

1 Developing an evaluation framework for public health
2 environmental surveillance: Protocol for an international,
3 multidisciplinary e-Delphi study
4 Public Health Environmental Surveillance Evaluation Framework (PHES-
5 EF)

6 Douglas G. Manuel^{1,2,3,4*} (0000-0003-0912-0845), Carol Bennett^{1,2} (0000-0002-3033-9860), Emma Brown¹
7 (0000-0002-0869-9907), David L. Buckeridge⁵ (0000-0003-1817-5047), Yoni Freedhoff² (0000-0001-
8 9391-3455), Sarah Funnell⁶, Farah Ishtiaq⁷ (0000-0002-6762-7014), Matthew J. Wade^{8,9} (0000-0001-
9 9824-7121), David Moher^{3,10} (0000-0003-2434-4206), on behalf of the PHES-EF Executive Group[^]

10 ¹ Methodological and Implementation Research Program, Ottawa Hospital Research Institute, Ottawa,
11 Ontario, Canada

12 ² Department of Family Medicine, University of Ottawa, Ottawa, Ontario, Canada

13 ³ School of Epidemiology and Public Health, Faculty of Medicine, University of Ottawa, Ottawa, Ontario,
14 Canada

15 ⁴ C.T. Lamont Primary Health Care Research Centre Program, Bruyère Research Institute, Ottawa,
16 Ontario, Canada

17 ⁵ School of Population and Global Health, McGill University, Montreal, Quebec, Canada

18 ⁶ Department of Family Medicine, Faculty of Health Sciences, Queen's University, Kingston, Ontario,

19 Canada

20 ⁷Tata Institute for Genetics and Society, National Centre for Biological Sciences, Bengaluru,

21 Karnataka, India

22 ⁸ Data Analytics and Surveillance Group, UK Health Security Agency, London, United Kingdom

23 ⁹ School of Engineering, Newcastle University, Newcastle-upon-Tyne, United Kingdom

24 ¹⁰ Centre for Journalology, Methodological and Implementation Research Program, Ottawa Hospital

25 Research Institute, Ottawa, Ontario, Canada

26 *Corresponding Author

27 Email: dmanuel@ohri.ca (DGM)

28 ^Membership of the PHES-EF Executive Group is provided in the Acknowledgments.

29 Metadata

30 Funding

31 This project is funded by the Canadian Institutes of Health Research-funded network, CoVaRR-Net

32 (Coronavirus Variants Rapid Response Network. FRN: 175622), and Health Canada (through the Safe

33 Restart Agreement Contribution Program. Arrangement #: 2223-HQ-000098). The Canadian Institutes of

34 Health Research and Health Canada have not been involved in the design or conduct of the study and

35 the views expressed herein do not necessarily represent the views of either funding organization.

36 Competing interests

37 The authors have declared that no competing interests exist.

38 Data availability

39 Anonymous aggregate data will be made publicly available when the study is completed and published.

40 Abstract

41 Introduction

42 Public health environmental surveillance has evolved, especially during the coronavirus disease
43 pandemic, with wastewater-based surveillance being a prominent example. As surveillance methods
44 diversify and expand, it is essential to have a robust evaluation of surveillance systems. This electronic
45 Delphi study will propose an evaluation framework for public health environmental surveillance,
46 informed by the expanding practice of wastewater-based surveillance during the pandemic.

47 Methods

48 The Public Health Environmental Surveillance Evaluation Framework (PHES-EF) will be developed in five
49 steps. In Step 1, a multinational and multidisciplinary executive group will be formed to guide the
50 framework development process. In Step 2, candidate items for Round 1 of the electronic Delphi will be
51 generated by conducting relevant scoping reviews and consultation with the study executive group. In
52 Step 3, an international electronic Delphi will be conducted over two rounds to develop consensus on
53 items for the framework. In Step 4, the executive group will reconvene to finalize the evaluation
54 framework, discuss standout items, and determine the dissemination strategies. Lastly, Step 5 will focus
55 on disseminating the evaluation framework to all parties involved with or affected by wastewater-based
56 surveillance using traditional and public-oriented methods.

57 Discussion

58 The electronic Delphi study will provide multidisciplinary and multinational consensus for the evaluation
59 framework for public health environmental surveillance by providing a set of minimum criteria required

60 for their evaluation. The evaluation framework is intended to support the sustainability of
61 environmental surveillance and improve its implementation, reliability, credibility, and value.

62 Introduction

63 Public health relies on robust surveillance systems to support informed decision-making and timely
64 interventions [1]. Environmental-based surveillance, particularly wastewater-based testing, has
65 historically informed public health responses to threats like typhoid, polio, and illicit drug use [2–6]. The
66 coronavirus disease (Covid-19) pandemic marked a turning point, with an expansion of wastewater-
67 based monitoring and surveillance to an unprecedented scale, highlighting its potential in broader public
68 health contexts [7].

69 Building on the advances and experiences from Covid-19 wastewater-based surveillance (wastewater
70 surveillance) during the pandemic, this study aims to develop a comprehensive evaluation framework.
71 Wastewater surveillance will be an example; however, it will apply to broader public health
72 environmental surveillance systems, including air and surface monitoring. Surveillance targets
73 considered include infectious agents and health risks beyond Covid-19.

74 Wastewater surveillance, once a peripheral tool, gained prominence during the pandemic. Its global
75 implementation in diverse settings [8], from metropolitan to rural areas [9,10], highlighted its strengths.
76 These strengths include the provision of insights into disease spread, unaffected by clinical testing
77 access or population health-seeking behaviours, and offering a non-invasive, resource-efficient
78 surveillance method [11,12].

