Canadian Journal of Health Technologies

CADTH

January 2023 Volume 3 Issue 1

CADTH Health Technology Review

Identifying Overused Lab Tests in Hospital Settings: A Delphi Study

Environmental Scan

CADTH

ISSN: 2563-6596

About CADTH: CADTH is an independent, not-for-profit organization responsible for providing Canada's health care decision-makers with objective evidence to help make informed decisions about the optimal use of drugs, medical devices, diagnostics, and procedures in our health care system.

Funding: CADTH receives funding from Canada's federal, provincial, and territorial governments, with the exception of Quebec.

Questions or requests for information about this report can be directed to **Requests@CADTH.ca**

CADTH

Authorship and Panellist Information

Authors

Calvin Young, **MSc** Clinical Research Officer (CADTH)

Francesca Brundisini, MA, PhD Scientific Advisor (CADTH)

Sarah C. McGill, BSc, MLIS Research Information Specialist (CADTH)

Contributors

The authors are grateful to CADTH staff **Kristen Moulton**, MSc; **Michelle Clark**, MSc; and **Chris Kamel**, MSc for their involvement in this project, including providing feedback on the content of this report, and **Irena Milenkovic** for project management. The CADTH team is additionally grateful to **Gillian Hurwitz** of Choosing Wisely Canada for support with identifying panellists and review of the report, and **Ping Mason-Lai**, MA, CMP, PMP for her facilitation of the workshop.

Panellists

These individuals were panellist participants and kindly provided comments on this report:

Adina S. Weinerman, MHSc, MD
Medical Director, Quality & Patient Safety, Sunnybrook Health Sciences Centre
General Internist, Sunnybrook Health Sciences Centre
Toronto, Ontario

Amanda VanSpronsen, MLT, PhD
Associate Professor, University of Alberta
Edmonton, Alberta

Andre Mattman, MD, FRCPC
Division Head Chemistry, Providence Health Care
Department of Pathology and Laboratory Medicine, University of British Columbia
Vancouver, British Columbia

Andrea Thoni, MD, CCFP, FRCP
Medical Director Esoterics, Dalhousie University
Halifax, Nova Scotia

Catherine Cheng, MD, FRCPC
Medical Biochemist, Vancouver Coastal Health
Vancouver, British Columbia

Christian Couture, MD, MSc, FRCPC
Anatomic Pathologist
Institut universitaire de cardiologie et de pneumologie de Québec – Université Laval
Laval, Quebec

Daniel Beriault, MSc, PhD, FCACB
Head of Biochemistry, Unity Health Toronto
Assistant Professor, University of Toronto
Toronto, Ontario

Daniel T. Holmes, MD, FRCPC
Head, Department of Pathology and Laboratory Medicine
Providence Health and the University of British Columbia
Vancouver, British Columbia

Edward Randell, PhD, DCC, FCACB
Clinical Chief and Professor, Department of Laboratory Medicine
Eastern Health Authority and Faculty of Medicine, Memorial University
St. John's, Newfoundland

Fergall Magee, MD, MRCP, FRCPC, MHSc
Provincial Head, Laboratory Medicine
Saskatchewan Health Authority / College of Medicine
Saskatoon, Saskatchewan

Janet Simons, MD, FRCPC
Medical Biochemist, University of British Columbia
Vancouver, British Columbia

Jeannie Callum, BA, MD, FRCPC
Director of Transfusion Medicine, Kingston Health Sciences Centre
Queens's University
Kingston, Ontario

Manal O. Elnenaei, MBChB, PhD, FRCPath
Division Head of Medical Biochemistry
Nova Scotia Health
Halifax, Nova Scotia

Michael Mengel, MD
Chair, Department of Laboratory Medicine and Pathology
University of Alberta
Edmonton, Alberta

Michelle Burleigh
Patient Partner
Georgetown, Ontario

Rosilene Kraft
Patient Partner
Coquitlam, British Columbia

Sam G. Campbell, MB BCH, CCFP(EM), FRCP(Edin)
Professor of Emergency Medicine

Dalhousie University
Halifax, Nova Scotia

Saranya Arnoldo, PhD, FCACB
Clinical Biochemist
Laboratory Medicine and Pathobiology, University of Toronto
Laboratory Medicine and Genetics Program, Trillium Health Partners
Toronto, Ontario

Conflicts of Interest

Three expert panellists declared conflicts of interest as listed below, whereas all other participants had no conflicts of interest to declare that were relevant to this report. One panellist received a grant from Canadian Blood Services, 1 panellist received funds from the Canadian Association of Emergency Physicians as a lecturer, 1 panellist received education funding from Mass Spectrometry & Advances in the Clinical Lab, 1 panellist received education funding from the American Association for Clinical Chemistry, 1 panellist received consulting fees from Waters Corporation, and 1 panellist received funding as part of a research collaboration with SCIEX.

CADTH

Table of Contents

Abbreviations 8

Key Messages 9

Background 9

Objectives 10

Scope, Key Definitions, and Lab Test Eligibility Criteria 10

Study Design 12

Findings 13

Stage 1: The Expert Panel 13

Stage 2: Generating the List of Highly Used Lab Tests 13

Stage 3: The Online Delphi Process 15

List of Hospital-Based Lab Tests Prioritized for Reduction or Elimination of Ordering 19

Discussion and Conclusions 23

Methods 25

Stage 1: Creating the Expert Panel 25

Stage 2: Generating the List of Highly Used Lab Tests 25

Stage 3: The Online Delphi Consensus-Building Process 26

Stakeholder Feedback 28

Limitations 28

References 30

Appendix 1: Additional Tables and Figures 34

CADTH

List of Tables

Table 1: Top 40 Highly Used Lab Tests 13
Table 2: Characteristics of the Subject Experts Involved in the Modified Delphi Process 34
Table 3: Delphi Survey 1 Voting Results 36
Table 4: Delphi Survey 2 Voting Results 37

List of Figures

Figure 1: Study Stages 12
Figure 2: Overview of the Modified Delphi Process 16
Figure 3: Delphi Survey 2 Results 18
Figure 4: Delphi Survey 1 Results 35

Abbreviations

CWC	Choosing Wisely Canada
FT3	free triiodothyronine
FT4	free thyroxine
INR	international normalized ratio
TSH	thyroid stimulating hormone

Key Messages

- Unnecessary lab testing may lead to inaccurate diagnoses, inappropriate treatment or further testing, and increased strain on our health care system, and can expose patients to potential harm.
- To identify lab tests that could be considered by Choosing Wisely Canada's Using Labs Wisely program, we convened a panel of experts that included physicians, medical laboratory professionals, patients, and decision-makers from across Canada. Our objective was to prioritize a list of high-volume laboratory tests that are often ordered unnecessarily in hospital settings.
- Through a consensus-generating process, the expert panel prioritized 7 lab tests as candidates for reduction or elimination in hospital settings. These were 25-hydroxy vitamin D testing, vitamin B12 testing, thyroid stimulating hormone (TSH) testing, free triiodothyronine (FT3) and free thyroxine (FT4) testing, international normalized ratio (INR) testing, and amylase testing.

Background

Laboratory tests provide health care professionals with important information used to make decisions regarding the prevention, diagnosis, treatment, and management of many diseases.[1] They are a critical component of health care services and are the highest volume medical activity in Canada.[2,3] Estimates suggest that each person in Canada receives an average of 14 to 20 laboratory tests per year, and that provincial and territorial governments spend approximately $5.9 billion annually on laboratory activities across health care settings.[3] Despite the importance of laboratory tests for the provision of high-quality health care, it has been reported that about 20% to 30% of lab tests are ordered unnecessarily (e.g., in situations where they may provide little to no benefit for patients, where they are not indicated, or where the test results may not be used to change or inform the course of treatment or management of conditions).[3-5] When used unnecessarily, lab tests can lead to inaccurate diagnoses, unneeded follow-up or treatment, wastage of health care resources (e.g., clinician time, laboratory supplies), and even patient harm (e.g., iatrogenic anemia).[2,5-8] Overuse of lab testing and other health care services also contributes to global pollution, avoidable greenhouse gas emissions, and climate change.[9-13]

Choosing Wisely Canada (CWC) is a national campaign that encourages clinicians and patients to engage in conversations about unnecessary tests, treatments, and procedures and to make sensible and effective care choices.[14] The campaign raises awareness of overused and unnecessary tests, treatments, and procedures by publishing recommendations and other resources produced by national clinician societies that identify specific medical interventions that are not supported by scientific evidence and may unnecessarily expose patients to harm. As of October 2022, more than 80 clinical societies have partnered with CWC to produce more than 400 recommendations.[15,16] Each recommendation is reviewed annually to ensure they are still relevant and reflect the latest evidence.[14] While the recommendations are not intended to establish coverage decisions or be followed in all clinical scenarios, they are meant to spur conversations and to encourage clinicians and patients to consider the potential risks, harms, and benefits before deciding on a treatment plan.[16]

In recognition of the importance of resource stewardship related to lab testing, CWC has established the Using Labs Wisely program, which is a national consortium of hospitals committed to reducing low-value hospital-based lab testing.[17,18] In 2022, CWC's Using Labs Wisely program worked closely with decision-maker funding partners and the lab and clinical community to develop the initial set of 5 recommendations for clinicians to consider when ordering laboratory tests in hospital settings. The 5 recommendations were as follows:

- Reduce aspartate transaminase testing (compared to alanine transaminase testing).
- Reduce partial thromboplastin time testing (compared to prothrombin time and INR testing).
- Eliminate creatine kinase-MB testing.
- Eliminate folate testing.
- Reduce urea testing (compared to creatinine testing).

(Source: Gillian Hurwitz, Choosing Wisely Canada, Toronto, ON: personal communication, Nov 3, 2022.)

In line with the scope of these previous recommendations, the aim of the current project is to prioritize additional hospital-based lab tests that could be considered as candidates for reduction or elimination of ordering in subsequent years of CWC's Using Labs Wisely program.

Objectives

The overall aim of this project was to identify at least 5 overused hospital-based lab tests for consideration for year 2 of the Using Labs Wisely program in Canada. To do so, we:

- identified 40 highly used lab tests in Canada that are potentially overused, low-value, and unnecessary
- prioritized at least 5 overused lab tests using expert consensus.

When developing the new list of unnecessary lab tests, we aimed to include tests that are frequently used, might expose patients to harm, might contribute to stress for patients, and create an increased strain on our health care system.

Scope, Key Definitions, and Lab Test Eligibility Criteria

Definition of Unnecessary, Overused, and Low-Value Lab Tests

For the purposes of this project, we defined overused lab tests as low value, clinically unnecessary lab tests. The following definitions of related concepts, including overused lab tests, unnecessary tests, and low-value care, informed our definition.

- Overused lab tests: "Overuse has been defined as the delivery of tests and procedures that provide little or no clinical benefit; are unlikely to have an effect on clinician decisions; increase health care spending without improving health outcomes; or risk patient harm in excess of potential benefits."[19]

- Unnecessary tests: "These are tests and treatments where strong scientific evidence demonstrates they are not helpful to patients in particular circumstances or may unnecessarily expose patients to harm."[20]
- Low-value care: "There is increasing recognition of the problem of overuse relating to 'low-value care' defined as a test or treatment for which there is no evidence of patient benefit or where there is evidence of more harm than benefit."[21]

Lab Test Eligibility Criteria

Our intention was to include lab tests that are ordered by health care practitioners working in a hospital setting (e.g., hospitalists, internists, emergency physicians), have their sample(s) retrieved from the patient in a hospital setting, and that are processed in a hospital lab by medical laboratory professionals. Lab tests (i.e., sampling and processing) conducted in the community, health care professional settings (in-office), and public health or private labs were not eligible. For the purposes of this project, we defined overused lab tests as low-value, unnecessary lab tests. Unnecessary does not mean a test is never necessary, but rather that there are situations where the test is ordered but not needed. We recognize that clinical judgment is an important component of patient care and emphasize that lab tests described as having low clinical value may still be important tools that clinicians and patients wish to pursue in certain clinical scenarios. The following criteria were used when determining which hospital-based lab tests are unnecessary, low-value, and overused:

- **Lab tests that are frequently used in Canada**. CWC's priority is to focus on frequently overused and unnecessary tests.[14] We limited our scope to frequently used lab tests as they consume a large amount of health care resources (e.g., clinician time, laboratory supplies), and any initiatives to reduce their use have larger potential impact.
- **Lab tests that have low clinical value**. Unnecessary lab tests offer little to no clinical benefits to patients, and may cause harm or risk patient harm in excess of potential benefits. In addition, they may be unlikely to have an effect on clinician decisions.
- **Lab tests that are costly and increase health care spending without improving patient outcomes.**
- **Lab tests that have potential for reduction**. Physicians, medical laboratory specialists, and decision-makers should have the power to reduce the use of these lab tests.

We intended to identify specific lab tests for inclusion in the final list, rather than laboratory and clinical practices related to lab testing. However, when considering lab tests for inclusion or exclusion in the final list of prioritized unnecessary lab tests, the expert panel considered laboratory practices to support deliberation and consensus. Practices may have included, but were not limited to, unnecessary initial testing (i.e., conducting a test during initial patient evaluation when it is not indicated), unnecessary duplicate testing (i.e., where a test is ordered even if there is a valid result on file), or unnecessary repeat testing (i.e., retesting a clinically stable patient before appropriate repeat intervals have elapsed). Practices may have also included decisions that inform test ordering, decisions which can be guided by clear indications for ordering a test, or decisions guided by the fact that there is no contraindication. However, we did not focus on practices, as we did not aim to draft clinical practice guidelines but aimed to indicate which lab tests could be candidates for elimination or reduction in the hospital setting.

