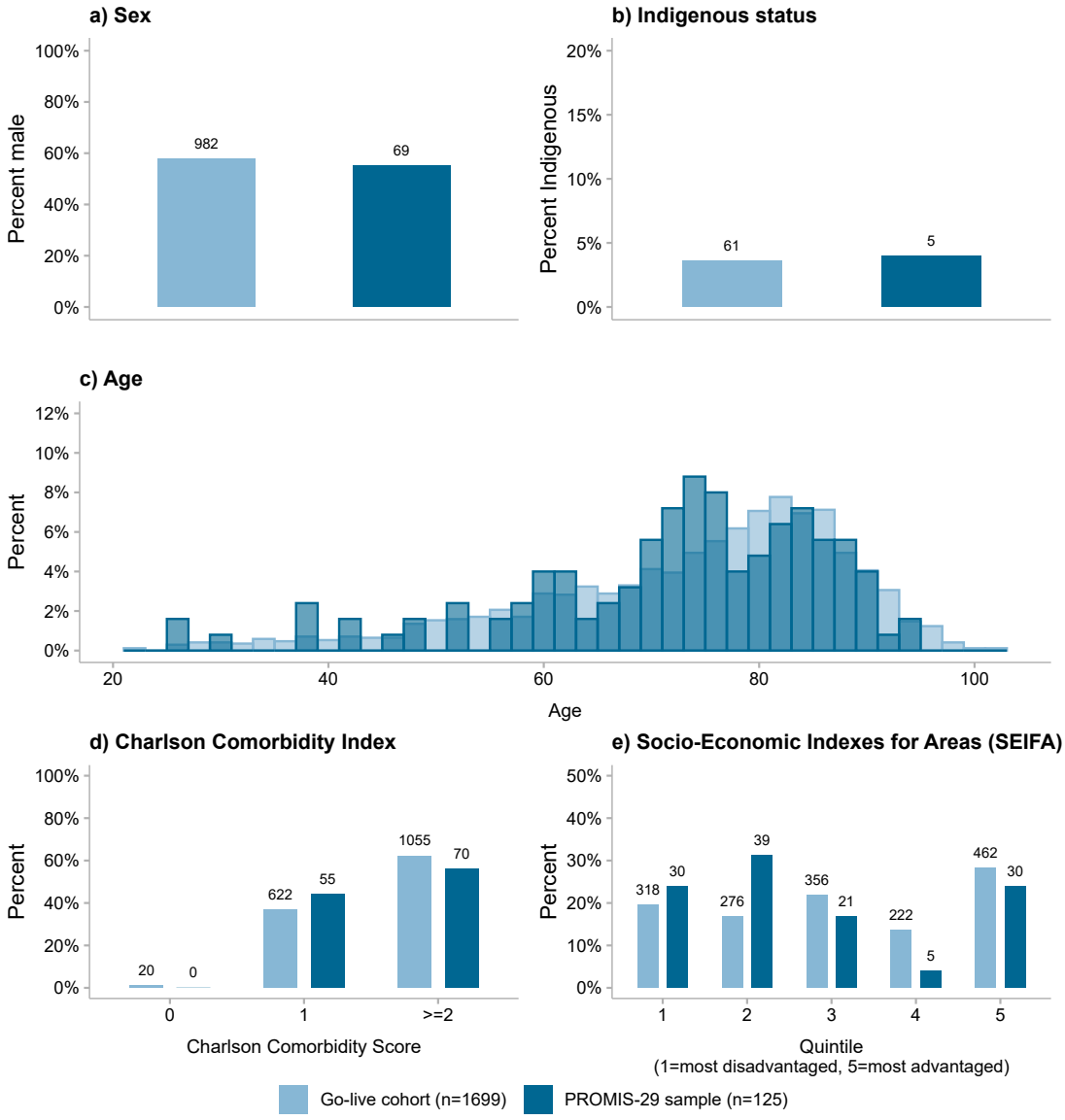
LBVC cohort	Total patients	Total patients after HOPE go-live	Survey type	Zero surveys (n)	Zero surveys (%)	One survey (n)	One survey (%)	Two or more surveys (n)	Two or more surveys (%)
CHF	11032	1699	KCQ-12	1513	89%	160	9%	26	2%
			PROMIS-29	1574	93%	101	6%	24	2%
COPD	9141	1119	COPD (CAT)	964	86%	121	11%	34	4%
			PROMIS-29	956	85%	114	10%	49	5%
HRFS	6792	2476	CWIS	2372	96%	78	3%	26	1%
			PROMIS-29	1917	77%	466	19%	93	5%
IMDM	13465	1966	PAID	1897	96%	69	4%	0	0%
OACCP	9860	5600	HOOS	5527	99%	35	1%	38	1%
			KOOS	5381	96%	98	2%	121	2%
			OHS	4820	86%	390	7%	390	8%
			OKS	3751	67%	902	16%	947	25%
			PROMIS-29	3543	63%	1184	21%	873	25%
ORP	9260	5187	FES-1	4509	87%	649	13%	29	1%
			PROMIS-29	4535	87%	622	12%	30	1%
RSC	2430	1375	EQ-5D-5L	1025	75%	252	18%	98	10%
			IPOS-Renal	894	65%	297	22%	184	21%
			PROMIS-29	1360	99%	15	1%	0	0%

Note. Some surveys are excluded due to small numbers for completion.