79 Despite these strengths, the widespread application of wastewater surveillance faces challenges
80 [5,13,14]. Environmental variability affects measurement consistency, and the need for multidisciplinary
81 collaboration brings logistic complexities [14]. Rapidly evolving methodologies have led to various
82 analytic techniques and interpretation approaches, resulting in inconsistencies and uncertainties [7].

83 These challenges underscore the need for a structured evaluation framework. Such a framework can
84 enhance accountability, transparency, and trust in wastewater surveillance systems, potentially
85 improving surveillance performance and identifying best practices [1].

86 Study objectives

87 This project will use a multidisciplinary and multinational consensus approach to develop an evaluation
88 framework for public health environmental surveillance (**Public Health Environmental Surveillance**
89 **Evaluation Framework**, PHES-EF). An evaluation framework provides structured guidance to facilitate a
90 systematic approach to the evaluation of the program [15–17]. See below for the study’s definition of an
91 evaluation framework.

92 Specific objectives for the framework development process are as follows: (1) engage a multidisciplinary
93 executive group of international experts to inform the development of the evaluation framework; (2)
94 review the literature to identify current guidance on surveillance evaluation items; (3) prioritize items
95 for an evaluation framework for environmental surveillance using an international electronic-Delphi (e-
96 Delphi) process; (4) conduct a consensus meeting, which engages experts across disciplines, to create an
97 evaluation framework for public health environmental surveillance based on feedback from the
98 international e-Delphi process; and, (5) develop a comprehensive dissemination plan for the evaluation
99 framework.

100 Definitions

101 We define key terms as follows:

- 102 • **Environmental surveillance:** Surveillance of substances or organisms in air, water, soil, and
103 living organisms to detect changes or threats that may affect human health, ecosystems, or both
104 [18].

- 105 • **Evaluation Framework:** A structured system or schema to systematically assess a program
106 [1,15,17]. An evaluation framework presents a clear list of items or elements that should appear
107 in an evaluation, such as constructs or concepts. It provides an organisation or structure for the
108 list of items; defines how the list and organisation were developed; and, potentially, includes
109 details about measuring the evaluation items. Importantly, an evaluation framework facilitates
110 informed decision-making by providing a comprehensive understanding of program
111 effectiveness and areas for improvement.
- 112 • **Public health surveillance:** The ongoing systematic collection, analysis, interpretation and
113 dissemination of health data for the planning, implementation and evaluation of public health
114 interventions and strategies to protect and improve the health of populations [19].
- 115 • **Wastewater:** Water that has been in contact with people (e.g., for washing) or used for
116 cleansing and sanitation (e.g., for flushing away faecal matter), and is discharged via sewers or
117 other sanitation systems [6].
- 118 • **Wastewater-based surveillance (wastewater surveillance):** A public health approach that
119 examines specific substances or organisms in wastewater (sewage). Wastewater surveillance
120 can be referred to as wastewater-based epidemiology; however, the latter can be specific to the
121 use of wastewater to study and analyse the distribution, patterns, and determinants of health
122 and disease conditions in defined populations [19].

123 Materials and methods

124 Study design

125 This protocol was developed *a priori*. The development of PHES-EF is guided by recommendations for
126 Delphi techniques in the health sciences [20]. An overview of the study flow is detailed in Fig 1.

127 **Fig 1. Overview of the development process for PHES-EF.**

128 **Step 1: Establish the executive group**

129 A multinational and multidisciplinary executive group will oversee the development of the evaluation
130 framework. The group will bring expertise from various fields, including public health, infectious disease,
131 epidemiology; environmental and physical sciences; mathematical sciences; social sciences;
132 communication, knowledge translation and exchange; knowledge users; and engaged public.

133 Content expert and knowledge user executive group members will be identified by examining relevant
134 author lists, internet-based searching, and the study working group's professional networks. In contrast,
135 the engaged member of the public will be identified through an open call for expressions of interest. The
136 composition of the study executive group is detailed in S1 Table.

137 The executive group will be responsible for providing feedback on the framework development process
138 (e.g., framework organization, completeness, additions, redundancy), supporting and advising working
139 groups in their area of expertise, connecting the study to other related projects, recruiting e-Delphi
140 panellists, participating in the consensus meeting, disseminating the evaluation framework, advising
141 and/or participating in post-publication activities, and promoting the adoption of the evaluation
142 framework among their professional networks. Co-chairs will also be responsible for decision-making
143 when the executive group cannot reach consensus, leading executive group activities, and the overall
144 framework development process. Any conflicts of interest among the research team, including executive
145 group members, will be assessed before beginning the study.

146 Step 2: Generate candidate evaluation items

147 We will systematically compile a list of e-Delphi candidate items through three processes: scoping
148 reviews, content analysis and extraction, and executive group consultation.

149 Scoping reviews

150 Two scoping reviews will be performed to provide a comprehensive view of public health environmental
151 surveillance system evaluations. The first review will identify frameworks (e.g., sources which guide the
152 planning, management, or conduct of processes) and criteria (e.g., principles on which a judgment may
153 be based) for evaluating public health or environmental surveillance systems. The second review will
154 explore insights from the Covid-19 pandemic, particularly focusing on the principles of equity, trust,
155 open science, community engagement, and One Health in the context of wastewater surveillance [21–
156 23]. Detailed protocols for the scoping reviews are available on OSF
157 (<https://doi.org/10.17605/OSF.IO/GSKF6>; <https://doi.org/10.17605/OSF.IO/C9A3Q>). The completed
158 scoping reviews will be reported following the Preferred Reporting Items for Systematic Reviews and
159 Meta-Analyses extension for scoping reviews (PRISMA-ScR) [24].