We excluded the 5 tests that were previously identified by CWC's Using Labs Wisely program (i.e., aspartate transaminase, partial thromboplastin time, creatine kinase-MB, folate, and

urea) for reduction or elimination of ordering (Gillian Hurwitz, personal communication, Nov 3, 2022).

Two invited subject experts reviewed the project scope and relevant criteria that were used to define unnecessary lab tests to confirm appropriateness.

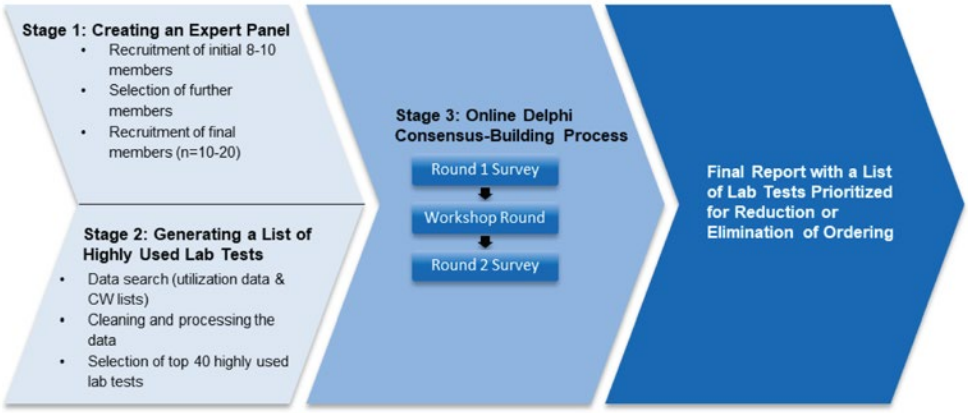
Study Design

We conducted an online modified Delphi process to identify, describe, and prioritize at least 5 hospital-based lab tests in Canada that are overused and could be considered for reduction or elimination of ordering. The Delphi process is a structured group facilitation technique to obtain consensus among a group of experts on a given topic through iterative rounds with feedback.[22,23]

CWC provides broad guidance on how to develop and prioritize the top 5 list by health professional practice societies.[14] The guidance outlines the following essential components: each item on the list should be within the specialty's scope of practice; there should be strong evidence to support each recommendation; and the process should be transparent, systematic, and consultative.[14]

In line with the essential components of CWC's guidance, we conducted a modified online Delphi study in 3 stages, as illustrated in Figure 1. In stage 1, we recruited a panel of participants who are experts in ordering or conducting lab tests, or implementing policies related to their use. Concurrently, in stage 2, we generated a list of 40 lab tests that are highly used in Canada from lab test order volumes provided by various hospitals throughout the country. Stage 3 was the modified online consensus-building process, where we conducted 2 survey rounds and an online workshop discussion to prioritize at least 5 unnecessary lab tests. A detailed description of the study process is provided in the Methods section.

Figure 1: Study Stages



CW = Choosing Wisely.

Findings

Stage 1: The Expert Panel

We recruited a total of 16 subject expert panellists who participated in the modified Delphi consensus process. Subject expert panellists were from diverse geographical settings across Canada (including Alberta [n = 2; 12.5%], British Columbia [n = 4; 25.0%], Newfoundland and Labrador [n = 1; 6.3%], Nova Scotia [n = 3; 18.8%], Ontario [n = 4; 25.0%], Quebec [n = 1; 6.3%], and Saskatchewan [n = 1; 6.3%]); had varying years of experience related to lab tests (6 subject experts [37.5%] had between 0 and 9 years of experience; 10 subject experts [62.5%] had 10 or more years of experience); and represented the professional stakeholder groups we aimed to recruit (i.e., medical laboratory professionals [n = 10; 62.5%], physicians [n = 9; 56.3%], and decision-makers [n = 9; 53.6%]). The expert panel included medical laboratory professionals from each traditional discipline of medical laboratory science (i.e., transfusion [n = 2], clinical chemistry [n = 9], hematology [n = 1], microbiology [n = 1], and histology [n = 2]) and physicians from our target medical specialties (i.e., internal medicine [n = 3], pathology [n = 4], and emergency medicine [n = 1]). Twelve subject experts participated in the round 1 survey, 15 subject experts participated in the online workshop, and all 16 subject experts participated in the round 2 survey. Relevant characteristics of expert panellists who participated in the survey and the workshop rounds are provided in Appendix 1, Table 2.
In addition to the 16 expert panellists who participated in the voting rounds of the Delphi process, 2 patient partners (1 from Ontario, 1 from British Columbia) with significant knowledge and understanding of hospital lab tests took part in the online workshop. They reflected on their experiences and offered their views on lab tests from the perspective of patients and caregivers to loved ones, to help inform the discussions.

Stage 2: Generating the List of Highly Used Lab Tests

We drafted an initial list of top 40 highly used lab tests for consideration by the expert panel. The list was generated using the cumulative test volumes from available data sources across Canada. The list is shown in Table 1.

Table 1: Top 40 Highly Used Lab Tests

Rank by order volume	Test or panel name	Cumulative test volume[a,b]	Normalized test volume[c]
1	Creatinine	658,704	100.0
2	Potassium	638,607	96.9
3	Sodium	638,207	96.9
4	Complete blood count	580,550	88.1
5	COVID-19 NAAT	540,798	82.1
6	Chloride	506,892	77.0
7	Total carbon dioxide	482,428	73.2
8	Random blood glucose	468,517	71.1
9	Phosphate	250,923	38.1

Rank by order volume	Test or panel name	Cumulative test volume[a,b]	Normalized test volume[c]
10	Magnesium	249,630	37.9
11	Alanine transaminase	233,897	35.5
12	Calcium	223,814	34.0
13	Total bilirubin	208,786	31.7
14	Alkaline phosphatase	191,723	29.1
15	INR	183,419	27.8
16	Albumin	163,823	24.9
17	Troponins	138,463	21.0
18	Arterial blood gases	118,130	17.9
19	Gamma-glutamyl transferase	91,797	13.9
20	Hemoglobin A1C	73,129	11.1
21	Urinalysis	72,830	11.1
22	C-reactive protein	65,918	10.0
23	Type and screen (ABO group, Rh group, and antibody screen)	64,009	9.7
24	Lactate	61,050	9.3
25	TSH	57,023	8.7
26	Amylase	53,118	8.1
27	Hepatitis B surface antigen	52,758	8.0
28	Lipid panel	45,993	7.0
29	Ionized calcium	45,484	6.9
30	White blood cell count[c]	44,632	6.8
31	Ferritin	41,867	6.4
32	Urine cultures	40,811	6.2
33	Blood cultures	38,980	5.9
34	Immunoperoxidase staining	36,562	5.6
35	Lipase	35,124	5.3
36	Crossmatch	33,781	5.1
37	Total protein	33,488	5.1
38	Brain natriuretic peptides	28,370	4.3
39	25-hydroxy vitamin D	28,289	4.3
40	HIV viral load	25,134	3.8

INR = international normalized ratio; NAAT = nucleic acid amplification test; TSH = thyroid stimulating hormone.

Note: This utilization data does not provide any information on which tests are unnecessarily or inappropriately ordered in routine clinical practice; however, it does provide insight into which tests consume high amounts of health care resources (e.g., clinician time, laboratory supplies). We relied on expert opinion during stage 3 of the study to

understand which of these high-volume tests are overused.

[a]Test volume data are from 5 hospitals across 3 provinces, including Mount Saint Joseph Hospital (Vancouver, British Columbia), St. Paul's Hospital (Vancouver, British Columbia), St. Joseph's Health Centre (Toronto, Ontario), St. Michael's Hospital (Toronto, Ontario), and Queen Elizabeth II Health Sciences Centre (Halifax, Nova Scotia). Cumulative test volumes represent the number of tests conducted between January 2021 and December 2021 at Mount Saint Joseph Hospital and St. Paul's Hospital, between November 2019 and October 2020 at St. Joseph's Health Centre, between April 2021 and March 2022 at St. Michael's Hospital, and between April 2021 and March 2022 at Queen Elizabeth II Health Sciences Centre.

[b]Test volumes were normalized by dividing the volume for each test by the volume for the highest volume test (i.e., creatinine) and multiplying by 100.

[c]White blood cell count is a component of the complete blood count; however, it may also be performed independently and was therefore included as a distinct lab test.

Validation of the List

Our targeted literature search yielded 12 publications that made recommendations regarding the unnecessary use of hospital-based lab tests. While many of the lab tests described as unnecessary in the literature were already captured in our list of 40 highly used lab tests (e.g., vitamin D testing, C-reactive protein, urine cultures, TSH, complete blood count), there were several tests that were not included in our list (e.g., free triiodothyronine [FT3], free thyroxine (FT4]), prostate-specific antigen, serum cortisol, blood mercury levels, and serum testosterone, among others). However, the retrieved publications provided little to no information on how frequently these tests may be used in Canadian hospitals. Because the scope of our project is specific to highly used tests (and we excluded low-volume tests), we decided not to add the additional lab tests for consideration in the Delphi round 1 survey. Instead, we relied on the tests identified as highly used and any additional tests that were nominated for consideration by the expert panel, who were asked to consider our lab test eligibility criteria when nominating tests (i.e., lab tests that are frequently used, have low clinical value, are costly and increase health care spending without improving patient outcomes, and have potential for reduction).

Stage 3: The Online Delphi Process

The online Delphi process comprised 3 rounds, including 2 survey rounds (round 1 and round 3) and 1 online workshop round (round 2). A high-level overview of the modified Delphi process is illustrated in Figure 2.

Delphi Round 1: Initial Survey

We drafted and distributed the Delphi round 1 survey to 16 subject experts on August 11, 2022. We received completed responses from 12 subject experts. Of the 4 remaining subject experts, 2 did not open the survey link and 2 started the survey but did not proceed beyond the launch page. No participants submitted partially completed surveys.

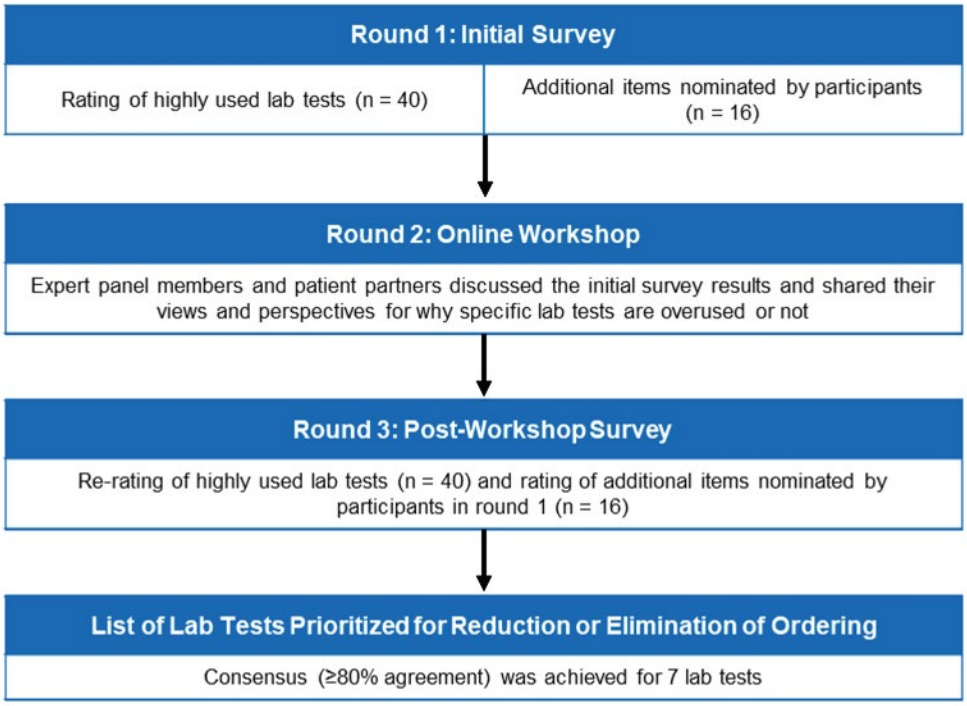
Of the 40 highly used lab tests voted on by participants in survey 1, 2 lab tests reached our cut-off for consensus (i.e., ≥ 80% of participants rating the lab test as “could be included” or “should definitely be included” in the final list). These were the 25-hydroxy vitamin D test and urine cultures. For the 25-hydroxy vitamin D test, 10 of the 12 participants (83.3%) voted that the 25-hydroxy vitamin D test should definitely be included in the final list, while 1 participant (8.3%) voted that it could be included in the final list. Urine cultures received 1 vote (8.3%) to definitely be included in the final list and 9 votes (75.0%) that they could be included in the final list. Although the 25-hydroxy vitamin D test and urine cultures achieved the threshold of consensus, they continued to be included in subsequent rounds of the Delphi process to allow participants to change their ratings in light of the information shared at the workshop, and would only be included in the final list if they maintained consensus in the post-workshop survey. The 38 lab tests that did not achieve consensus in round 1 of the Delphi process had variable vote distributions. The voting results of the Delphi round 1 survey (i.e., response frequencies) are presented in Appendix 1, Figure 3, and are further detailed in Appendix 1,

Table 3. In addition to the ratings for each item, we received 127 open-ended responses providing justification for the highly used lab tests that were rated a 3 (i.e., could be included in the final list) or 4 (i.e., should definitely be included in the final list) by participants.