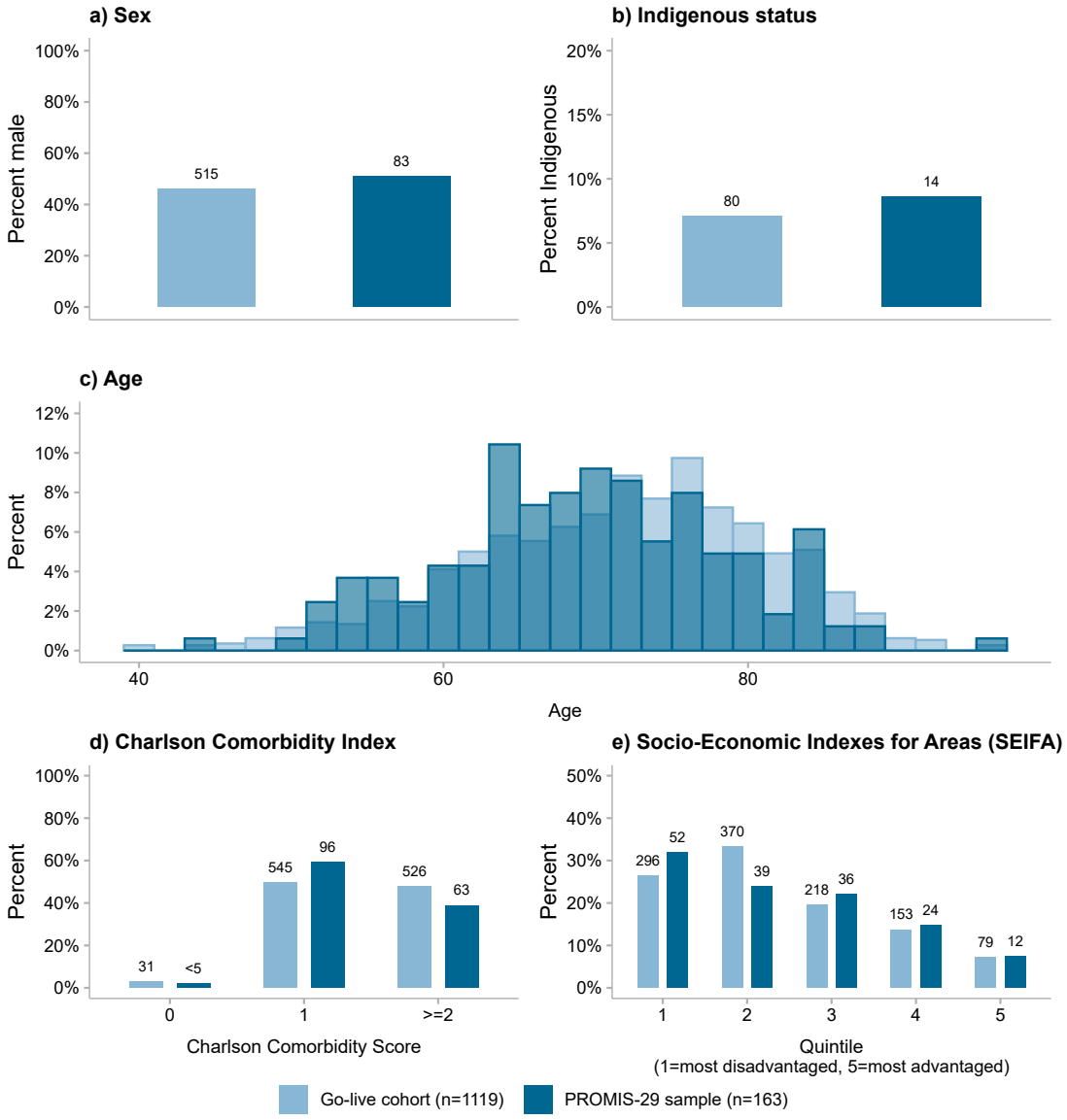
Appendix 4: Cohort and sample characteristics – additional results

Figure A4.1: Representative comparison for CHF, COPD, HRFS, ORP (total patients after HOPE go-live) and sample (at least one PROMIS-29), 2021