160 Content analysis and extraction

161 We will conduct a content analysis on items identified during the data extraction phase for both scoping
162 reviews. The content analyses will identify and group criteria for evaluating public health and
163 environmental surveillance systems. In the first step, one team member will group all the extracted
164 items into initial domains and themes. In the second step, two researchers will independently compare
165 the category system of the first researcher to agree on the final category system consensually. In the
166 third step, all criteria will be reviewed again to merge or modify duplicate concepts and create a
167 parsimonious framework for the initial list of e-Delphi candidate items.

168 Executive group consultation

169 The executive group will review the initial list of e-Delphi candidate items and discuss the initial
170 structure and presentation of Round 1 items. Their review will focus on creating an evidence-informed
171 public health environmental surveillance framework centred around SARS-CoV-2 wastewater
172 surveillance, and incorporating lessons learned from the pandemic (e.g., equity, open science,
173 community engagement, and One Health). Additional items will be added if the executive group
174 determines that key evaluation areas have not been addressed.

175 Step 3: International e-Delphi survey

176 A Delphi technique will be used to develop an internationally accepted evaluation framework for public
177 health environmental surveillance. The Delphi method is an iterative multi-round approach that uses a
178 series of sequential surveys, interspersed by controlled feedback, to elicit consensus among a group of
179 individuals while maintaining anonymity [25,26]. An e-Delphi method will be used to overcome
180 geographic barriers and allow us to engage panellists internationally across various time zones. Delphi
181 surveys relating to health sciences have typically been conducted for 2-3 rounds [20,27,28]. Therefore, a
182 two-round e-Delphi survey and a subsequent consensus meeting will be conducted to balance feasibility
183 with rigour. Each round will be active for approximately 4-6 weeks to allow panellists sufficient time to
184 participate in the survey while balancing study timeline constraints [29].

185 Recruitment of panellists

186 We will recruit a multinational, multidisciplinary panel of wastewater surveillance experts, knowledge
187 users, and engaged members of the public to complete the e-Delphi survey. Content experts are adult
188 individuals with knowledge in their respective subdiscipline groups (see S2 Table). Panellists will be
189 selected to capture the multiple perspectives of those that influence the design, implementation,

190 evaluation, use, and reporting of wastewater surveillance activities, including the following
191 subdisciplines groups: public health, infectious disease, epidemiology; environmental and physical
192 sciences; mathematical sciences; social sciences; and communication, knowledge translation and
193 exchange. Targeted recruitment will also be conducted to identify knowledge users and engaged
194 members of the public. We will use purposive sampling to identify potentially eligible panellists by
195 examining relevant author lists from the scoping reviews, professional networks, and internet-based
196 searching, among other potential methods. Potential panellists may also be recruited using social media,
197 a study website or calls for expressions of interest during relevant meetings, presentations, or
198 correspondence. Snowball sampling may also be employed by executive group members and potential
199 panellists to maximize recruitment.

200 Potentially eligible panellists will receive a recruitment email from the study working group, executive
201 group members, or other potential panellists. A call for panellists to share the e-Delphi with other
202 potentially eligible panellists may be included in recruitment materials and the e-Delphi survey.

203 Inclusion criteria for e-Delphi panellists are outlined in Table 1.

204 **Table 1. Eligibility criteria for e-Delphi survey panellists**

Discipline group	Discipline subgroup	Inclusion criteria
Content experts	Public health, infectious disease, epidemiology Environmental and physical sciences Mathematical sciences	An adult ^a who is proficient in English and has a graduate degree in one of the listed specializations ^b , or ≥ 3 years of professional experience, or ≥ 2 peer-reviewed publications relating to wastewater surveillance.

	Social sciences Communication, knowledge translation and exchange	
Knowledge users		An adult ^a who is proficient in English and who is a professional who does not have specialized training or qualifications in wastewater surveillance, but who uses surveillance to inform policy and action in their workplace.
Engaged public		An adult ^a who is proficient in English and has relevant lived experience.

205 Definitions: (1) Professional experience – Paid employment or professional practice in a listed specialization (current
 206 or former); (2) Graduate or professional degree – Master’s or Doctor of Philosophy (PhD) relating to a listed discipline
 207 subgroup.

208 ^aAdult: ≥ 18 years of age

209 ^bSee S2 Table for specializations associated with each discipline.

210

211 We will conduct an all-rounds e-Delphi invitation approach, meaning panellists will be invited to every
 212 round regardless of whether they participated in the previous round [30]. Previous Delphi studies have
 213 commonly ranged from 20 to 30 panellists; however, there is no agreed upon standard [28,31]. Studies
 214 have also shown that the number of panellists and their diversity can impact the quality of outcomes
 215 [31]. Therefore, assuming only half of registered panellists will participate in at least one round, we will
 216 aim to recruit at least 50 panellists. Preferably, with at least 8-10 people per discipline subgroup. The
 217 study working group will monitor the distribution of registered panellists based on their demographic

218 information and will try to distribute appropriately across discipline subgroups and other demographic
219 markers.

220 Registration to complete the e-Delphi will open approximately four weeks before the first round, and e-
221 Delphi round surveys will be active for roughly 4-6 weeks. The tentative start and end dates for
222 recruitment are 04/11/2024 and 21/02/2025, respectively. Recruitment or retention efforts will occur
223 throughout each e-Delphi round and the four weeks of targeted recruitment before Round 1. Email
224 reminders to complete the survey will be sent twice per e-Delphi round. Panellists may also receive up
225 to two reminders to complete the screening questionnaire during the targeted recruitment period. The
226 time required to complete each e-Delphi round survey will be approximately one hour. The estimated
227 time commitment will be provided to all potential panellists during recruitment and before starting the
228 consensus process.