As part of section II of survey 1, participants nominated 16 items for consideration in subsequent rounds of the Delphi process. The lab tests that were nominated were as follows:

- vitamin B12 test (3 nominations)
- venous blood gases (2 nominations)
- anti-nuclear antibody test (2 nominations)
- serum protein electrophoresis (1 nomination)
- lactate dehydrogenase test (1 nomination)
- pathologic examination of orthopedic devices (1 nomination)
- bacterial vaginosis test (1 nomination)
- iron panel (1 nomination)
- fecal occult blood test (1 nomination)
- pathologic examination of placentas (1 nomination)
- pathologic examination of skin or fatty tissue (1 nomination)
- corrected calcium (1 nomination)
- FT3 and FT4 tests (1 nomination)
- prostate-specific antigen test (1 nomination)
- wound swabs (1 nomination)

Figure 2: Overview of the Modified Delphi Process



- gastrointestinal tract biopsies (1 nomination).

One respondent also suggested we consider "extra tubes in anticipation of orders" as an unnecessary test. While this clinical practice may contribute to wrong blood in tube errors and unnecessarily expose patients to harm (e.g., iatrogenic anemia),[24-27] it was considered out of scope for the current project as it is not attributed to a specific lab test.

Several respondents also nominated tests that were previously identified as unnecessary by the Using Labs Wisely program (Gillian Hurwitz, personal communication, Nov 3, 2022), including partial thromboplastin time (3 respondents), creatine kinase-MB (3 respondents), folate testing (3 respondents), and urea (1 respondent). However, these tests were not added to subsequent rounds of the Delphi process as they were considered out of scope for the current project.

Delphi Round 2: Online Workshop

Fifteen subject expert panellists and 2 patient partners attended the online workshop on September 13, 2022. While we had hoped to achieve a 100% participation rate of expert panellists, 1 subject expert panellist was unavailable due to scheduling conflicts.

The online workshop provided participants with an opportunity to discuss the clinical value of each of the 40 highly used lab tests and the 16 items nominated for consideration during the survey round 1. Generally, the discussions focused on how frequently each test is used inappropriately (e.g., unnecessary duplicate or repeat testing, in clinical scenarios where the test is not indicated), the capacity for tests results to influence patient management, the balance of clinical benefits and harms associated with each test, the costs of each test, and whether more reliable tests are available.

Delphi Round 3: Post-Workshop Survey

We distributed the Delphi survey 2 to 16 experts on September 19, 2022. We received completed responses from 15 experts by October 7, 2022, when the survey closed. One participant submitted a partially complete survey, where responses to the demographic questions were provided but none of the lab tests were rated.

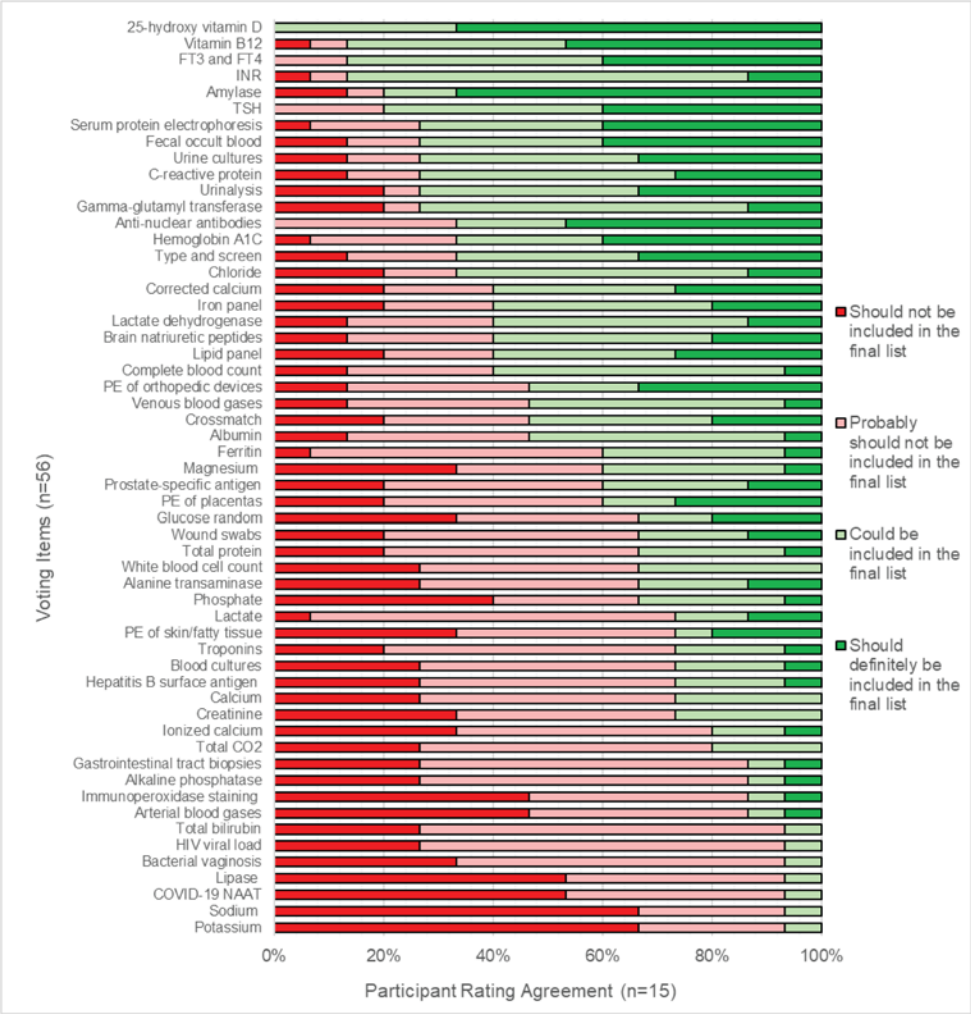
The 15 participants who submitted completed surveys voted on 56 items, including 40 from our list of highly used lab tests and 16 that were nominated for consideration during survey 1. The voting results of the Delphi survey 2 (i.e., response frequencies) are presented in Figure 3, and are further detailed in Appendix 1, Table 4.

Seven lab tests achieved our cut-off for consensus (i.e., ≥ 80% of participants rating the lab test as "could be included" or "should definitely be included" in the final list) and were included in our final list of hospital-based lab tests prioritized for reduction or elimination of ordering. These were as follows:

- 25-hydroxy vitamin D test (10 votes for definitely include [66.7%]; 5 votes for could be included [33.3%])
- vitamin B12 test (7 votes for definitely include [46.7%]; 6 votes for could be included [40.0%])
- FT3 and FT4 tests (6 votes for definitely include [40.0%]; 7 votes for could be included [46.7%])
- INR testing (2 votes for definitely include [13.3%]; 11 votes for could be included [73.3%])

- amylase test (10 votes for definitely include [66.7%]; 2 votes for could be included [13.3%])
- TSH test (6 votes for definitely include [40.0%]; 6 votes for could be included [40.0%]).

Figure 3: Delphi Survey 2 Results



CO_2 = carbon dioxide; FT3 = free triiodothyronine; FT4 = free thyroxine; INR = international normalized ratio; NAAT = nucleic acid amplification test; PE = pathologic examination; TSH = thyroid stimulating hormone.

Note: Lab tests are ordered from top to bottom by their level of consensus to include in the final list of unnecessary lab tests.

Four of the 7 tests identified for inclusion in the final list were from the initial list of the top 40 highly used lab tests, which was generated during stage 2 of the study. The remaining prioritized tests (i.e., FT3 and FT4, vitamin B12) were nominated by subject experts in the initial survey and subsequently achieved consensus for inclusion during survey 2.

Compared to survey 1, there were no substantial changes to the response frequencies for each lab test. For example, no tests listed in the bottom 10% of tests when ordered by their level of consensus were rated in the top 10% of survey 2. One notable difference between the results from survey 1 and survey 2 was that urine cultures did not achieve the cut-off for

consensus in survey 2, despite reaching the cut-off in survey 1. This can likely be explained due to the discussion at the workshop and the additional responses from expert panellists who did not respond to survey 1.

List of Hospital-Based Lab Tests Prioritized for Reduction or Elimination of Ordering

25-Hydroxy Vitamin D Test

The 25-hydroxy vitamin D test is a blood test that measures levels of vitamin D, a fat-soluble vitamin used by the body for calcium homeostasis and bone mineralization.[28] 25-hydroxy vitamin D is the most abundant circulating form of vitamin D and has been recognized as a functional indicator of vitamin D status.[29]

Throughout the modified Delphi process, expert panellists described 25-hydroxy vitamin D testing as being overutilized, unnecessary, non-informative, clinically irrelevant, and not likely to improve patient outcomes. It provides limited information of value and has no benefits for subsequent inpatient management. Relative to other routine tests, 25-hydroxy vitamin D testing is expensive; thus, it represents a substantial waste of resources. Experts noted that while there are certain populations that may benefit from vitamin D replacement (e.g., individuals with limited sun exposure, people with renal conditions, people on medications that can impact vitamin D status),[30,31] literature suggests that supplementation with vitamin D generally does not improve outcomes and may be associated with harms.[32-34] Additionally, vitamin D supplementation can be reasonably initiated in some populations without testing;[35,36] thus, the test should be highly restricted. Experts indicated that while some jurisdictions have successfully curtailed 25-hydroxy vitamin D testing, it is still overused in many jurisdictions relative to the number of appropriate indications. The Delphi voting results indicated that the 25-hydroxy vitamin D test achieved the highest consensus to include in the final list of prioritized unnecessary hospital-based lab tests, as all 15 participants of the round 2 survey indicated that it should definitely be included (n = 10 of 15; 66.7%) or could be included (n = 5 of 15; 33.3%) in the list.

Vitamin D testing has been articulated as having low value in several previous resource stewardship campaigns, including CWC,[37] Choosing Wisely Alberta,[38] Choosing Wisely Manitoba,[39] Choosing Wisely Italy,[40] Choosing Wisely Australia,[41] Choosing Wisely New Zealand,[42,43] and Quality of Care NL.[44] These previous campaigns have recommended against routine vitamin D testing in low-risk adults,[38,44] population-based screening for vitamin D deficiency,[38,40,43] and vitamin D testing at the first antenatal visit.[42] In addition to these campaigns, CADTH has conducted a review on the use of vitamin D testing in the general population.[45] While much of the previous work has not been specific to hospital settings, the limited clinical value of 25-hydroxy vitamin D testing has been accentuated.

Vitamin B12 Test

Vitamin B12, also known as cobalamin, is essential for DNA synthesis and energy metabolism.[46] Vitamin B12 levels are assessed using the vitamin B12 test, which measures the amount of vitamin B12 in the blood. The test is often conducted to check for vitamin B12 deficiency, which may occur in individuals who present with pernicious anemia or certain nervous system symptoms (e.g., numbness in the limbs, decreased mental abilities, or loss of balance).[47]

While vitamin B12 testing did not make our list of 40 highly used lab tests in Canada, it was nominated for consideration in the Delphi round 1 survey by 3 respondents. Panellists suggested that vitamin B12 testing is often inappropriately used and overused. Experts noted that this test is highly prone to unnecessary testing, such as duplicate testing within 48 hours of patient admission despite having a recent normal result on file, and that these types of unnecessary orders should be limited. According to the panellists and published literature,[48,49] vitamin B12 testing correlates poorly with biochemical deficiency of vitamin B12 and it generally only makes sense biochemically to check once a year.[50-52] When used in appropriate patient populations, such as those with anemia or neuropathy, vitamin B12 is a useful diagnostic test; however, in other scenarios it causes resource waste and can lead to unnecessary treatment with vitamin B12 injection, which has been identified as an overprescribed treatment in Canada.[53] Of the 15 subject experts who participated in the round 2 survey, 7 [46.7%] voted that vitamin B12 testing should definitely be included in the final list and 6 [40.0%] voted that it could be included in the final list.

While we did not systematically search for CW lists related to the prioritized hospital-based lab tests, our literature search did not find any CW lists that made recommendations related to vitamin B12 testing. Previous CADTH reviews have examined the use of vitamin B12 testing in the general population[54] and in people with suspected vitamin B12 deficiency.[55]

TSH Test

TSH is a hormone produced and released into the bloodstream by the pituitary gland that regulates the production of the thyroid hormones, triiodothyronine and thyroxine. The TSH test measures the level of TSH in the blood. Elevated levels of TSH may be a sign of primary hypothyroidism, while low levels of TSH may be suggestive of hyperthyroidism or, in some cases, secondary hypothyroidism.[56]

Expert panellists noted that TSH testing is ordered too frequently in inpatients and is often difficult to interpret in the context of acute medical illness in hospital settings. It was discussed that TSH testing is particularly prone to unnecessary duplicate testing, as it is often administered multiple times during the admission process (e.g., once in the emergency department and then again in the ward). TSH levels are often abnormal in hospitalized patients due to acute illness and are not usually indicative of a thyroid problem. The experts suggested that reducing unnecessary TSH testing would decrease the impact of further unnecessary tests or treatments (e.g., FT3 and FT4, investigations for panhypopituitarism, further endocrinological consultations) in the context of acute illness or euthyroid sick syndrome, and could incur substantial cost savings to health systems as TSH testing is relatively expensive compared to other routine chemistry tests. Several panellists also indicated that the use of TSH testing should be restricted in asymptomatic patients (i.e., not used as a screening test) and that testing should not be repeated more often than every 12 months in this group.