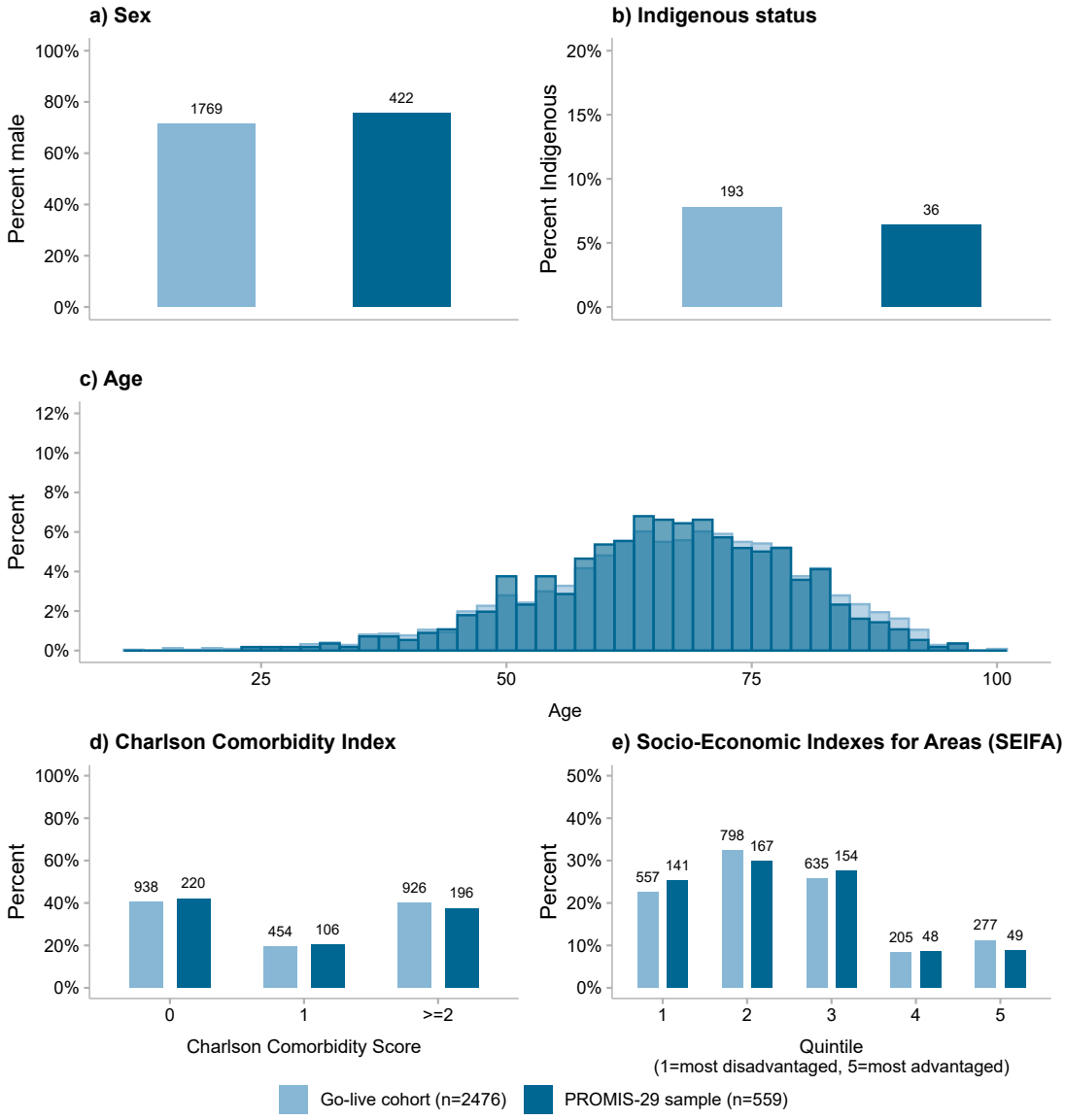
CHF



COPD


a) Sex
Percent male
100%
80%
60%
40%
20%
0%
515
83
b) Indigenous status
Percent Indigenous
20%
15%
10%
5%
0%
80
14
c) Age
Percent
12%
10%
8%
6%
4%
2%
0%
40
60
80
Age
d) Charlson Comorbidity Index
Percent
100%
80%
60%
40%
20%
0%
31
<5
545
96
526
63
0
1
>=2
Charlson Comorbidity Score
e) Socio-Economic Indexes for Areas (SEIFA)
Percent
50%
40%
30%
20%
10%
0%
296
52
370
39
218
36
153
24
79
12
1
2
3
4
5
Quintile
(1=most disadvantaged, 5=most advantaged)
Go-live cohort (n=1119)
PROMIS-29 sample (n=163)

HRFS


a) Sex
Percent male
100%
80%
60%
40%
20%
0%
1769
422
b) Indigenous status
Percent Indigenous
20%
15%
10%
5%
0%
193
36
c) Age
Percent
12%
10%
8%
6%
4%
2%
0%
25
50
75
100
Age
d) Charlson Comorbidity Index
Percent
100%
80%
60%
40%
20%
0%
938
220
454
106
926
196
0
1
>=2
Charlson Comorbidity Score
e) Socio-Economic Indexes for Areas (SEIFA)
Percent
50%
40%
30%
20%
10%
0%
557
141
798
167
635
154
205
48
277
49
1
2
3
4
5
Quintile
(1=most disadvantaged, 5=most advantaged)
Go-live cohort (n=2476)
PROMIS-29 sample (n=559)