229 **Panellist withdrawal or termination**

230 Panellists will not be able to withdraw from the study due to the safeguards in place to maintain
231 anonymity; however, they may choose to stop actively participating in the study at any time.

232 **Panellist compensation**

233 Content experts and knowledge users who complete all rounds of the e-Delphi process will receive a
234 \$100 CAD pre-paid Visa card in compensation for the time required to complete the e-Delphi process.
235 Engaged members of the public who complete all rounds will be entered to win one of five \$100 CAD
236 pre-paid Visa cards. Remuneration has been shown to encourage participation and retention from
237 diverse groups, thereby minimizing bias, and motivating thoughtful, high-quality responses, thereby
238 improving data quality [32].

239 Procedure

240 We will conduct a two-round e-Delphi survey to generate consensus on evaluation criteria (see Fig 2).

241 The full survey will be pre-tested and validated before administration. Summaries of Round 1 will be

242 compiled for the subsequent round.

243 **Fig 2. E-Delphi and consensus meeting methodology, methods to achieve consensus on core items for**

244 **PHES-EF.** Consensus: if $\geq 70\%$ of panellist votes fall within the same category (irrelevant, equivocal,

245 relevant); no consensus if $<70\%$ agreement. Item relevancy: *core item*: highly or moderately

246 relevant; *further discussion*: slightly relevant, neutral, slightly irrelevant; *excluded item*: moderately or

247 highly irrelevant.

248 We will use the Surveylet Delphi platform to administer the survey [33]. The platform facilitates

249 interactive dialogue between panellists by allowing them to justify their responses and comment on the

250 responses of others. Panellists can also save their progress and answer survey questions across multiple

251 sessions. Custom survey pathways will be generated for each panellist discipline subgroup (i.e.,

252 panellists from different discipline subgroups will be shown a different collection of candidate items).

253 Within each custom discipline subgroup survey pathway, panellists can skip items or self-declare that

254 they are not qualified to assess certain candidate items. To ensure the independent evaluation of items,

255 the order of candidate items may be randomized for each panellist.

256 Individuals will complete the screening questionnaire separate from the e-Delphi platform using

257 LimeSurvey [34]. The study working group will review eligibility, and if eligible to participate in the study,

258 individuals will receive an individualized link for anonymous participation in the e-Delphi survey. Implied

259 consent will be obtained electronically for all panellists using the Surveylet Delphi platform before

260 participation in the survey. Panellists who are eligible for compensation or wish to receive

261 acknowledgement for their contributions will be able to do so by clicking a link at the end of each survey

262 round. The link will bring them to a separate LimeSurvey form that will not be linked to their survey
263 responses, allowing for e-Delphi responses to remain anonymous.

264 Defining consensus

265 We will use a 7-point Likert scale (1 = highly irrelevant to 7 = highly relevant). Panellist responses will be
266 categorised as excluded item (irrelevant: 1–2), further discussion (equivocal: 3–5), or core item
267 (relevant: 6–7). Consensus for each item is defined as $\geq 70\%$ of the panellist votes falling within the
268 same category (1–2, 3–5, or 6–7).

269 The approach to use a 7-point fully-labeled Likert scale – a higher number of categories compared to
270 many studies – is informed by the consideration that panellists are professionals and engaged members
271 of the public with good cognitive skills; therefore, more categories and labels will be more discriminating
272 and reproducible [35–38]. The $\geq 70\%$ consensus cut-off – a lower cut-off compared to many studies – is
273 informed by the rapidly evolving nature of environmental surveillance for public health and the wide
274 range of disciplines invited to the panel; therefore, a high level of agreement may not occur [35–38].

275 Stopping rules

276 When consensus is reached for a given item, no subsequent rounds of ranking for that item will be
277 performed. Consensus items that are considered equivocal, in addition to remaining non-consensual
278 items, will be further deliberated during the consensus meeting.

279 Round 1

280 Panellists will be invited to rate their level of agreement with candidate items generated from scoping
281 review results and consultation with the study executive group. Free-text boxes will be included for
282 panellists to provide feedback or identify additional candidate items to be included in the next e-Delphi
283 survey round.

284 Round 2

285 Regardless of whether they participated in the previous round, panellists will be invited to participate in
286 Round 2 of the e-Delphi survey. All panelists will be invited to rate new items and re-rate previous items
287 that did not reach consensus. When re-rating their level of agreement, panellists will be presented with
288 their previous round scores alongside the aggregate group results. Anonymous feedback from Round 1
289 will also be compiled and presented during Round 2. Any newly suggested items during Round 2 will be
290 deliberated on during the consensus meeting.

291 Summary panelist characteristics

292 At the start of the e-Delphi survey, panellists will provide basic demographic details to ascertain their
293 discipline subgroup, ensuring they receive the correct list of candidate items. To maintain anonymity,
294 this demographic data will not contain identifiable information, and only aggregate responses of 5 or
295 more panelists will be analysed. We will report aggregate panel characteristics based on gender,
296 discipline group or subgroup (depending on the number of panelists), level and type of experience,
297 country income level, and World Bank work regions.

298 Quantitative analysis

299 We will report response rates, agreement levels, consensus, and stability across rounds using median,
300 range, and IQR for each criterion item. The proportion of panelists who skipped items or selected “not
301 qualified to respond” will be reported. Fisher’s exact test will be used to evaluate differences in item
302 ratings by panellist characteristics: primary panellist disciplines (see S2 Table), country income level
303 (high-income versus low-and-middle income), and World Bank world regions (see S3 Table).