Despite the scenarios where TSH testing was described as overutilized, panellists indicated it can be a valuable test in patients where there is a high clinical suspicion for thyroid disease. The expert panellists noted that there are many published thyroid testing algorithms and guidelines, and that there is a need for raising clinician awareness of these resources.[57-59] Similarly, experts indicated that previous resource stewardship campaigns have highlighted the overutilization of TSH in various clinical scenarios, including recommendations against ordering thyroid function tests in asymptomatic patients[38,60] and ordering TSH testing at the first antenatal visit.[42]

FT3 and FT4 Tests

Triiodothyronine and thyroxine are the 2 main hormones released by the thyroid gland. They play an important role in the regulation of metabolism.[61] The amount of triiodothyronine and thyroxine in the body can be estimated using FT3 and FT4 tests, which quantify the amount of hormone not bound to serum proteins that is readily available to affect the body tissues. FT3 and FT4 testing may be used to investigate thyroid disease (e.g., hyperthyroidism, hypothyroidism) and disorders of the pituitary gland.[62]

FT3 and FT4 were nominated for consideration in our list by an expert panellist during round 1 of the Delphi process. The tests were subsequently discussed at the online workshop and voted on during round 2, where they achieved consensus for inclusion in the final list of unnecessary tests. Because these tests were not identified using our list of 40 highly used lab tests, it was less clear to what extent FT3 and FT4 are used in Canada; however, discussion among the expert panellists during the online workshop indicated that these tests are frequently used in the hospital setting, and thus meet the scope of the project.

During the Delphi process, expert panellists stated that FT3 and FT4 tests are overused in screening for thyroid dysfunction despite the guidelines and algorithms that exist, including 1 publication[63] that was shared during the workshop by an expert panellist. The experts suggested that FT3 and FT4 levels are often abnormal in unwell people and are not usually indicative of a thyroid problem, and that these tests should not be performed in hospitalized patients except where the clinical suspicion is that the patient is admitted for the primary reason of a thyroid disorder. Similar to other tests prioritized for the final list, unnecessary duplicate or repeat testing was also identified as a problem for FT3 and FT4 tests. While 1 expert noted that it may be more impactful to focus resource stewardship efforts on TSH (also prioritized for inclusion in our list), other experts suggested campaigns could package all thyroid tests (i.e., TSH, FT3, FT4) as candidates for reduction. Finally, the panel suggested that the implementation of FT3 and FT4 testing algorithms based on initial TSH results would result in large savings and prevent resource wastage.[64-66]

Many initiatives have previously recommended against the routine use of FT3 or FT4 testing in specific clinical scenarios, including CWC,[67] Choosing Wisely Manitoba,[68] Choosing Wisely Alberta,[38] Choosing Wisely Italy,[40] Choosing Wisely Australia,[41] and Quality of Care NL.[44] Specifically, previous campaigns have recommended against the use of FT3 or FT4 to screen for hypothyroidism or to monitor and adjust levothyroxine dose in patients with known primary hypothyroidism.[38,40,41,44,67,68] Previous CADTH reviews[69,70] have examined the literature on the use of FT3 and FT4 testing to screen for hypothyroidism or to monitor levothyroxine dosing in people with hypothyroidism.

INR Testing

An INR is a standardized method for expressing the result of a prothrombin time test, which assesses the coagulation status of patients by measuring the amount of time for plasma to clot after clotting reagents are added.[71] INR testing can be used to diagnose bleeding or clotting disorders (e.g., hemophilia), monitor patients being treated with anticoagulant therapy (e.g., warfarin), or assess liver function.[71]

The expert panellists suggested that INRs are often ordered unnecessarily and too frequently, particularly for hospitalized patients and in the pre-procedural setting.[72] In hospitalized patients, INRs are useful for monitoring the efficacy and safety of warfarin therapy; however, the experts indicated that INRs are often ordered daily for inpatients and

can lead to medication dosing errors, as the half-life of warfarin can be up to 60 hours. It was reported that, in many cases, INR results are non-informative and are not used to make clinical decisions. As for pre-procedural settings, experts noted that there is a lack of evidence to support INRs before many low-risk procedures, yet clinicians routinely request a measurement. To support this, they referenced a 2019 guideline[73] from the Society for Interventional Radiology that recommends against routinely screening coagulation laboratory testing before procedures with low bleeding risk. Regardless of the patient population, overuse of INR testing can lead to potentially unnecessary further testing and transfusions. Another expert panellist noted that 1 strategy to potentially reduce unnecessary INR testing would be to review hospital policies, order sets, and pre-printed orders that require INRs in specific settings (e.g., standing orders in the emergency department) to ensure the policies are appropriate.

One expert panellist explained that, unlike other common lab tests, INR testing requires a relatively large volume of blood obtained using fastidious collection. In situations where insufficient blood is collected on the first draw, a second sample is required. Re-collection requires additional clinician and patient time, and may expose the patient to harm.

Despite the many scenarios where the expert panellists agreed there is overuse of INRs, they acknowledged that this test is a valuable diagnostic tool when ordered appropriately, for example in patients with an inherent (or iatrogenic) bleeding tendency or severe liver disease, and in those receiving warfarin. This sentiment was shared by the patient partners during the online workshop session, where they expressed that INRs were a useful tool for managing their therapies and informing treatment decisions. The patient partners acknowledged that they may be part of the small patient population for whom the test is helpful.

One expert noted that reduction of unnecessary ordering will require a coordinated effort among all specialties that perform procedures (e.g., interventional radiologists, clinicians performing procedures like biopsies), and agreement that INRs are not indicated before many low-risk procedures. The panellists suggested that an overall reduction in ordering INRs would improve blood collection workflow and overall costs, and would decrease unnecessary delays in care.

INR testing was identified as the 15th highest cumulative volume lab test across our data sources used in stage 2 of this study (where a list of highly used lab tests was generated), suggesting that efforts to reduce unnecessary INR ordering could have a substantial impact. The experts also noted that initiatives to reduce INR ordering may synergize well with the previous efforts of CWC's Using Labs Wisely program to reduce partial thromboplastin time ordering, as INR and partial thromboplastin time are often performed in combination.

Several previous Choosing Wisely campaigns[41,44,74,75] have signalled that INR ordering is often unnecessary and that a reduction of this test in various settings would benefit patients and health systems. For example, previous campaigns have suggested INR should not be ordered for asymptomatic patients before minor or low-risk surgery,[74,75] recommended against coagulation studies in emergency department patients unless there is a clearly defined specific clinical indication,[41] and recommended monitoring INR monthly in stable patients on warfarin and as infrequently as every 12 weeks in very stable patients.[44]

Amylase Test

Amylase is a digestive enzyme primarily secreted by the pancreas and the salivary glands that breaks down starches into simple sugars.[76] The amylase test measures the amount of this

enzyme in a blood or urine sample. The test is often ordered to diagnose pancreatitis or other pancreatic diseases, as elevated levels of amylase are indicative of these conditions.[76]

The expert panel emphasized that amylase is an obsolete test in most clinical scenarios and should be phased out or eliminated, as lipase testing is superior in several ways (e.g., lipase testing has increased sensitivity, specificity, and clinical utility). They also noted that amylase and lipase testing are often ordered together, and that this practice is unnecessary as lipase testing alone is preferable when evaluating patients for pancreatic disease. The notion that lipase testing has increased clinical value compared to amylase testing is supported in the literature.[77-82]

One expert panellist suggested that the hesitancy of some hospitals and labs to switch from amylase testing to lipase testing may be related to the cost of the lipase test, which is substantially more expensive than amylase testing. Expert panellists also indicated it may be related to the comfort level of some physicians to stop ordering tests that they are accustomed to using for clinical decision-making.

Previous Choosing Wisely campaigns[17,83,84] have recommended against ordering amylase in addition to lipase to detect pancreatitis and have instead suggested using just lipase as it has superior diagnostic performance.

Discussion and Conclusions

Health care systems throughout the world — including in Canada — are challenged by imbalances between demand for services and available resources, particularly throughout the COVID-19 pandemic, which has strained systems in unprecedented ways. Efforts to reduce low-value care are important for increasing the availability of high-quality care for everyone.

The purpose of this study was to identify at least 5 unnecessary hospital-based lab tests for consideration for year 2 of the Using Labs Wisely program in Canada. Through a 3-stage process that involved convening a panel of lab test experts, generating a list of 40 lab tests that are highly used in Canadian hospitals, and a 3-round modified Delphi process, we have prioritized a list of 7 unnecessary lab tests that often provide low clinical value in hospital settings. The final prioritized list includes testing for 25-hydroxy vitamin D, vitamin B12, TSH, FT3 and FT4, INR, and amylase. Our findings align with other resource stewardship campaigns, including provincial and national Choosing Wisely initiatives,[17,37-39,41-44,60,67,68,74,75,83,84] as many of prioritized tests have previously been identified as overused and as having low clinical value.

The findings summarized in this report are primarily related to the lab tests that achieved our threshold for consensus and were included in the final list of prioritized unnecessary lab tests. However, the online workshop included rich and robust information on other tests that were discussed in detail but did not achieve expert consensus for inclusion in the list. For example, serum protein electrophoresis was identified as a high-volume and resource-intensive test that is often unnecessarily performed daily in patients who are acutely ill. The experts suggested that serum protein electrophoresis is prone to false-positive results that may lead to further downstream testing, which is often costly and has the potential to lead to patent harm. Similarly, urinalysis was identified as a test that is frequently ordered without clinical

indication (e.g., without signs or symptoms of a urinary tract infection or hematuria) and as having a high false-positive rate in inpatients due to the methods used for sample collection. Experts emphasized the risk for harms associated with over-testing, as false-positive results may lead to unnecessary treatment with antibiotics and potential *Clostridium difficile* infection. Additionally, experts noted that urine dip stick tests may be a useful alternative, but many clinicians will still always confirm positive or negative dipstick tests with a urinalysis. The patient partners shared some of their views and experiences related to urinalysis, which included that urine samples are often collected without a clear understanding of why the test is being done. Although our final list was limited to 7 lab tests, there may be value in reviewing practices and policies related to other lab tests that were not included in the final list but were close to achieving consensus for inclusion.

Throughout the modified Delphi process, subject experts commented on health system-level considerations that have implications for overutilization of laboratory services that are not specific to any particular lab tests. For example, subject experts emphasized that unnecessary duplicate testing can lead to strain on laboratory services as patients are seen by multiple health care professionals through the care pathway. It is possible that the implementation of health information systems with clinical decision support tools that enable improved health data sharing between clinicians and prevent repeat testing within evidence-informed time frames could help to reduce unnecessary duplicate or repeat testing.[85,86] Furthermore, improved tracking of lab test ordering per hospitalized patient could facilitate the benchmarking of ordering practices across hospitals and be used to identify targets for quality improvement initiatives.[87,88] Another consideration is that hospital and jurisdictional policies, including legal regulations, may require that lab tests are conducted in specific circumstances even when they are not supported by evidence. Subject experts suggested that these policies can act as a barrier to initiatives aimed at reducing unnecessary testing. Furthermore, the importance of clinician engagement and education was mentioned throughout the online workshop.[89] Subject experts indicated that some physicians will order lab tests out of habit or because they are accustomed to seeing the result, rather than because the test will improve patient outcomes or because the test result is critical to informing clinical decisions. Previous literature has suggested that educational interventions, including those that aim to improve the teaching of pathology and laboratory medicine during medical school and residency, may reduce inappropriate lab testing over time.[3,90,91]

The findings from this study are meant to identify hospital-based lab tests that could be considered by CWC's Using Labs Wisely program. Choosing Wisely recommendations are published to encourage thoughtful conversations among clinicians, patients, and decision-makers about the use of low-value interventions. Clinicians and decision-makers who are involved in conducting lab tests and establishing policies related to their use may find the results of this study relevant to their settings and informative for quality improvement initiatives. While the scope of this study was specific to hospital settings, the discussions related to the clinical value of each of the prioritized lab tests may be relevant to other settings (e.g., community settings). Previous initiatives that have implemented recommendations from Choosing Wisely campaigns related to reducing unnecessary lab testing have contributed to health systems outcomes, such as achieving significant reductions in unnecessary daily basic metabolic panel and complete blood count testing,[92] increased adherence to guideline-concordant ordering of cardiac biomarkers,[93] and significant decreases in the proportion of inappropriate vitamin D screening.[94] Future studies that evaluate the impact of efforts to reduce unnecessary testing with the prioritized lab tests would further contextualize the findings of this study and provide additional insights for clinicians and decision-makers.

Methods

Stage 1: Creating the Expert Panel

Expert Panel Sample and Recruitment Process

We used purposive maximum variation sampling to select and recruit subject experts from diverse geographical settings across Canada, and with diverse years and variety of working experience and professional designation. We defined subject experts as individuals who had knowledge and/or experience of ordering and conducting lab tests and were directly affected by the use of lab tests in their profession. Following this rationale, we aimed for the expert sample to include medical laboratory professionals (including those who specialize in transfusion, clinical chemistry, hematology, microbiology, and histology), physicians (including internal medicine physicians, pathologists, and emergency physicians), and decision-makers (that is, those decision-makers working on or participating in quality improvement projects and who have knowledge of the barriers to quality, performance, or flaws in system design). Additionally, experts included patients and caregivers who had lived experience with lab testing. We sought to recruit a diverse group of experts with between 10 and 20 members.

We first recruited subject experts who had collaborated with CWC previously, as their past experiences with resource stewardship campaigns was expected to increase their engagement with our project. We invited them via email to participate in the panel. We then identified and recruited further panel members using a snowball sampling technique. Subject experts agreed by email to participate in the study.