ORP

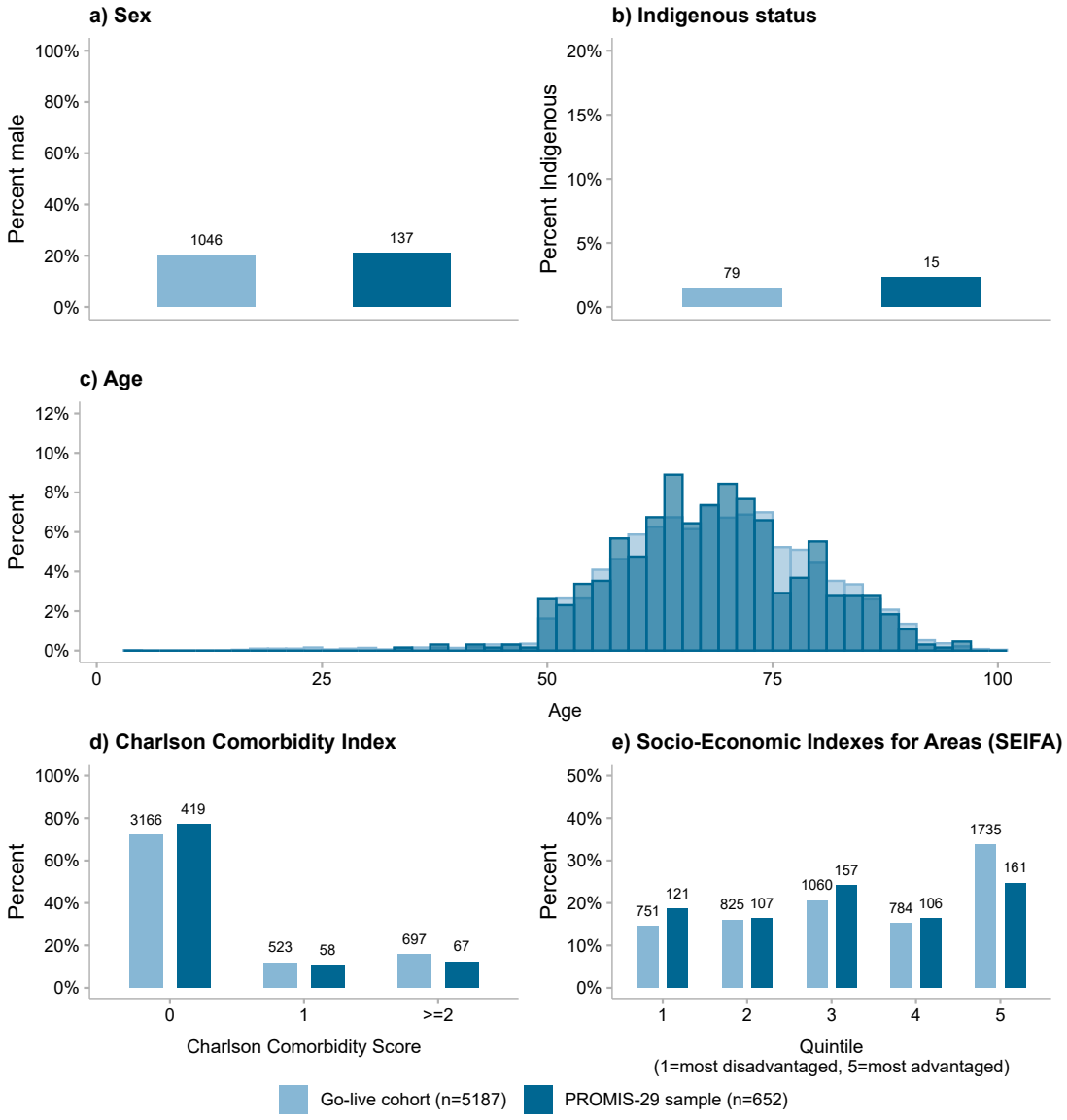