304 Qualitative analysis

305 Content analyses will be conducted on panelists' feedback to extract common themes and insights. A
306 summary will describe the evolution of each criterion across rounds, including reasons for item
307 modification or elimination.

308 Grouping of subgroups

309 Subgroups with fewer than 5 panellists will be combined with similar groups based on panellist numbers
310 and group likeness. Decisions on combined subgroups will be made prior to analyses of the items.

311 Exclusion of data

312 All panellists that participate in at least one round will have their responses included in analyses unless
313 the study working group determines that a panellist's response warrants being excluded from analyses
314 for reasons that are newly developed or not previously recognised.

315 Step 4: Consensus meeting

316 In preparation for the consensus meeting, the study working group will categorize the remaining
317 candidate items for consideration into a preliminary evaluation framework. This will include merging
318 and/or modifying the remaining items based on gathered feedback from the e-Delphi rounds.

319 An in-person or hybrid (in-person and online) consensus meeting will be held using the Surveylet Delphi
320 survey platform after the results of the e-Delphi have been compiled and analysed. If in a hybrid format,
321 Microsoft Teams will be used for messaging and videoconferencing. Video recording and transcription
322 services may also be used to aid notetaking. The primary objective will be to achieve expert consensus
323 on the final list of items for PHES-EF, through review and discussion of salient items. This process will be
324 guided by the empirical evidence that was identified during the scoping reviews, and the opinions

325 gathered during the e-Delphi process. The secondary objective will be to discuss publication and
326 dissemination strategies for the final evaluation framework.

327 Steps to produce the final list of items are as follows: (i) present the results of the e-Delphi exercise
328 (name, rationale, and score of each item); (ii) discuss the rationale and relevance for including the items
329 in the framework; and (iii) vote on equivocal and non-consensual items. We will invite all members of
330 the executive group to participate in the consensus meeting. PHES-EF will be developed based on the
331 final list of items that receive consensus during the executive group consensus meeting.

332 **Step 5: Dissemination of evaluation framework**

333 The dissemination of this evaluation framework will start with the publication of the PHES-EF protocol.
334 The development of the evaluation framework will be reported in a statement document that will
335 include the rationale and a brief description of the meeting and the panellists involved. While the
336 publication strategy will be finalized based on the consensus meeting discussion, the preliminary
337 approach will be to post to OSF, as a preprint, before publication in an open-access peer-reviewed
338 journal. During this time, we will actively seek feedback from experts, knowledge users, and the public
339 within our networks and via social media. PHES-EF will be made available using an open license (CC-BY-
340 SA-4.0 license).

341 We will develop a website and may create a Discourse server or explore other social media platforms for
342 ongoing user interactions. Active dissemination approaches will include presenting at relevant scientific
343 conferences, holding webinars, and conducting workshops. Additional dissemination practices will be
344 determined during the consensus meeting.

345 Equity, diversity, and inclusion (EDI)

346 “Equity is defined as the removal of systemic barriers and biases enabling all individuals to have equal
347 opportunity to access and benefit from the program” [39]. To achieve this, members of the research
348 team are committed to developing a strong understanding of the systemic barriers faced by individuals
349 from underrepresented groups. “Diversity is defined as differences in race, colour, place of origin,
350 religion, immigrant and newcomer status, ethnic origin, ability, sex, sexual orientation, gender identity,
351 gender expression and age” [39]. Panellists will be asked to provide characteristic information (e.g.,
352 country of residence, gender, primary discipline) to assess and help ensure (through more targeted
353 recruitment) a large diversity in perspectives. “Inclusion is defined as the practice of ensuring that all
354 individuals are valued and respected for their contributions and are equally supported” [39]. To help
355 ensure that all research team members are integrated and supported, we will have a code of conduct
356 statement in the executive group Terms of Reference.

357 All study working group members will complete EDI-related Canadian Institutes of Health Research
358 (CIHR) training [40], and the First Nations Principles of OCAP® (Ownership, Control, Access, and
359 Possession) training for research [41]. The executive group will receive an overview of EDI and OCAP®
360 and will be encouraged and supported to complete such training. Furthermore, staff trained in EDI and
361 OCAP® principles will review the framework from an EDI perspective to ensure appropriate language.

362 Knowledge user and public involvement

363 The international e-Delphi survey will collect responses from content experts, knowledge users, and
364 engaged members of the public to ensure that all parties involved with or affected by wastewater
365 surveillance are represented in the consensus-building approach. Apart from what has already been
366 described, more detailed recruitment or public engagement methods may be informed by the executive

367 group engaged member of the public. Additionally, to ensure accurate and transparent reporting of
368 knowledge user and public involvement throughout the study, we will refer to the Guidance for
369 Reporting Involvement of Patients and the Public (GRIPP2) checklist [42]. We will document the
370 methods used to engage knowledge users and members of the public, report the impacts and outcomes
371 of their engagement, and report on lessons learned from the experience.

372 Acknowledgements

373 E-Delphi participation will not satisfy authorship criteria; however, panellists will have the option to
374 receive individual acknowledgement for their contributions to the framework, given they provide
375 consent to be named. Individuals who participate in the consensus meeting will have the opportunity to
376 be named as authors.

377 Research data management

378 The project will follow an open-science approach and make available the data, code, and materials for
379 all project stages, including the search strategy, search findings, initial curation of evaluation measures,
380 summary of executive group discussions, e-Delphi processes, and agreement procedures. The
381 management of research data will follow a data management and sharing plan.