Patient partners were recruited by CADTH's Patient Engagement team. Interested individuals completed a form describing their interest in participating and how their experiences could add to the diversity of ideas shared in the modified Delphi workshop session.

Stage 2: Generating the List of Highly Used Lab Tests

Concurrently with stage 1, we generated a list of 40 highly used lab tests for consideration by the expert panel during the modified Delphi process. While utilization data does not provide any information on which tests are unnecessarily ordered in routine clinical practice, it does provide insight into which tests consume high amounts of health care resources (e.g., clinician time, laboratory supplies). Ideally, we would refer to the literature for published studies that have assessed the degree to which commonly used lab tests are unnecessarily ordered, and the clinical or cost-effectiveness of laboratory tests when used inappropriately. However, clinical studies that provide this type of evidence are rarely conducted and are often not feasible.[95] Therefore, we relied on expert opinion to understand and prioritize unnecessary, overused tests using our list of 40 highly used tests as a starting point.

Volume Testing Data Collection

We collected hospital lab test utilization data from available data sources across Canada via CWC and CADTH's professional networks. We retrieved lab test volume data from 5 hospitals across 3 provinces. These included Mount Saint Joseph Hospital (Vancouver, British Columbia), St. Paul's Hospital (Vancouver, British Columbia), St. Joseph's Health Centre (Toronto, Ontario), St. Michael's Hospital (Toronto, Ontario), and Queen Elizabeth II Health Sciences Centre (Halifax, Nova Scotia).

The data from the British Columbia hospitals (i.e., Mount Saint Joseph Hospital and St. Paul's Hospital) included lab tests conducted between January 2021 and December 2021. The data from the Ontario hospitals, St. Joseph's Health Centre and St. Michael's Hospital, included labs tests conducted from November 2019 to October 2020 and April 2021 to March 2022, respectively. Finally, the data from Queen Elizabeth II Health Sciences Centre in Halifax, Nova Scotia, included lab tests conducted between April 2021 and March 2022.

Volume Data Cleaning and Drafting the List of 40 Highly Used Lab Tests

The retrieved lab utilization data were filtered and cleaned to ensure comparable lab tests were identified and labelled using a single name. Test codes that corresponded to calculations (e.g., estimated glomerular filtration rate) and not specific lab tests were removed from the datasets. In some instances, common lab panels were grouped into a single data point (i.e., the volume of test ordering was not available for each component of the panel, only the panel itself), while each individual test was listed in others. To resolve this, the individual tests (e.g., triglycerides, HDL cholesterol, LDL cholesterol, total cholesterol) were grouped into a panel (e.g., lipid panel) and assigned the volume of the most highly ordered component of the panel. When needed, we sought clarification from the contacts who submitted the data to understand and filter the data. Once the data were sufficiently filtered and homogenous to permit combining, the number of lab tests conducted at each hospital were combined, and we selected the top 40 highly used lab tests across all data sources using cumulative test volumes. We excluded 5 tests that were previously identified for reduction or elimination of ordering by the Using Labs Wisely program (i.e., aspartate transaminase testing, partial thromboplastin time testing, creatine kinase-MB testing, folate testing, and urea testing) (Gillian Hurwitz, personal communication, Nov 3, 2022).

Literature Search and Validation of the List of 40 Highly Used Lab Tests

A targeted internet search was conducted by an information specialist for Canadian and major international lists of unnecessary or inappropriate laboratory tests. The search was completed on July 15, 2022, and limited to English-language documents.

We reviewed lab test recommendations from existing Choosing Wisely campaigns or other health care stewardship initiatives, together with supporting evidence, to support validation of the identified top overused lab tests in Canada.

Stage 3: The Online Delphi Consensus-Building Process

We adopted a modified Delphi method that combines the anonymous structured multi-stage collection of expert opinion using surveys with open exchange during a workshop moderated by a facilitator.[23] The use of surveys ensured a degree of anonymity for the individual responses, supported by individualized feedback and a clarification stage through open exchanges during a workshop. During the workshop, participants had the opportunity to discuss the items and ensure shared understanding.

Delphi Round 1: Initial Survey

In the first round, we asked members of the expert panel to identify which frequently used lab tests are unnecessary via an anonymous online survey. This was the first of 2 surveys distributed to the expert panel members. The first survey provided a list of 40 highly used lab tests in Canada (generated during stage 2 of the project) and asked the expert panel members to indicate which ones are unnecessary and why. The survey first instructed respondents to consider the list of criteria for including or excluding lab tests from the

final list. Then, the survey invited experts to respond to a set of closed-ended demographic questions and a series of closed-ended questions asking panellists to rate each of the 40 highly used lab tests on a 4-point Likert scale (indicating whether each test should not be considered for inclusion, probably should not be included, could be included, or should definitely be included in the final list of prioritized unnecessary lab tests). When respondents indicated a test could or definitely should be included in the final list, the survey asked respondents to provide a written rationale explaining why via open-ended answers. Finally, the survey also offered respondents the opportunity to add up to 5 additional unnecessary lab tests for consideration in subsequent rounds of the Delphi process (if they thought we missed any of relevance).

A draft of the survey was reviewed internally and was pilot-tested by CADTH staff. A final version of the questionnaire used in the survey is available on request to CADTH.

Once the survey closed, we collected the responses, determined the level of agreement between the responses of the panellists for each listed item, and summarized the responses. Open-ended responses indicating why items were deemed for inclusion in the final list were narratively summarized for each item using a descriptive approach to identify categories of similar meanings.

When interpreting results of the survey, responses were considered partially complete if 1 or more of the highly used lab tests were not rated by the respondent. Participants who did not provide responses to any demographic questions or ratings for at least 1 lab test, were considered to not have participated in the Delphi round but were still invited to participate in subsequent rounds.

Delphi Round 2: Online Workshop

We held a virtual, online, face-to-face workshop with all expert panel members and patient partners to discuss the first-round results and share insights to ensure shared understanding and clarification.[23] Prior to the workshop, we shared with each panel member the individualized summary of the results of the first survey included the expert panel's collective response to each lab test (i.e., proportion of responses for each test), the median response to each test, the participants' own responses from round 1, a list of lab tests that were not included in survey 1 but were nominated for consideration by respondents, and narrative summaries of any justifications for including lab tests in the top 5 list provided as open-ended responses by participants in survey 1. We also shared a summary of interviews conducted with the patient partners, which described some of their views and experiences related to hospital-based lab testing. The patient partners had lived experience of travelling to hospitals, undergoing repeated testing, coordinating appointments around test results, and more. They offered unique knowledge and perspectives that had implications beyond any specific lab tests being considered.

During the workshop, the facilitator presented the survey results and encouraged panellists to discuss which lab tests were the ideal candidates for inclusion in the final list of unnecessary lab tests and why. The panel facilitator prioritized items where there was considerable disagreement in the panel ratings to promote a sharing of divergent ideas, and the additional lab tests nominated by panellists, inviting further discussion on whether these items would be appropriate for inclusion in our final list. Patient partners shared their experiences as patients and/or caregivers with hospital-based lab tests during the workshop, to inform the discussions and the consensus-generating process. A written summary of the

workshop discussions was provided to subject expert panellists for them to refer to when completing survey 2.

Delphi Round 3: Post-Workshop Survey

In the second survey round, we asked members of the expert panel to reconsider their responses in light of the group's responses to the first survey and the discussion at the online workshop via an anonymous online survey. Despite including much of the same information and many of the same questions as survey 1, there were some key differences between the surveys:

- All tests that were suggested for addition to the list of unnecessary tests by respondents of survey 1 were now included as items to be rated by each participant on the 4-point Likert scale.
- The second survey no longer provided an opportunity to add additional overused lab tests for consideration in subsequent rounds of the Delphi process.

Stopping criterion and consensus definition: As it represents a standard threshold of consensus, we considered lab tests that were rated as "could be included" or "should definitely be included" in the final list by 80% or more of the expert panellists to have reached the cut-off for consensus.[96] All other scores were regarded as not having reached consensus. The top 5 (or more) unnecessary lab tests with high agreement for inclusion (≥ 80%) were included in the final list.

Stakeholder Feedback

A draft version of this report was sent to expert participants to elicit stakeholder feedback. Relevant stakeholder feedback was incorporated into the final version of this report based on the input received.

Limitations

The final list of hospital-based lab tests prioritized for reduction or elimination of ordering was informed by the views, experiences, and expertise of 16 professionals who routinely order, conduct, or interpret the results of hospital-based lab tests, and 2 patient partners who have lived experience with lab tests. While there is no consensus on the optimal number of panellists in Delphi studies[97] and many Delphi panels in health science research use between 15 to 20 panellists,[97-99] recruiting a larger pool of experts may have allowed for a more diverse range of views and perspectives to be incorporated into the consensus-generating process. We attempted to mitigate the impact of this limitation by carefully selecting and recruiting Delphi participants from diverse geographical settings across Canada, and with diverse years and variety of working experience and professional designations. However, given the massive scope of laboratory testing and the reality that laboratory professionals usually only specialize in 1 area (e.g., clinical chemistry),[100] some professional sub-specialties may have been under-represented in our expert panel (e.g., the panel included 1 expert with professional specialization in hematology and 1 expert with professional specialization in microbiology).

We instructed members of the expert panel to consider the clinical value (i.e., the balance of clinical benefits and harms, and the cost-effectiveness) of each lab test when filling out

the surveys; however, systematic literature searches were not conducted to identify relevant publications and this report is not intended as a comprehensive review of the clinical effectiveness or cost-effectiveness for any of the aforementioned hospital-based lab tests.

Throughout the online Delphi process, we combined FT3 and FT4 tests as a single item, as these tests were nominated together by an expert during Delphi round 1 and they have appeared together in other Choosing Wisely lists.[38,67,68] Specifically, they were discussed together during the online workshop (i.e., Delphi round 2) and were voted on as a single item during the Delphi round 2 survey, where they achieved our threshold for consensus. Expert participants had conflicting views on whether these tests should be grouped or not. Some suggested that grouping the 2 together (or even together with TSH) was preferable, as they are related thyroid hormones, while others recommended listing FT3 and FT4 as separate items, as they have different ordering indications. It is unclear if and how our findings or conclusions would have been impacted if FT3 and FT4 were listed as distinct voting items.

Delphi studies are inherently susceptible to several forms of bias, such as dominance (i.e., when group members are able to exhibit a disproportionate amount of control over the ratings of other members, often due to vocality or outspoken nature) and group conformity (i.e., the tendency for people to behave or vote in line with other members of the group rather than using their own personal judgment).[99,101,102] While we attempted to mitigate these biases by using anonymous surveys and by promoting equitable and inclusive discussion during the online workshop, these may have impacted the Delphi process.

While we aimed to include experts from across most Canadian provinces and territories, our Delphi panel did not include any experts from Manitoba, New Brunswick, Northwest Territories, Nunavut, Prince Edward Island, or Yukon. These provinces and territories may have unique circumstances related to the use of hospital-based lab tests. The applicability of the final list of unnecessary lab tests to all Canadian jurisdictions should be further explored.