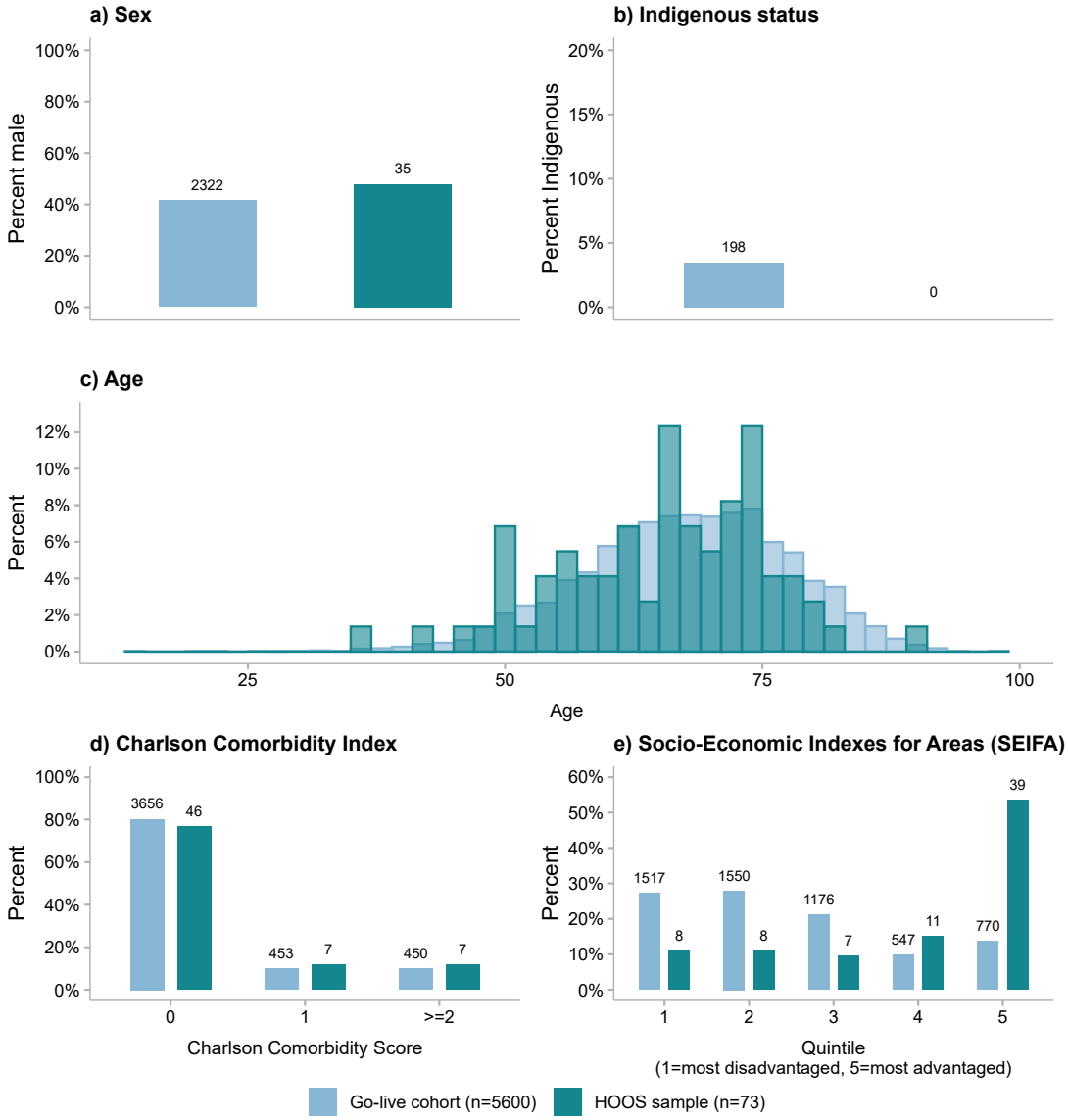
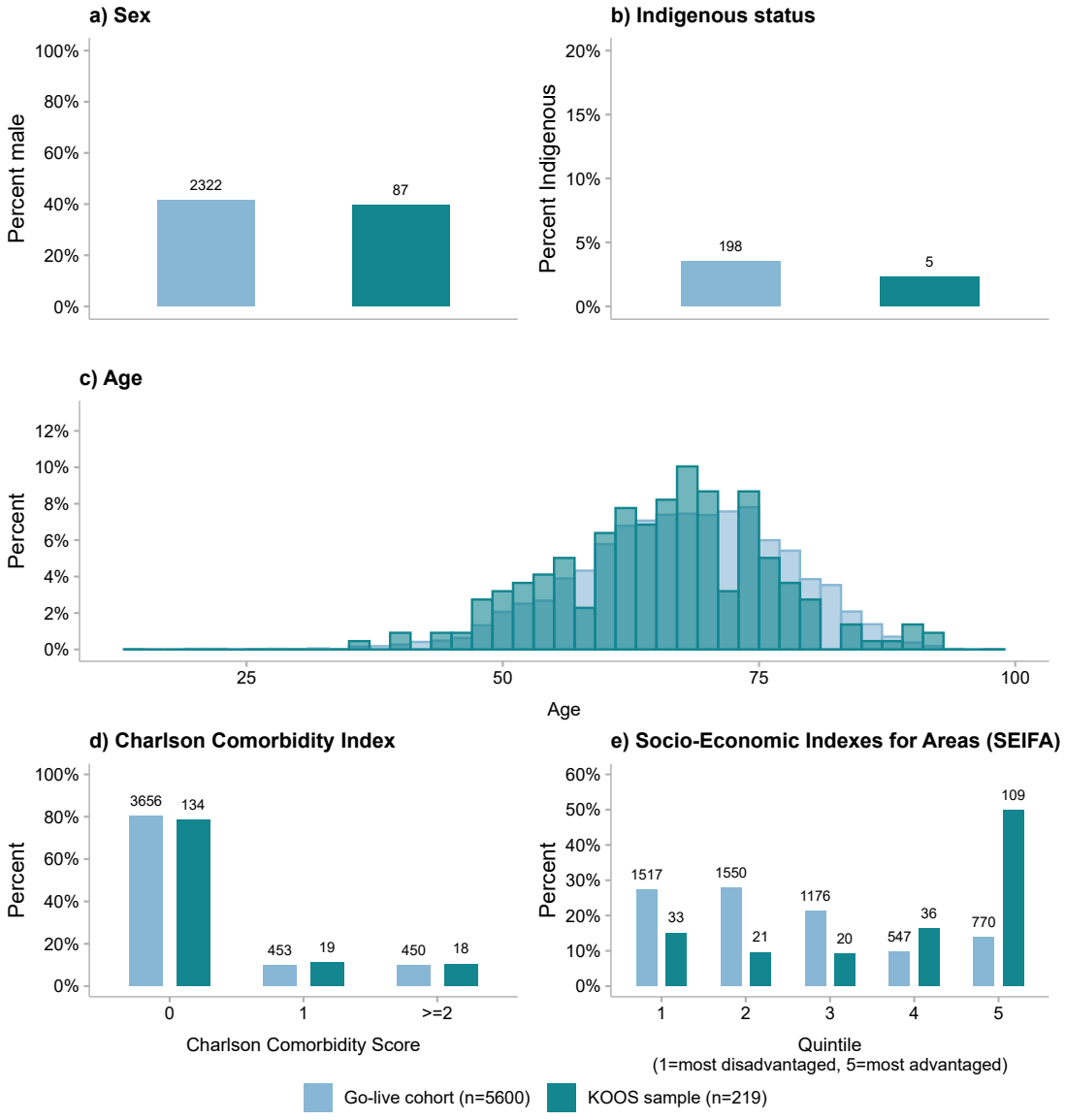