382 All personal data will be managed using LimeSurvey. Personal data will only be collected from
383 individuals during the screening questionnaire, and from panellists who wish to receive compensation
384 and/or individual acknowledgement for their contributions to the evaluation framework. LimeSurvey
385 data management practices are overseen by The Ottawa Hospital (TOH) Research Institute Methods
386 Centre that protects privacy and confidentiality while assisting authorized users in collecting and
387 managing survey data. LimeSurvey is a free, open-source, online survey tool. It has been installed on a
388 TOH server in the TOH Demilitarized Zone and all data collected is stored in a TOH database within the

389 TOH firewall. LimeSurvey provides a streamlined process to build and manage online surveys. Users can
390 create and publish surveys, collect responses, view analytics, and export survey data for further
391 analysis. Because LimeSurvey is installed outside of the TOH firewall in the TOH Demilitarized Zone,
392 surveys can be responded to from anywhere using an internet connection.

393 Calibrum's Surveylet Delphi platform will be used to manage e-Delphi responses. It may also be used for
394 the consensus meeting. The e-Delphi will only collect anonymous data using a modified Delphi
395 approach; meanwhile, the consensus meeting may collect identifiable data using a real-time Delphi
396 approach. Surveylet complies with the EU-U.S. Privacy Shield Framework and Swiss-U.S. Privacy Shield
397 Framework regarding the collection, use, and retention of personal information. The Surveylet platform
398 also allows its users to be General Data Protection Regulation (GDPR) compliant. Calibrum's servers are
399 protected by high-end firewall systems and vulnerability scans are conducted regularly. Calibrum uses
400 Transport Layer Security (TLS) encryption (also known as HTTPS) for all transmitted data. Surveys are
401 also protected with passwords and HTTP referrer checking. All data are encrypted at rest and hosted by
402 SSAE-16 SOC II certified third-party data centres. Data on deprecated hard drives are destroyed by U.S.
403 DOD methods and delivered to a third-party data destruction service.

404 [Study record retention](#)

405 Any data that is exported from either survey company will be stored securely in a password protected
406 location within TOH's SharePoint. Upon study completion, all study records will be retained securely for
407 10 years, then destroyed as per the Ottawa Hospital Research Institute (OHRI) requirements.

408 [Ethical considerations](#)

409 This study has received ethical approval (20230428-01H) from the Ottawa Health Science Network
410 Research Ethics Board (OHSN-REB). Implied consent will be obtained from all panellists prior to

411 participating in the e-Delphi survey. Information about the survey and consent practices will be provided
412 on the first page of the survey. Panellists will be informed that by providing their demographic
413 information and proceeding to the next page, they are providing their implied consent to participate in
414 the study.

415 Only anonymous panellist responses will be used for analyses. All panellist data and responses will be
416 exported from relevant survey platforms to a shared folder on the Ottawa Hospital SharePoint. The
417 shared folder contents will be password protected and only REB-approved study working group
418 members from the Ottawa Hospital will have access to the data.

419 Discussion

420 This paper outlines the protocol for a two-round e-Delphi study that will be used to develop a
421 multinational, multidisciplinary evaluation framework for public health environmental surveillance.
422 Developing an evaluation framework will contribute to a better understanding of the ideal scope and
423 use cases for wastewater surveillance and improve the reliability and credibility of surveillance data for
424 public health officials. An evaluation framework will also help promote the sustainability of wastewater
425 surveillance systems by providing an effective performance measurement tool.

426 We strove to design the PHES-EF protocol using a rigorous consensus approach. The development of the
427 PHES-EF study protocol was informed by ACCORD (ACcurate COnsensus Reporting Document): A
428 reporting guideline for consensus methods in biomedicine developed via a modified Delphi [43].
429 Surveillance evaluation frameworks are uncommonly developed using a consensus approach. PHES-EF
430 will be developed for environmental surveillance and informed by wastewater surveillance during the
431 pandemic. However, associated scoping reviews and prior surveillance evaluation frameworks are

432 related largely to any public health surveillance system [44]. Therefore, PHES-EF may be applicable for
433 evaluating public health surveillance beyond environmental systems.

434 Limitations

435 Although the Delphi technique uses a systematic process to generate consensus among a panel of
436 experts, the method is not without its limitations. One of the primary concerns when using this
437 technique is the ability to implement a truly representative expert panel. Despite our plan to use various
438 recruitment methods to enroll panellists from all discipline subgroups with varying demographic
439 characteristics, consensus results may ultimately be influenced by the final panel composition [45].

440 Another potential limitation in using a modified Delphi approach is the potential influence of controlled
441 feedback on the convergence of panellist opinions. Although providing controlled feedback between
442 survey rounds aids in generating consensus, this approach may bias consensus results by influencing
443 individual panellists' opinions [46].

444 Furthermore, due to feasibility and resource limitations, only sources that are available in English will be
445 included in associated scoping reviews. Additionally, only individuals who are proficient in English will be
446 eligible to participate in the e-Delphi survey and consensus meeting. Therefore, the findings of this study
447 may not reflect relevant knowledge that is not available in English.

448 Authors' contributions

449 **Conceptualization:** Douglas G. Manuel, Carol Bennett.

450 **Funding acquisition:** Douglas G. Manuel.

451 **Methodology:** Douglas G. Manuel, Carol Bennett, Emma Brown, David Moher.

452 **Supervision:** Douglas G. Manuel.

453 **Writing – original draft:** Douglas G. Manuel, Carol Bennett, Emma Brown.