References

1. Hallworth MJ. The '70% claim': what is the evidence base? *Ann Clin Biochem.* 2011;48(6):487-488. PubMed
2. Freedman DB. Towards Better Test Utilization - Strategies to Improve Physician Ordering and Their Impact on Patient Outcomes. *EJIFCC.* 2015;26(1):15-30. PubMed
3. Naugler C, Wyonch W. What the Doctor Ordered: Improving the Use and Value of Laboratory Testing. Toronto (ON): C.D. HOWE Institute; 2019: https://www.cdhowe.org/sites/default/files/2021-11/Commentary_533.pdf. Accessed 2022 Dec 5.
4. Zhi M, Ding EL, Theisen-Toupal J, Whelan J, Arnaout R. The landscape of inappropriate laboratory testing: a 15-year meta-analysis. *PLoS One.* 2013;8(11):e78962. PubMed
5. Koch C, Roberts K, Petruccelli C, Morgan DJ. The Frequency of Unnecessary Testing in Hospitalized Patients. *Am J Med.* 2018;131(5):500-503. PubMed
6. Bodley T, Chan M, Levi O, et al. Patient harm associated with serial phlebotomy and blood waste in the intensive care unit: A retrospective cohort study. *PLOS ONE.* 2021;16(1):e0243782. PubMed
7. Makam AN, Nguyen OK, Clark C, Halm EA. Incidence, Predictors, and Outcomes of Hospital-Acquired Anemia. *J Hosp Med.* 2017;12(5):317-322. PubMed
8. Beriault DR, Gilmour JA, Hicks LK. Overutilization in laboratory medicine: tackling the problem with quality improvement science. *Crit Rev Clin Lab Sci.* 2021;58(6):430-446. PubMed
9. Yusuf E, Luijendijk A, Roo-Brand G, Friedrich AW. The unintended contribution of clinical microbiology laboratories to climate change and mitigation strategies: a combination of descriptive study, short survey, literature review and opinion. *Clin Microbiol Infect.* 2022;28(9):1245-1250. PubMed
10. Hensher M, Canny B, Zimitat C, Campbell J, Palmer A. Health care, overconsumption and uneconomic growth: A conceptual framework. *Soc Sci Med.* 2020;266:113420. PubMed
11. Yeo SC, Ooi XY, Tan TSM. Sustainable kidney care delivery and climate change - a call to action. *Global Health.* 2022;18(1):75. PubMed
12. Hensher M, Tisdell J, Zimitat C. "Too much medicine": Insights and explanations from economic theory and research. *Soc Sci Med.* 2017;176:77-84. PubMed
13. Health Care Sustainability Metrics: Building A Safer, Low-Carbon Health System. *Health Affairs.* 2020;39(12):2080-2087. PubMed
14. Medical Professional Society Handbook. Toronto (ON): Choosing Wisely Canada; 2020: https://choosingwiselycanada.org/wp-content/uploads/2022/02/2020-10-01_Medical-professional-society-handbook-V1.4.pdf. Accessed 2022 Dec 5.
15. Choosing Wisely Canada. About. https://choosingwiselycanada.org/about/. Accessed 2022 Dec 5.
16. Choosing Wisely Canada. Choosing Wisely Canada Recommendations. https://choosingwiselycanada.org/recommendations/. Accessed 2022 Dec 5.
17. Choosing Wisely. Medical Biochemistry: Five Tests and Treatments to Question. 2022; https://choosingwiselycanada.org/recommendation/medical-biochemistry/. Accessed 2022 Dec 5.
18. Choosing Wisely Canada. Using Labs Wisely. https://choosingwiselycanada.org/hospitals/using-labs-wisely/. Accessed 2022 Dec 5.
19. Brownlee SM, Korenstein D. Better understanding the downsides of low value healthcare could reduce harm. *BMJ.* 2021;372:n117. PubMed
20. Choosing Wisely Canada. FAQ. https://choosingwiselycanada.org/faq/. Accessed 2022 Dec 5.
21. Grimshaw JM, Patey AM, Kirkham KR, et al. De-implementing wisely: developing the evidence base to reduce low-value care. *BMJ Quality & Safety.* 2020;29(5):409-417. PubMed
22. Jones J, Hunter D. Qualitative Research: Consensus methods for medical and health services research. *BMJ.* 1995;311(7001):376-380. PubMed
23. McMillan SS, King M, Tully MP. How to use the nominal group and Delphi techniques. *Int J Clin Pharm.* 2016;38(3):655-662. PubMed
24. Snozek CLH, Hernandez JS, Traub SJ. "Rainbow Draws" in the Emergency Department: Clinical Utility and Staff Perceptions. *J Appl Lab Med.* 2019;4(2):229-234. PubMed
25. Dunbar NM, Delaney M, Murphy MF, et al. Emergency departments are higher-risk locations for wrong blood in tube errors. *Transfusion.* 2021;61(9):2601-2610. PubMed
26. Dunbar NM, Kaufman RM. Factors associated with wrong blood in tube errors: An international case series - The BEST collaborative study. *Transfusion.* 2022;62(1):44-50. PubMed
27. Grady D. Let's Get Another Tube—Just in Case. *JAMA Internal Medicine.* 2017;177(1):130-130. PubMed
28. Sizar O, Khare S, Goyal A, Givler A. Vitamin D Deficiency. Treasure Island (FL): StatPearls Publishing; 2022: https://www.ncbi.nlm.nih.gov/books/NBK532266/. Accessed 2022 Dec 5.
29. Greene-Finestone LS, Berger C, de Groh M, et al. 25-Hydroxyvitamin D in Canadian adults: biological, environmental, and behavioral correlates. *Osteoporos Int.* 2011;22(5):1389-1399. PubMed
30. Pludowski P, Takacs I, Boyanov M, et al. Clinical Practice in the Prevention, Diagnosis and Treatment of Vitamin D Deficiency: A Central and Eastern European Expert Consensus Statement. *Nutrients.* 2022;14(7). PubMed

31. Wakeman M. A Literature Review of the Potential Impact of Medication on Vitamin D Status. *Risk Manag Healthc Policy.* 2021;14:3357-3381. PubMed

32. Pilz S, Trummer C, Theiler-Schwetz V, et al. Critical Appraisal of Large Vitamin D Randomized Controlled Trials. *Nutrients.* 2022;14(2). PubMed

33. Bouillon R, Manousaki D, Rosen C, Trajanoska K, Rivadeneira F, Richards JB. The health effects of vitamin D supplementation: evidence from human studies. *Nature Reviews Endocrinology.* 2022;18(2):96-110. PubMed

34. Letavernier E, Daudon M. Vitamin D, Hypercalciuria and Kidney Stones. *Nutrients.* 2018;10(3). PubMed

35. Kennel KA, Drake MT, Hurley DL. Vitamin D deficiency in adults: when to test and how to treat. *Mayo Clin Proc.* 2010;85(8):752-757; quiz 757-758. PubMed

36. Vitamin D Testing. Victoria (BC): BCGuidelines.ca; 2019: https://www2.gov.bc.ca/assets/gov/health/practitioner-pro/bc-guidelines/vitamin_d_testing_for_website_hlbclinkupdate.pdf. Accessed 2022 Dec 5.

37. Vitamin D Tests: When you need them—and when you don't. Toronto (ON): Choosing Wisely Canada; 2017: https://choosingwiselycanada.org/wp-content/uploads/2017/06/Vitamin-D-EN.pdf. Accessed 2022 Dec 5.

38. Choosing Wisely Alberta. Choosing Wisely Alberta Resources. https://www.albertadoctors.org/choosing-wisely-resources. Accessed 2022 Dec 5.

39. Choosing Wisely Manitoba. Appropriate Vitamin D Testing. *Local Initiatives* https://choosingwiselymanitoba.ca/cwmb-initiatives/#1548802912890-b794910d-e3ea. Accessed 2022 Dec 5.

40. Choosing Wisely Italy. Non usare la determinazione della 25OH vitamina D come esame di screening per valutare lo status di ipovitaminosi nella popolazione generale. *Recommendations*. Accessed 2022 Dec 5.

41. Choosing Wisely Australia. Tests, treatments, and procedures for healthcare providers and consumers to question. *Recommendations* https://www.choosingwisely.org.au/recommendations?category=medicalTest. Accessed 2022 Dec 5.

42. The Royal Australian and New Zealand College of Obstetricians and Gynaecologists - Te Kāhui Oranga ō Nuku. Recommendations and Resources for Clinicians. 2021; https://www.hqsc.govt.nz/resources/choosing-wisely/recommendations-and-resources/for-clinitians/the-royal-australian-and-new-zealand-college-of-obstetricians-and-gynaecologists-te-kahui-oranga-o-nuku/. Accessed 2022 Dec 5.

43. The Royal College of Pathologists of Australasia. Recommendations and Resources for Clinicians. 2021; https://www.hqsc.govt.nz/resources/choosing-wisely/recommendations-and-resources/for-clinitians/the-royal-college-of-pathologists-of-australasia/. Accessed 2022 Dec 5.

44. Guidelines for Medical Laboratory Testing. St. John's (NL): Quality Of Care NL; 2022: https://qualityofcarenl.ca/wp-content/uploads/2022/07/PracticePoints_Vol8_p62_Guidelines_for_Medical_Laboratory_Testing.pdf. Accessed 2022 Dec 5.

45. Vitamin D Testing in the General Population: A Review of the Clinical and Cost-Effectiveness and Guidelines. Ottawa (ON): CADTH; 2015: https://www.cadth.ca/sites/default/files/pdf/htis/jan-2015/RC0626%20Vitamin%20D%20Testing%20Final.pdf. Accessed 2022 Dec 5.

46. Ankar A, Kumar A. Vitamin B12 Deficiency. Treasure Island (FL): StatPearls Publishing; 2022: https://www.ncbi.nlm.nih.gov/books/NBK441923/#:~:text=Once%20absorbed%2C%20vitamin%20B12%20is,decreasing%20the%20likelihood%20of%20deficiency. Accessed 2022 Dec 5.

47. Wolffenbuttel BHR, Wouters H, Heiner-Fokkema MR, van der Klauw MM. The Many Faces of Cobalamin (Vitamin B(12)) Deficiency. *Mayo Clin Proc Innov Qual Outcomes.* 2019;3(2):200-214. PubMed

48. Devalia V, Hamilton MS, Molloy AM, Haematology tBCfSi. Guidelines for the diagnosis and treatment of cobalamin and folate disorders. *Br J Haematol.* 2014;166(4):496-513. PubMed

49. Heil SG, de Jonge R, de Rotte MC, et al. Screening for metabolic vitamin B12 deficiency by holotranscobalamin in patients suspected of vitamin B12 deficiency: a multicentre study. *Ann Clin Biochem.* 2012;49(Pt 2):184-189. PubMed

50. Gill P, Guo M, Lau CK, Naugler C. Implementation of an educational province-wide intervention to reduce redundant vitamin B12 testing: A cross-sectional study. *Clinical Biochemistry.* 2020;76:1-4. PubMed

51. National Minimum Re-testing Interval Project: A final report detailing consensus recommendations for minimum re-testing intervals for use in Clinical Biochemistry. London (UK): The Royal College of Pathologists; 2013: https://www.rcpath.org/uploads/assets/bbcd0eb4-e250-4a09-80ec5e7139ab4fb8/ACB-Minimum-Retesting-Intervals-for-Clinical-Biochemistry-recommendations.pdf. Accessed 2022 Dec 5.

52. Vitamin B12 screening and appropriate repeat intervals. *Laboratory Bulletin*. Edmonton (AB): Alberta Health Services; 2017: https://www.albertahealthservices.ca/assets/wf/lab/wf-lab-bulletin-vitamin-b12-screening-and-appropriate-repeat-intervals.pdf. Accessed 2022 Dec 5.

53. Silverstein WK, Lin Y, Dharma C, Croxford R, Earle CC, Cheung MC. Prevalence of Inappropriateness of Parenteral Vitamin B12 Administration in Ontario, Canada. *JAMA Internal Medicine.* 2019;179(10):1434-1436. PubMed

54. Vitamin B12 Testing in the General Population: Clinical and Cost-Effectiveness and Guidelines. Ottawa (ON): CADTH; 2015: https://www.cadth.ca/sites/default/files/pdf/htis/jan-2015/RB0781%20Vitamin%20B12%20Testing%20Final.pdf. Accessed 2022 Dec 5.

55. Vitamin B12 Testing in People With Suspected Vitamin B12 Deficiency. Ottawa (ON): CADTH; 2022: https://www.cadth.ca/sites/default/files/pdf/htis/2022/RC1413%20vitamin%20B12%20testing.pdf. Accessed 2022 Dec 5.

56. Pirahanchi Y, Toro F, Jialal I. Physiology, Thyroid Stimulating Hormone. *StatPearls*. Treasure Island (FL): StatPearls Publishing; 2022: https://www.ncbi.nlm.nih.gov/books/NBK499850/. Accessed 2022 Dec 5.