a) Sex
Percent male
100%
80%
60%
40%
20%
0%
1046
137
b) Indigenous status
Percent Indigenous
20%
15%
10%
5%
0%
79
15
c) Age
Percent
12%
10%
8%
6%
4%
2%
0%
0
25
50
75
100
Age
d) Charlson Comorbidity Index
Percent
100%
80%
60%
40%
20%
0%
3166
419
523
58
697
67
0
1
>=2
Charlson Comorbidity Score
e) Socio-Economic Indexes for Areas (SEIFA)
Percent
50%
40%
30%
20%
10%
0%
751
121
825
107
1060
157
784
106
1735
161
1
2
3
4
5
Quintile
(1=most disadvantaged, 5=most advantaged)
Go-live cohort (n=5187)
PROMIS-29 sample (n=652)

Figure A4.2: Representative comparison for OACCP Condition Specific Surveys (total patients after HOPE go-live) and sample (at least one PROM), 2021

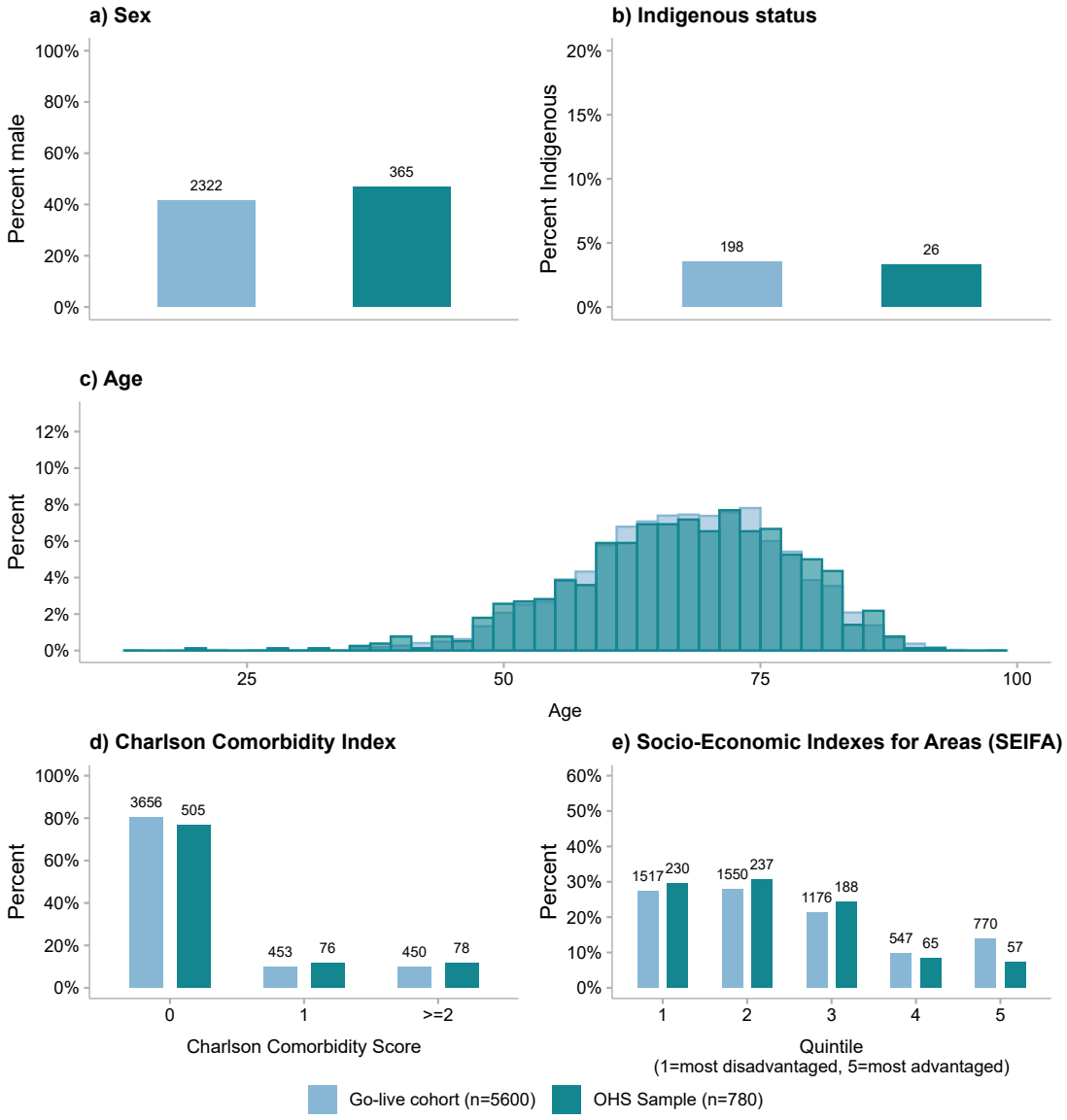
HOOS



KOOS


a) Sex
Percent male
100%
80%
60%
40%
20%
0%
2322
87
b) Indigenous status
Percent Indigenous
20%
15%
10%
5%
0%
198
5
c) Age
Percent
12%
10%
8%
6%
4%
2%
0%
25
50
75
100
Age
d) Charlson Comorbidity Index
Percent
100%
80%
60%
40%
20%
0%
3656
134
453
19
450
18
0
1
>=2
Charlson Comorbidity Score
e) Socio-Economic Indexes for Areas (SEIFA)
Percent
60%
50%
40%
30%
20%
10%
0%
1517
33
1550
21
1176
20
547
36
770
109
1
2
3
4
5
Quintile
(1=most disadvantaged, 5=most advantaged)
Go-live cohort (n=5600)
KOOS sample (n=219)

OHS


a) Sex
Percent male
100%
80%
60%
40%
20%
0%
2322
365
b) Indigenous status
Percent Indigenous
20%
15%
10%
5%
0%
198
26
c) Age
Percent
12%
10%
8%
6%
4%
2%
0%
25
50
75
100
Age
d) Charlson Comorbidity Index
Percent
100%
80%
60%
40%
20%
0%
3656
505
453
76
450
78
0
1
>=2
Charlson Comorbidity Score
e) Socio-Economic Indexes for Areas (SEIFA)
Percent
60%
50%
40%
30%
20%
10%
0%
1517
230
1550
237
1176
188
547
65
770
57
1
2
3
4
5
Quintile
(1=most disadvantaged, 5=most advantaged)
Go-live cohort (n=5600)
OHS Sample (n=780)