454 **Writing – review and editing:** David Moher, Matthew J. Wade, Yoni Freedhoff, David L. Buckeridge,

455 Farah Ishtiaq, Sarah Funnell.

456 Acknowledgements

457 We thank Mary Jessome (PhD candidate, University of British Columbia) for her contributions to the
458 study as a past executive group member with expertise in social sciences. We would also like to thank
459 Peter Farrell (Research Librarian, University of Ottawa) for developing the associated scoping review
460 search strategies for this study.

461 PHES-EF Executive Group (ordered alphabetically according to surname except for the first two
462 authors):

463 Douglas G. Manuel* (0000-0003-0912-0845), David L. Buckeridge* (0000-0003-1817-5047), Yoni
464 Freedhoff (0000-0001-9391-3455), Sarah Funnell, Bernd M. Gawlik (0000-0002-4406-5362), Farah
465 Ishtiaq (0000-0002-6762-7014), Amy E. Kirby, Kerrigan McCarthy (0000-0001-8958-9795), David Moher

466 (0000-0003-2434-4206), Beate Sander (0000-0003-2128-9133), Jeremy Veillard, Matthew J. Wade

467 (0000-0001-9824-7121)

468 *Co-chairs of PHES-EF Executive Group

469 Supporting information

470 **S1 Table. Executive group composition.** Note: The regional classifications are used by the World Bank
471 for organizing economic data, development projects, and policy analysis. These classifications are not
472 strictly geographical but also consider economic and developmental similarities and relationships.

473 **S2 Table. Executive group and e-Delphi panellist disciplines.** The list of specializations or related
474 concepts is not extensive.

475 **S3 Table. World region classifications.** Note: The regional classifications are used by the World Bank for
476 organizing economic data, development projects, and policy analysis. These classifications are not
477 strictly geographical but also consider economic and developmental similarities and relationships.

478 References

- 479 1. Woodward G, Manuel DG, Goel V. Developing a balanced scorecard for public health. Toronto:
480 Institute for Clinical Evaluative Sciences; 2004.
- 481 2. McClary-Gutierrez JS, Mattioli MC, Marcenac P, Silverman AI, Boehm AB, Bibby K, et al. SARS-CoV-2
482 Wastewater Surveillance for Public Health Action. *Emerg Infect Dis.* 2021;27: e210753.
483 doi:10.3201/eid2709.210753
- 484 3. Diamond MB, Keshaviah A, Bento AI, Conroy-Ben O, Driver EM, Ensor KB, et al. Wastewater
485 surveillance of pathogens can inform public health responses. *Nat Med.* 2022;28: 1992–1995.
486 doi:10.1038/s41591-022-01940-x
- 487 4. Ettore Zuccato, Chiara Chiabrando, Sara Castiglioni, Renzo Bagnati, Roberto Fanelli. Estimating
488 Community Drug Abuse by Wastewater Analysis. *Environ Health Perspect.* 2008;116: 1027–1032.
489 doi:10.1289/ehp.11022

- 490 5. Hruday SE, Bischel HN, Charrois J, Chik AHS, Conant B, Delatolla R, et al. Wastewater Surveillance for
491 SARS-CoV-2 RNA in Canada. *FACETS*. 2022;7: 1493–1597. doi:10.1139/facets-2022-0148
- 492 6. World Health Organization. Environmental surveillance for SARS-COV-2 to complement public health
493 surveillance. Geneva, Switzerland: World Health Organization; 2023 Sep p. 55. Available:
494 <https://www.who.int/publications/i/item/9789240080638>
- 495 7. Manuel DG, Delatolla R, Fisman DN, Fuzzen M, Graber T, Katz GM, et al. The Role of Wastewater
496 Testing for SARS-CoV-2 Surveillance. *Ontario COVID-19 Science Advisory Table*; 2021 Aug.
497 doi:10.47326/ocsat.2021.02.40.1.0
- 498 8. Woytus A. What’s in the wastewater? In: St. Louis [Internet]. 22 Sep 2021 [cited 1 Jun 2023].
499 Available: <https://www.stlmag.com/api/content/3f60d498-1b26-11ec-a8ea-1244d5f7c7c6/>
- 500 9. D’Aoust PM, Towhid ST, Mercier É, Hegazy N, Tian X, Bhatnagar K, et al. COVID-19 wastewater
501 surveillance in rural communities: Comparison of lagoon and pumping station samples. *Sci Total*
502 *Environ*. 2021;801: 149618. doi:10.1016/j.scitotenv.2021.149618
- 503 10. Shah S, Gwee SXW, Ng JQX, Lau N, Koh J, Pang J. Wastewater surveillance to infer COVID-19
504 transmission: A systematic review. *Sci Total Environ*. 2022;804: 150060.
505 doi:10.1016/j.scitotenv.2021.150060
- 506 11. Liu P, Ibaraki M, VanTassell J, Geith K, Cavallo M, Kann R, et al. A sensitive, simple, and low-cost
507 method for COVID-19 wastewater surveillance at an institutional level. *Sci Total Environ*. 2022;807:
508 151047. doi:10.1016/j.scitotenv.2021.151047
- 509 12. Larsen DA, Wigginton KR. Tracking COVID-19 with wastewater. *Nat Biotechnol*. 2020;38: 1151–1153.
510 doi:10.1038/s41587-020-0690-1
- 511 13. Gretchen Vogel. Signals from the sewer—Measuring virus levels in wastewater can help track the
512 pandemic. But how useful is that? *Science*. 9 Mar 2022;375. Available:
513 [https://www.science.org/content/article/pandemic-signals-sewer-what-virus-levels-wastewater-](https://www.science.org/content/article/pandemic-signals-sewer-what-virus-levels-wastewater-tell-us)
514 [tell-us](https://www.science.org/content/article/pandemic-signals-sewer-what-virus-levels-wastewater-tell-us). Accessed 1 Jun 2023.
- 515 14. Manuel D, Amadei CA, Campbell JR, Brault J, Zierler A, Veillard J. Strengthening Public Health
516 Surveillance Through Wastewater Testing: An Essential Investment for the COVID-19 Pandemic
517 and Future Health Threats. World Bank Group. The World Bank; 2022 pp. 1–67.
- 518 15. Fynn JF, Hardeman W, Milton K, Jones AP. A scoping review of evaluation frameworks and their
519 applicability to real-world physical activity and dietary change programme evaluation. *BMC Public*
520 *Health*. 2020: 1000. doi:<https://doi.org/10.1186/s12889-020-09062-0>
- 521 16. Nilsen P. Making sense of implementation theories, models, and frameworks. *Implement Sci* 30.
522 2020; 53–79.
- 523 17. Calba C, Goutard FL, Hoinville L, Hendrikx P, Lindberg A, Saegerman C, et al. Surveillance systems
524 evaluation: a systematic review of the existing approaches. *BMC Public Health*. 2015;15.
525 doi:10.1186/S12889-015-1791-5