57. Routine Prenatal Thyroid-Stimulating Hormone Testing: Evidence-Based Guidelines. Ottawa (ON): CADTH; 2016: https://www.cadth.ca/sites/default/files/pdf/htis/2016/RB1038%20Thyroid%20Screening%20Final.pdf. Accessed 2022 Dec 5.
58. Birtwhistle R, Morissette K, Dickinson JA, et al. Recommendation on screening adults for asymptomatic thyroid dysfunction in primary care. *CMAJ.* 2019;191(46):E1274-E1280. PubMed
59. Screening for Thyroid Dysfunction: U.S. Preventive Services Task Force Recommendation Statement. *Annals of Internal Medicine.* 2015;162(9):641-650. PubMed
60. Choosing Wisely Canada. Family Medicine: Thirteen Tests and Treatments to Question. 2022; https://choosingwiselycanada.org/recommendation/family-medicine/. Accessed 2022 Dec 5.
61. Mullur R, Liu YY, Brent GA. Thyroid hormone regulation of metabolism. *Physiol Rev.* 2014;94(2):355-382. PubMed
62. Kurian M, Kapoor N. Interpretation of thyroid function tests. *Current Medical Issues.* 2018;16(2):34-38.
63. Thyroid Function Testing in the Diagnosis and Monitoring of Thyroid Function Disorder. Victoria (BC): BCGuidelines.ca; 2018: https://www2.gov.bc.ca/assets/gov/health/practitioner-pro/bc-guidelines/thyroid-function-testing.pdf. Accessed 2022 Dec 5.
64. Taher J, Beriault DR, Yip D, Tahir S, Hicks LK, Gilmour JA. Reducing free thyroid hormone testing through multiple Plan-Do-Study-Act cycles. *Clin Biochem.* 2020;81:41-46. PubMed
65. Taher J, Brinc D, Gilmour JA, Beriault DR. Validating thyroid-stimulating hormone (TSH) reflexive testing cutpoints in a tertiary care institution. *Clin Chem Lab Med.* 2019;58(1):e11-e13. PubMed
66. Feldt-Rasmussen U, Klose M. Clinical Strategies in the Testing of Thyroid Function. In: Feingold KR, Anawalt B, Boyce A, et al., eds. *Endotext. South Dartmouth (MA): MDText.com, Inc.*; 2000.
67. Choosing Wisely Canada. Endocrinology and Metabolism: Five Tests and Treatments to Question. 2020; https://choosingwiselycanada.org/recommendation/endocrinology-and-metabolism/. Accessed 2022 Dec 5.
68. Choosing Wisely Manitoba. Thyroid Function Testing. *Local Initiatives* https://choosingwiselymanitoba.ca/cwmb-initiatives/#1570471224194-f61d4c91-ed50. Accessed 2022 Dec 5.
69. Thyroid Testing for Hypothyroidism: Clinical Effectiveness and Guidelines. Ottawa (ON): CADTH; 2018: https://www.cadth.ca/sites/default/files/pdf/htis/2018/RB1281%20Thyroid%20Testing%20Final.pdf. Accessed 2022 Dec 5.
70. Free T3 or T4 Levels for Hypothyroidism: Clinical Effectiveness and Guidelines. Ottawa (ON): CADTH; 2015: https://www.cadth.ca/sites/default/files/pdf/htis/apr-2015/RB0847%20Free%20T3%20or%20T4%20Levels%20for%20Hypothyroidism%20Final.pdf. Accessed 2022 Dec 5.
71. Riley RS, Rowe D, Fisher LM. Clinical utilization of the international normalized ratio (INR). *J Clin Lab Anal.* 2000;14(3):101-114. PubMed
72. Ambasta A, Pancic S, Wong BM, Lee T, McCaughey D, Ma IWY. Expert Recommendations on Frequency of Utilization of Common Laboratory Tests in Medical Inpatients: a Canadian Consensus Study. *J Gen Intern Med.* 2019;34(12):2786-2795. PubMed
73. Patel IJ, Rahim S, Davidson JC, et al. Society of Interventional Radiology Consensus Guidelines for the Periprocedural Management of Thrombotic and Bleeding Risk in Patients Undergoing Percutaneous Image-Guided Interventions--Part II: Recommendations: Endorsed by the Canadian Association for Interventional Radiology and the Cardiovascular and Interventional Radiological Society of Europe. *J Vasc Interv Radiol.* 2019;30(8):1168-1184.e1161. PubMed
74. Choosing Wisely Manitoba. Improving Preoperative Diagnostic Testing. *Local Initiatives* https://choosingwiselymanitoba.ca/cwmb-initiatives/#1548798013116-ee26595a-a067. Accessed 2022 Dec 5.
75. Lab Wisely. Routine preoperative chest x-rays and baseline laboratory studies, such as complete blood count, metabolic panel, or coagulation studies, should not be obtained in patients undergoing low-risk surgery with no significant systemic disease (ASA I or II) and the absence of symptom. https://labwisely.ca/routine-preoperative-chest-x-rays-and-baseline-laboratory-studies-such-as-complete-blood-count-metabolic-panel-or-coagulation-studies-should-not-be-obtained-in-patients-undergoing-low-risk-surgery/. Accessed 2022 Dec 5.
76. Akinfemiwa O, Muniraj T. Amylase. Treasure Island (FL): StatPearls Publishing; 2022: https://www.ncbi.nlm.nih.gov/books/NBK557738/. Accessed 2022 Dec 5.
77. Batra HS, Kumar A, Saha TK, Misra P, Ambade V. Comparative study of serum amylase and lipase in acute pancreatitis patients. *Indian J Clin Biochem.* 2015;30(2):230-233. PubMed
78. Hammami MB, Khoury M, Sule SS, Haire HM. Superfluous Amylase/Lipase Testing at a University-Affiliated Teaching Hospital: A Retrospective Review. *Ochsner J.* 2019;19(2):102-106. PubMed
79. Sutton PA, Humes DJ, Purcell G, et al. The role of routine assays of serum amylase and lipase for the diagnosis of acute abdominal pain. *Ann R Coll Surg Engl.* 2009;91(5):381-384. PubMed
80. Ismail OZ, Bhayana V. Lipase or amylase for the diagnosis of acute pancreatitis? *Clin Biochem.* 2017;50(18):1275-1280. PubMed
81. Smith RC, Southwell-Keely J, Chesher D. Should serum pancreatic lipase replace serum amylase as a biomarker of acute pancreatitis? *ANZ J Surg.* 2005;75(6):399-404. PubMed
82. Yadav D, Agarwal N, Pitchumoni CS. A critical evaluation of laboratory tests in acute pancreatitis. *Am J Gastroenterol.* 2002;97(6):1309-1318. PubMed
83. Choosing Wisely Italy. Raccomandazioni. https://choosingwiselyitaly.org/raccomandazioni/. Accessed 2022 Dec 5.

84. Choosing Wisely. Do not test for amylase in cases of suspected acute pancreatitis. Instead, test for lipase. 2016; https://www.choosingwisely.org/clinician-lists/american-society-clinical-pathology-testing-for-amylase/. Accessed 2022 Dec 5.
85. Procop GW, Keating C, Stagno P, et al. Reducing Duplicate Testing: A Comparison of Two Clinical Decision Support Tools. *Am J Clin Pathol.* 2015;143(5):623-626. PubMed
86. Stewart BA, Fernandes S, Rodriguez-Huertas E, Landzberg M. A preliminary look at duplicate testing associated with lack of electronic health record interoperability for transferred patients. *J Am Med Inform Assoc.* 2010;17(3):341-344. PubMed
87. Signorelli H, Straseski JA, Genzen JR, Walker BS, Jackson BR, Schmidt RL. Benchmarking to Identify Practice Variation in Test Ordering: A Potential Tool for Utilization Management. *Laboratory Medicine.* 2015;46(4):356-364. PubMed
88. Bevins NJ, Luevano DR, Nuspl R, Wang-Rodriguez J. Test Volume Ratio Benchmarking to Identify and Reduce Low-Value Laboratory Utilization. *Am J Clin Pathol.* 2021;156(4):708-714. PubMed
89. Hofstede H, van der Burg HAM, Mulder BC, et al. Reducing unnecessary vitamin testing in general practice: barriers and facilitators according to general practitioners and patients. *BMJ Open.* 2019;9(10):e029760. PubMed
90. Khromova V, Gray TA. Learning needs in clinical biochemistry for doctors in foundation years. *Ann Clin Biochem.* 2008;45(Pt 1):33-38. PubMed
91. Cadogan SL, Browne JP, Bradley CP, Cahill MR. The effectiveness of interventions to improve laboratory requesting patterns among primary care physicians: a systematic review. *Implementation Science.* 2015;10(1):167. PubMed
92. Iams W, Heck J, Kapp M, et al. A Multidisciplinary Housestaff-Led Initiative to Safely Reduce Daily Laboratory Testing. *Acad Med.* 2016;91(6):813-820. PubMed
93. Larochelle MR, Knight AM, Pantle H, Riedel S, Trost JC. Reducing excess cardiac biomarker testing at an academic medical center. *J Gen Intern Med.* 2014;29(11):1468-1474. PubMed
94. Felcher AH, Gold R, Mosen DM, Stoneburner AB. Decrease in unnecessary vitamin D testing using clinical decision support tools: making it harder to do the wrong thing. *J Am Med Inform Assoc.* 2017;24(4):776-780. PubMed
95. Baird GS. The Choosing Wisely initiative and laboratory test stewardship. *Diagnosis.* 2019;6(1):15-23. PubMed
96. Diamond IR, Grant RC, Feldman BM, et al. Defining consensus: a systematic review recommends methodologic criteria for reporting of Delphi studies. *J Clin Epidemiol.* 2014;67(4):401-409. PubMed
97. Taylor E. We Agree, Don't We? The Delphi Method for Health Environments Research. *HERD.* 2020;13(1):11-23. PubMed
98. Niederberger M, Spranger J. Delphi Technique in Health Sciences: A Map. *Front Public Health.* 2020;8:457. PubMed
99. Hsu C, Sandford B. The Delphi technique: making sense of consensus—Practical assessment, research & evaluation. *Pract Assessment. Res Eval.* 2007;12.
100. VanSpronsen AD, Villatoro V, Zychla L, et al. A New List for Choosing Wisely Canada From the "Hidden Profession" of Medical Laboratory Science. *Am J Clin Pathol.* 2022;157(1):40-46. PubMed
101. Nasa P, Jain R, Juneja D. Delphi methodology in healthcare research: How to decide its appropriateness. *World J Methodol.* 2021;11(4):116-129. PubMed
102. Murphy MK, Black NA, Lamping DL, et al. Consensus development methods, and their use in clinical guideline development. *Health Technol Assess.* 1998;2(3):i-iv, 1-88. PubMed

Appendix 1: Additional Tables and Figures

Note that this appendix has not been copy-edited.

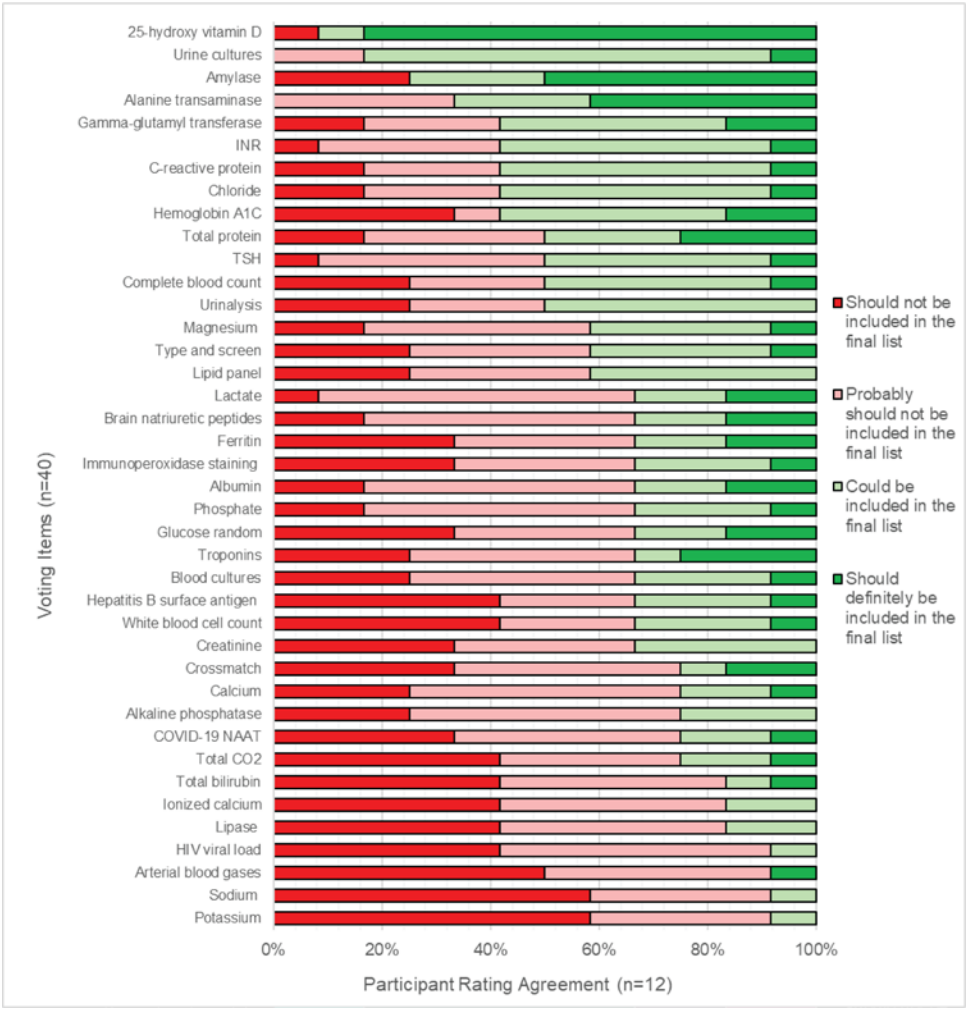
Table 2: Characteristics of the Subject Experts Involved in the Modified Delphi Process

Panellist characteristics	Delphi round		
	Survey 1 (N = 12), n (%)	Workshop (N = 15), n (%)	Survey 2 (N = 16), n (%)
Province/Territory			
Alberta	2 (16.7%)	1 (6.7%)	2 (12.5%)
British Columbia	3 (25.0%)	4 (26.7%)	4 (25.0%)
Manitoba	0 (0%)	0 (0%)	0 (0%)
New Brunswick	0 (0%)	0 (0%)	0 (0%)
Newfoundland and Labrador	1 (8.3%)	1 (6.7%)	1 (6.3%)
Northwest Territories	0 (0%)	0 (0%)	0 (0%)
Nova Scotia	2 (16.7%)	3 (20.0%)	3 (18.8%)
Nunavut	0 (0%)	0 (0%)	0 (0%)
Ontario	3 (25.0%)	4 (26.7%)	4 (25.0%)
Prince Edward Island	0 (0%)	0 (0%)	0 (0%)
Quebec	1 (8.3%)	1 (6.7%)	1 (6.3%)
Saskatchewan	0 (0%)	1 (6.7%)	1 (6.3%)
Yukon	0 (0%)	0 (0%)	0 (0%)
Years of experience			
0 to 9 years	4 (33.3%)	6 (40.0%)	6 (37.5%)
10 or more years	8 (66.7%)	9 (60.0%)	10 (62.5%)
Stakeholder perspective[a]			
Medical laboratory professional	8 (66.7%)	10 (66.7%)	10 (62.5%)
Physician	9 (75.0%)	9 (60.0%)	9 (56.3%)
Decision-maker	8 (66.7%)	8 (53.3%)	9 (56.3%)

Note: Patients and caregivers (n = 2) participated in the workshop round where they shared their views and experiences related to lab testing. Their insight contributed to the discussions, but they did not participate in either of the Delphi survey rounds.

[a]Panellists were able to select multiple responses.

Figure 4: Delphi Survey 1 Results



CO_2 = carbon dioxide; INR = international normalized ratio; NAAT = nucleic acid amplification test; TSH = thyroid stimulating hormone.

Note: Lab tests are ordered from top to bottom by their level of consensus to include in the final list of unnecessary lab tests.