OKS

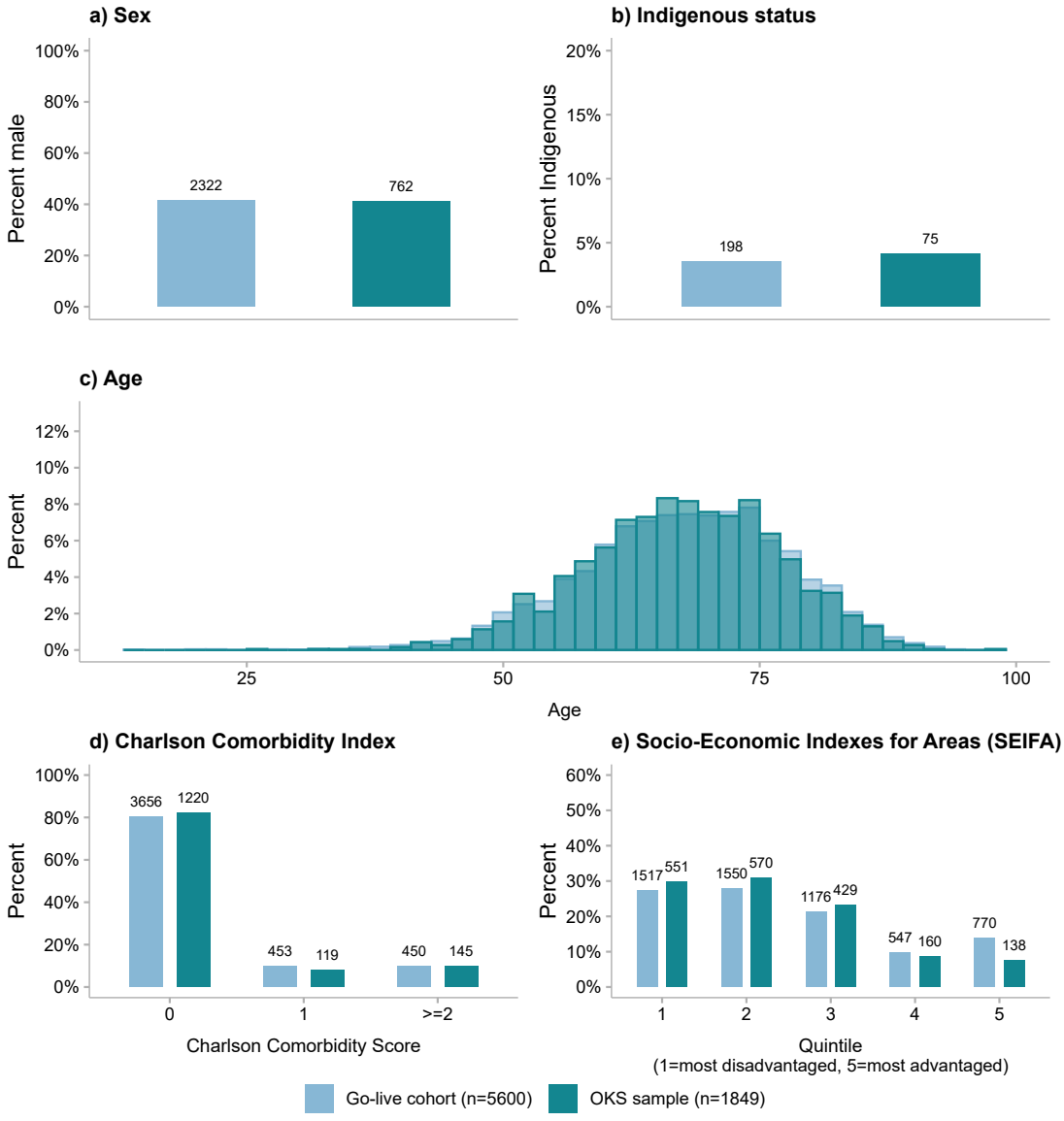
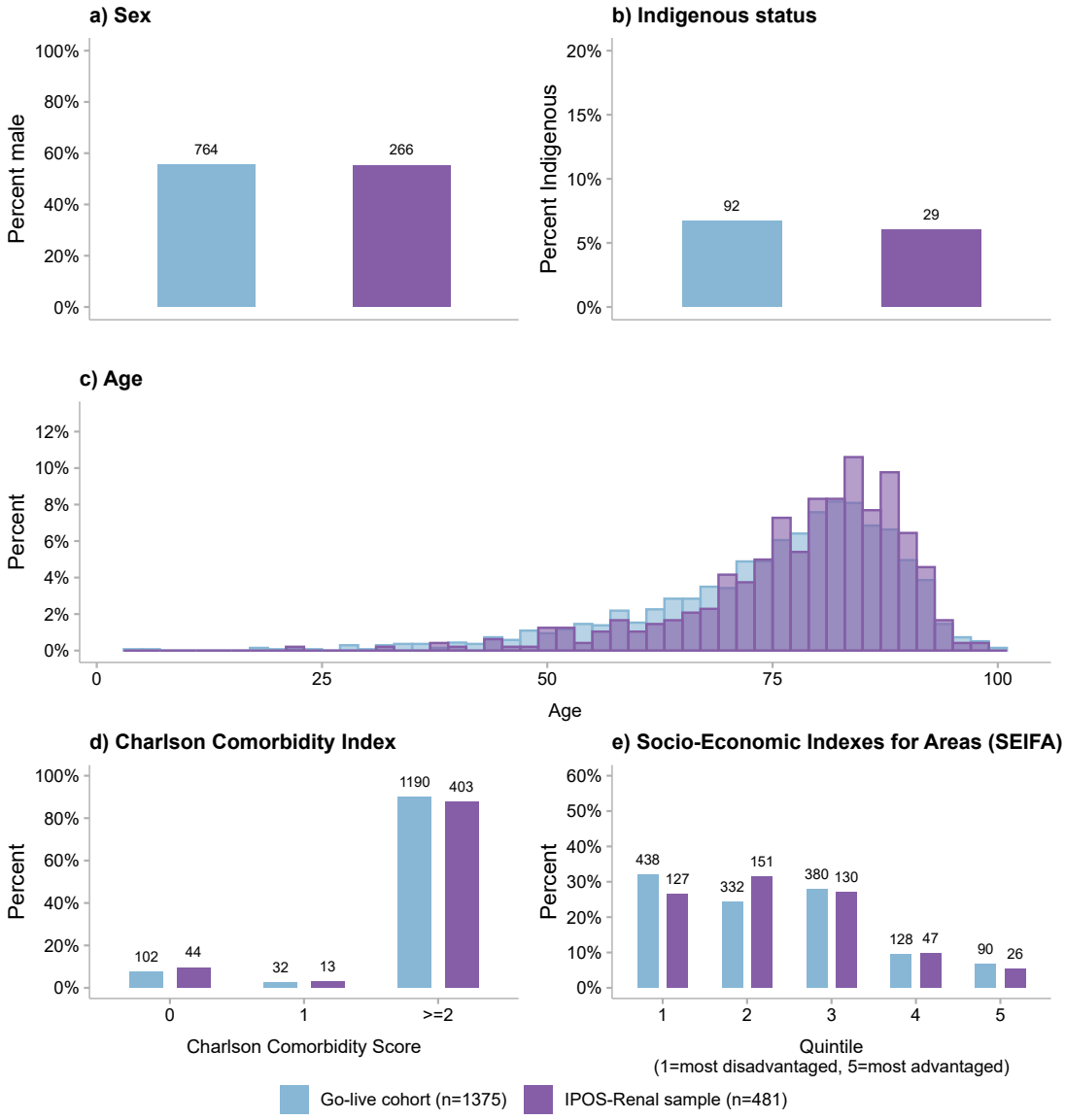

a) Sex
Percent male
100%
80%
60%
40%
20%
0%
2322
762
b) Indigenous status
Percent Indigenous
20%
15%
10%
5%
0%
198
75
c) Age
Percent
12%
10%
8%
6%
4%
2%
0%
25
50
75
100
Age
d) Charlson Comorbidity Index
Percent
100%
80%
60%
40%
20%
0%
3656
1220
453
119
450
145
0
1
>=2
Charlson Comorbidity Score
e) Socio-Economic Indexes for Areas (SEIFA)
Percent
60%
50%
40%
30%
20%
10%
0%
1517
551
1550
570
1176
429
547
160
770
138
1
2
3
4
5
Quintile
(1=most disadvantaged, 5=most advantaged)
Go-live cohort (n=5600)
OKS sample (n=1849)

Figure A4.3: Representative comparison for RSC IPOS-Renal (total patients after HOPE go-live) and sample (at least one PROM), 2021