Table 3: Delphi Survey 1 Voting Results

Lab test	Number of expert panel members assigning each rank (%)				Median response (IQR)[a]
	Should not be included in the final list	Probably should not be included in the final list	Could be included in the final list	Should definitely be included in the final list	
40 highly used lab tests					
Creatinine	4 (33.33%)	4 (33.33%)	4 (33.33%)	0 (0%)	2 (2)
Potassium	7 (58.33%)	4 (33.33%)	1 (8.33%)	0 (0%)	1 (1)
Sodium	7 (58.33%)	4 (33.33%)	1 (8.33%)	0 (0%)	1 (1)
Complete blood count	3 (25.00%)	3 (25.00%)	5 (41.67%)	1 (8.33%)	2.5 (1.25)
COVID-19 NAAT	4 (33.33%)	5 (41.67%)	2 (16.67%)	1 (8.33%)	2 (1.25)
Chloride	2 (16.67%)	3 (25.00%)	6 (50.00%)	1 (8.33%)	3 (1)
Total carbon dioxide	5 (41.67%)	4 (33.33%)	2 (16.67%)	1 (8.33%)	2 (1.25)
Random blood glucose	4 (33.33%)	4 (33.33%)	2 (16.67%)	2 (16.67%)	2 (2)
Phosphate	2 (16.67%)	6 (50.00%)	3 (25.00%)	1 (8.33%)	2 (1)
Magnesium	2 (16.67%)	5 (41.67%)	4 (33.33%)	1 (8.33%)	2 (1)
Alanine transaminase	0 (0%)	4 (33.33%)	3 (25.00%)	5 (41.67%)	3 (2)
Calcium	3 (25.00%)	6 (50.00%)	2 (16.67%)	1 (8.33%)	2 (0.5)
Total bilirubin	5 (41.67%)	5 (41.67%)	1 (8.33%)	1 (8.33%)	2 (1)
Alkaline phosphatase	3 (25.00%)	6 (50.00%)	3 (25.00%)	0 (0%)	2 (0.5)
INR	1 (8.33%)	4 (33.33%)	6 (50.00%)	1 (8.33%)	3 (1)
Albumin	2 (16.67%)	6 (50.00%)	2 (16.67%)	2 (16.67%)	2 (1)
Troponins	3 (25.00%)	5 (41.67%)	1 (8.33%)	3 (25.00%)	2 (1.5)
Arterial blood gases	6 (50.00%)	5 (41.67%)	0 (0%)	1 (8.33%)	1.5 (1)
Gamma-glutamyl transferase	2 (16.67%)	3 (25.00%)	5 (41.67%)	2 (16.67%)	3 (1)
Hemoglobin A1C	4 (33.33%)	1 (8.33%)	5 (41.67%)	2 (16.67%)	3 (2)
Urinalysis	3 (25.00%)	3 (25.00%)	6 (50.00%)	0 (0%)	2.5 (1.25)
C-reactive protein	2 (16.67%)	3 (25.00%)	6 (50.00%)	1 (8.33%)	3 (1)
Type and screen (ABO group, Rh group, and antibody screen)	3 (25.00%)	4 (33.33%)	4 (33.33%)	1 (8.33%)	2 (1.25)
Lactate	1 (8.33%)	7 (58.33%)	2 (16.67%)	2 (16.67%)	2 (1)
TSH	1 (8.33%)	5 (41.67%)	5 (41.67%)	1 (8.33%)	2.5 (1)
Amylase	3 (25.00%)	0 (0%)	3 (25.00%)	6 (50.00%)	3.5 (1.5)
Hepatitis B surface antigen	5 (41.67%)	3 (25.00%)	3 (25.00%)	1 (8.33%)	2 (2)

Lab test	Number of expert panel members assigning each rank (%)				Median response (IQR)[a]
	Should not be included in the final list	Probably should not be included in the final list	Could be included in the final list	Should definitely be included in the final list	
Lipid panel	3 (25.00%)	4 (33.33%)	5 (41.67%)	0 (0%)	2 (1.25)
Ionized calcium	5 (41.67%)	5 (41.67%)	2 (16.67%)	0 (0%)	2 (1)
White blood cell count	5 (41.67%)	3 (25.00%)	3 (25.00%)	1 (8.33%)	2 (2)
Ferritin	4 (33.33%)	4 (33.33%)	2 (16.67%)	2 (16.67%)	2 (2)
Urine cultures	0 (0%)	2 (16.67%)	9 (75.00%)	1 (8.33%)	3 (0)
Blood cultures	3 (25.00%)	5 (41.67%)	3 (25.00%)	1 (8.33%)	2 (1.25)
Immunoperoxidase staining	4 (33.33%)	4 (33.33%)	3 (25.00%)	1 (8.33%)	2 (2)
Lipase	5 (41.67%)	5 (41.67%)	2 (16.67%)	0 (0%)	2 (1)
Crossmatch	4 (33.33%)	5 (41.67%)	1 (8.33%)	2 (16.67%)	2 (1.25)
Total protein	2 (16.67%)	4 (33.33%)	3 (25.00%)	3 (25.00%)	2.5 (1.25)
Brain natriuretic peptides	2 (16.67%)	6 (50.00%)	2 (16.67%)	2 (16.67%)	2 (1)
25-hydroxy vitamin D	1 (8.33%)	0 (0%)	1 (8.33%)	10 (83.33%)	4 (0)
HIV viral load	5 (41.67%)	6 (50.00%)	1 (8.33%)	0 (0%)	2 (1)

INR = international normalized ratio; IQR = interquartile range; NAAT = nucleic acid amplification test; TSH = thyroid stimulating hormone.

[a]Respondents ranked each item using a Likert scale from 1 to 4, with 1 (should not be considered for inclusion in the final list), 2 (probably should not be included), 3 (could be included), or 4 (should definitely be included in the final list).

Table 4: Delphi Survey 2 Voting Results

Lab test	Number of expert panel members assigning each rank (%)				Median response (IQR)[a]
	Should not be included in the final list	Probably should not be included in the final list	Could be included in the final list	Should definitely be included in the final list	
40 highly used lab tests					
Creatinine	5 (33.33%)	6 (40.00%)	4 (26.67%)	0 (0%)	2 (2)
Potassium	10 (66.67%)	4 (26.67%)	1 (6.67%)	0 (0%)	1 (1)
Sodium	10 (66.67%)	4 (26.67%)	1 (6.67%)	0 (0%)	1 (1)
Complete blood count	2 (13.33%)	4 (26.67%)	8 (53.33%)	1 (6.67%)	3 (1)
COVID-19 NAAT	8 (53.33%)	6 (40.00%)	1 (6.67%)	0 (0%)	2 (1)
Chloride	3 (20.00%)	2 (13.33%)	8 (53.33%)	2 (13.33%)	3 (1)
Total carbon dioxide	4 (26.67%)	8 (53.33%)	3 (20.00%)	0 (0%)	2 (1)
Random blood glucose	5 (33.33%)	5 (33.33%)	2 (13.33%)	3 (20.00%)	2 (2)
Phosphate	6 (40.00%)	4 (26.67%)	4 (26.67%)	1 (6.67%)	2 (2)
Magnesium	5 (33.33%)	4 (26.67%)	5 (33.33%)	1 (6.67%)	2 (2)

Lab test	Number of expert panel members assigning each rank (%)				Median response (IQR)[a]
	Should not be included in the final list	Probably should not be included in the final list	Could be included in the final list	Should definitely be included in the final list	
Alanine transaminase	4 (26.67%)	6 (40.00%)	3 (20.00%)	2 (13.33%)	2 (2)
Calcium	4 (26.67%)	7 (46.67%)	4 (26.67%)	0 (0%)	2 (2)
Total bilirubin	4 (26.67%)	10 (66.67%)	1 (6.67%)	0 (0%)	2 (1)
Alkaline phosphatase	4 (26.67%)	9 (60.00%)	1 (6.67%)	1 (6.67%)	2 (1)
INR	1 (6.67%)	1 (6.67%)	11 (73.33%)	2 (13.33%)	3 (0)
Albumin	2 (13.33%)	5 (33.33%)	7 (46.67%)	1 (6.67%)	3 (1)
Troponins	3 (20.00%)	8 (53.33%)	3 (20.00%)	1 (6.67%)	2 (1)
Arterial blood gases	7 (46.67%)	6 (40.00%)	1 (6.67%)	1 (6.67%)	2 (1)
Gamma-glutamyl transferase	3 (20.00%)	1 (6.67%)	9 (60.00%)	2 (13.33%)	3 (1)
Hemoglobin A1C	1 (6.67%)	4 (26.67%)	4 (26.67%)	6 (40.00%)	3 (2)
Urinalysis	3 (20.00%)	1 (6.67%)	6 (40.00%)	5 (33.33%)	3 (1.5)
C-reactive protein	2 (13.33%)	2 (13.33%)	7 (46.67%)	4 (26.67%)	3 (2)
Type and screen (ABO group, Rh group, and antibody screen)	2 (13.33%)	3 (20.00%)	5 (33.33%)	5 (33.33%)	3 (2)
Lactate	1 (6.67%)	10 (66.67%)	2 (13.33%)	2 (13.33%)	2 (1)
TSH	0 (0%)	3 (20.00%)	6 (40.00%)	6 (40.00%)	3 (1)
Amylase	2 (13.33%)	1 (6.67%)	2 (13.33%)	10 (66.67%)	4 (1)
Hepatitis B surface antigen	4 (26.67%)	7 (46.67%)	3 (20.00%)	1 (6.67%)	2 (2)
Lipid panel	3 (20.00%)	3 (20.00%)	5 (33.33%)	4 (26.67%)	3 (2)
Ionized calcium	5 (33.33%)	7 (46.67%)	2 (13.33%)	1 (6.67%)	2 (1)
White blood cell count	4 (26.67%)	6 (40.00%)	5 (33.33%)	0 (0%)	2 (2)
Ferritin	1 (6.67%)	8 (53.33%)	5 (33.33%)	1 (6.67%)	2 (1)
Urine cultures	2 (13.33%)	2 (13.33%)	6 (40.00%)	5 (33.33%)	3 (2)
Blood cultures	4 (26.67%)	7 (46.67%)	3 (20.00%)	1 (6.67%)	2 (2)
Immunoperoxidase staining	7 (46.67%)	6 (40.00%)	1 (6.67%)	1 (6.67%)	2 (1)
Lipase	8 (53.33%)	6 (40.00%)	1 (6.67%)	0 (0%)	1 (1)
Crossmatch	3 (20.00%)	4 (26.67%)	5 (33.33%)	3 (20.00%)	3 (1)
Total protein	3 (20.00%)	7 (46.67%)	4 (26.67%)	1 (6.67%)	2 (1)
Brain natriuretic peptides	2 (13.33%)	4 (26.67%)	6 (40.00%)	3 (20.00%)	3 (1)

Lab test	Number of expert panel members assigning each rank (%)				Median response (IQR)[a]
	Should not be included in the final list	Probably should not be included in the final list	Could be included in the final list	Should definitely be included in the final list	
25-hydroxy vitamin D	0 (0%)	0 (0%)	5 (33.33%)	10 (66.67%)	4 (1)
HIV viral load	4 (26.67%)	10 (66.67%)	1 (6.67%)	0 (0%)	2 (1)
Tests nominated for consideration by respondents of survey 1					
Vitamin B12	1 (6.67%)	1 (6.67%)	6 (40.00%)	7 (46.67%)	3 (1)
Venous blood gases	2 (13.33%)	5 (33.33%)	7 (46.67%)	1 (6.67%)	3 (1)
Anti-nuclear antibodies	0 (0%)	5 (33.33%)	3 (20.00%)	7 (46.67%)	3 (2)
PE of placentas	3 (20.00%)	6 (40.00%)	2 (13.33%)	4 (26.67%)	2 (2)
Serum protein electrophoresis	1 (6.67%)	3 (20.00%)	5 (33.33%)	6 (40.00%)	3 (2)
Lactate dehydrogenase	2 (13.33%)	4 (26.67%)	7 (46.67%)	2 (13.33%)	3 (1)
PE of orthopedic devices	2 (13.33%)	5 (33.33%)	3 (20.00%)	5 (33.33%)	3 (2)
Bacterial vaginosis	5 (33.33%)	9 (60.00%)	1 (6.67%)	0 (0%)	2 (1)
Iron panel	3 (20.00%)	3 (20.00%)	6 (40.00%)	3 (20.00%)	3 (1)
Fecal occult blood	2 (13.33%)	2 (13.33%)	5 (33.33%)	6 (40.00%)	3 (2)
PE of skin or fatty tissue	5 (33.33%)	6 (40.00%)	1 (6.67%)	3 (20.00%)	2 (2)
Corrected calcium	3 (20.00%)	3 (20.00%)	5 (33.33%)	4 (26.67%)	3 (2)
Wound swabs	3 (20.00%)	7 (46.67%)	3 (20.00%)	2 (13.33%)	2 (1)
Gastrointestinal tract biopsies	4 (26.67%)	9 (60.00%)	1 (6.67%)	1 (6.67%)	2 (1)
Prostate-specific antigen	3 (20.00%)	6 (40.00%)	4 (26.67%)	2 (13.33%)	2 (1)
FT3 and FT4	0 (0%)	2 (13.33%)	7 (46.67%)	6 (40.00%)	3 (1)

FT3 = free triiodothyronine; FT4 = free thyroxine; INR = international normalized ratio; IQR = interquartile range; NAAT = nucleic acid amplification test; PE = pathologic examination; TSH = thyroid stimulating hormone.

[a]Respondents ranked each item using a Likert scale from 1 to 4, with 1 (should not be considered for inclusion in the final list), 2 (probably should not be included), 3 (could be included), or 4 (should definitely be included in the final list).