IPOS-Renal



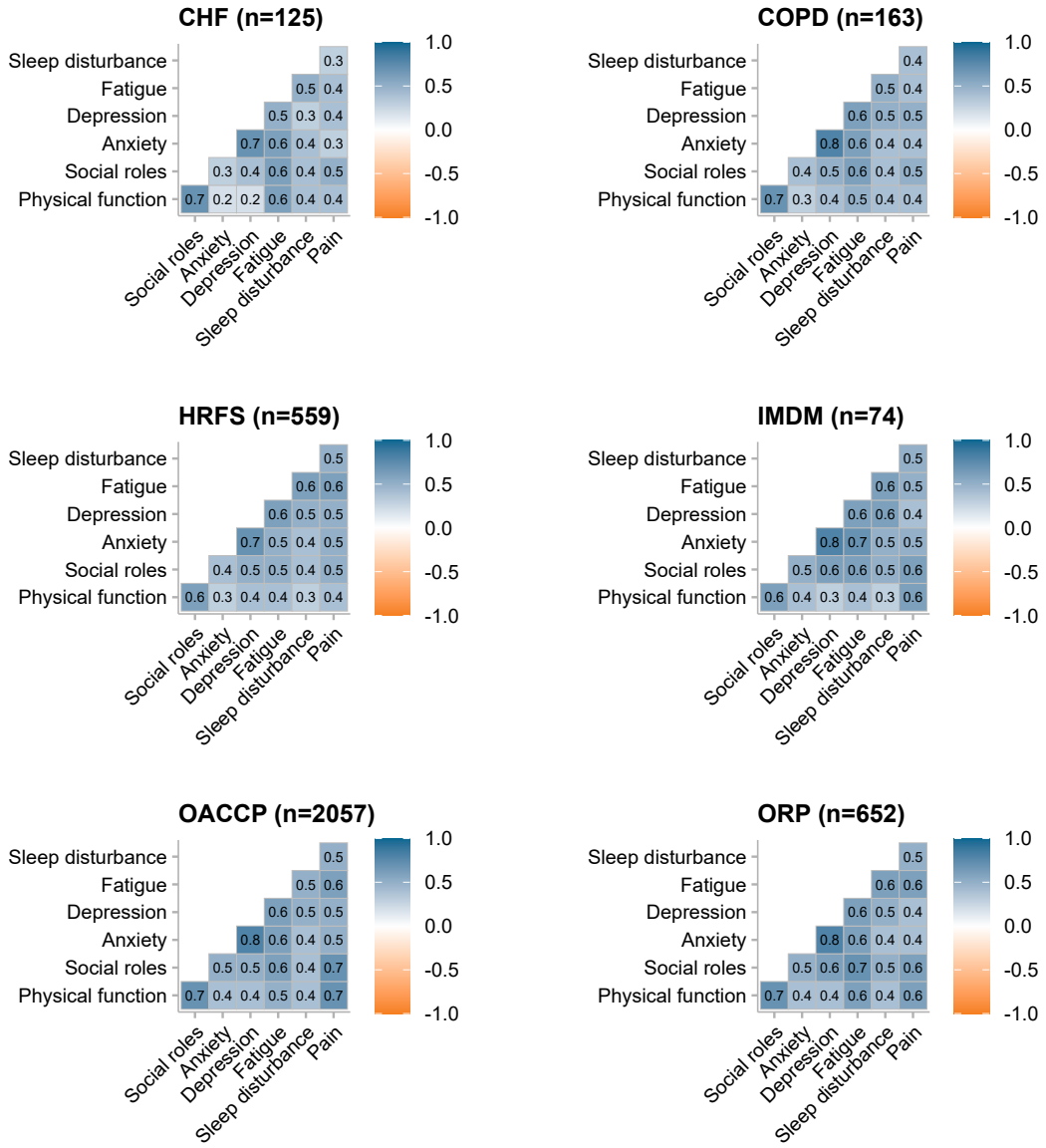
Appendix 5: Cross sectional analysis – additional results

PROMIS-29 domain correlation

There is positive correlation between all domains on the PROMIS-29 survey, meaning the higher a respondent is in one domain, the more likely they are to be higher in another domain (Figure A5.1).

There is strong correlation between responses to anxiety and depression on the PROMIS-29 survey (Figure A5.1).

Figure A5.1: PROMIS-29 correlations between domain T-scores, first survey completed, 2021



Glossary

ACI	Agency for Clinical Innovation
CHeReL	Centre for Health Record Linkage
CHF	Chronic heart failure
COPD	Chronic obstructive pulmonary disease
Future Health	The NSW Health roadmap for the delivery of health services over the coming decade (2022-2032)
HOOS	Hip dysfunction and Osteoarthritis Outcome Score
HOPE	Health Outcomes and Patient Experience - the patient-reported measures information technology platform
HRFS	High risk foot services
IMDM	Inpatient management of diabetes mellitus
IPOS Renal	Integrated Palliative Outcome Score Renal Survey
KOOS	Knee injury and Osteoarthritis Outcome Score
LBVC	Leading Better Value Care - a NSW Health value-based healthcare program
LHD	Local health district
MID	Minimally important difference - the smallest change in a treatment outcome that an individual patient would identify as important and would indicate a change in the patient's management
NAP	Non-admitted patient
OACCP	Osteoarthritis chronic care program
OHS	Oxford Hip Score

OKS	Oxford Knee Score
ORP	Osteoporotic refracture prevention
PREMs	Patient-reported experience measures - asks patients to describe, rather than simply evaluate, what happened during their encounters with health services
PRMs	Patient-reported measures – surveys that help us to understand what matters most to patients and to find out if the care we deliver supports the outcomes and experiences that patients expect
PROMs	Patient-reported outcome measures – capture information about a patient's quality of life or condition-specific measures (e.g. measuring how diabetes is impacting their life). Responses are directly reported, without interpretation by a clinician or anyone else
PROMIS-29	Patient Reported Outcomes Information System 29 – a generic health-related quality of life survey, assesses each of the seven PROMIS domains with four questions. The questions are ranked on a five-point Likert scale. There is also one 11-point rating scale for pain intensity.
ROVE	Register of Outcomes Value and Experience – an enduring NSW Health Public Health Register
RSC	Renal supportive care
SHN	Specialty health network
Value-based healthcare	NSW Health program that aims to improve health outcomes that matter to patients, experiences of receiving and providing care, and effectiveness and efficiency of care

The Agency for Clinical Innovation (ACI) is the lead agency for innovation in clinical care.

We bring consumers, clinicians and healthcare managers together to support the design, assessment and implementation of clinical innovations across the NSW public health system to change the way that care is delivered.

The ACI's clinical networks, institutes and taskforces are chaired by senior clinicians and consumers who have a keen interest and track record in innovative clinical care.

We also work closely with the Ministry of Health and the four other pillars of NSW Health to pilot, scale and spread solutions to healthcare system-wide challenges. We seek to improve the care and outcomes for patients by re-designing and transforming the NSW public health system.

Our innovations are:

- person-centred
- clinically-led
- evidence-based
- value-driven.

aci.health.nsw.gov.au

NSW GOVERNMENT

AGENCY FOR CLINICAL INNOVATION

Our vision is to create the future of healthcare, and healthier futures for the people of NSW.