- 526 18. Andrews JR, Yu AT, Saha S, Shakya J, Aiemjoy K, Horng L, et al. Environmental Surveillance as a Tool
527 for Identifying High-risk Settings for Typhoid Transmission. *Clin Infect Dis*. 2020;71: S71–S78.
528 doi:10.1093/cid/ciaa513
- 529 19. Porta M. *A Dictionary of Epidemiology*. A Dictionary of Epidemiology. Oxford University Press; 2016.
530 Available:
531 [https://www.oxfordreference.com/display/10.1093/acref/9780199976720.001.0001/acref-](https://www.oxfordreference.com/display/10.1093/acref/9780199976720.001.0001/acref-9780199976720)
532 [9780199976720](https://www.oxfordreference.com/display/10.1093/acref/9780199976720.001.0001/acref-9780199976720)
- 533 20. Spranger J, Homberg A, Sonnberger M, Niederberger M. Reporting guidelines for Delphi techniques
534 in health sciences: A methodological review. *Z Evid Fortbild Qual Gesundh Wesen ZEFQ*. 172: 1–11.
- 535 21. How epidemiology has shaped the COVID pandemic. *Nature*. 2021;589: 491–492.
536 doi:10.1038/d41586-021-00183-z
- 537 22. Devine D, Gaskell J, Jennings W, Stoker G. Trust and the Coronavirus Pandemic: What are the
538 Consequences of and for Trust? An Early Review of the Literature. *Polit Stud Rev*. 2021;19: 274–
539 285. doi:10.1177/1478929920948684
- 540 23. World Health Organization. Defining collaborative surveillance: a core concept for strengthening the
541 global architecture for health emergency preparedness, response, and resilience (HEPR). Geneva:
542 World Health Organization; 2023. Available: <https://iris.who.int/handle/10665/367927>
- 543 24. Tricco AC, Lillie E, Zarin W, O’Brien KK, Colquhoun H, Levac D, et al. PRISMA Extension for Scoping
544 Reviews (PRISMA-ScR): Checklist and Explanation. *Ann Intern Med*. 2018;169: 467–473.
545 doi:10.7326/M18-0850
- 546 25. Nasa P, Jain R, Juneja D. Delphi methodology in healthcare research: How to decide its
547 appropriateness. *World J Methodol*. 2021;11: 116–129. doi:10.5662/wjm.v11.i4.116
- 548 26. Shang Z. Use of Delphi in health sciences research: A narrative review. *Medicine (Baltimore)*.
549 2023;102: e32829. doi:10.1097/MD.00000000000032829
- 550 27. Guidance on Conducting and REporting DELphi Studies (CREDES) in palliative care:
551 Recommendations based on a methodological systematic review. 2017;31: 684–706.
- 552 28. Niederberger M, Spranger J. Delphi Technique in Health Sciences: A Map. *Front Public Health*.
553 2020;8: 457. doi:10.3389/fpubh.2020.00457
- 554 29. Schmalz U, Spinler S, Ringbeck J. Lessons Learned from a Two-Round Delphi-based Scenario Study.
555 *MethodsX*. 2021;8: 101179. doi:10.1016/j.mex.2020.101179
- 556 30. Boel A, Navarro-Compán V, Landewé R, van der Heijde D. Two different invitation approaches for
557 consecutive rounds of a Delphi survey led to comparable final outcome. *J Clin Epidemiol*.
558 2021;129: 31–39. doi:10.1016/j.jclinepi.2020.09.034
- 559 31. Gattrell WT, Hungin AP, Price A, Winchester CC, Tovey D, Hughes EL, et al. ACCORD guideline
560 for reporting consensus-based methods in biomedical research and clinical practice: a study
561 protocol. 2022;7: 3.

- 562 32. Permeth-Wey J, Borenstein AR. Financial Remuneration for Clinical and Behavioral Research
563 Participation: Ethical and Practical Considerations. *Ann Epidemiol.* 2009;19: 280–285.
564 doi:10.1016/j.annepidem.2009.01.004
- 565 33. Surveylet. United States: Calibrium; Available: <https://calibrium.com>
- 566 34. LimeSurvey: An open source survey tool. Hamburg, Germany: LimeSurvey GmbH; Available:
567 <https://www.limesurvey.org>
- 568 35. Diamond IR, Grant RC, Feldman BM, Pencharz PB, Ling SC, Moore AM, et al. Defining consensus: A
569 systematic review recommends methodologic criteria for reporting of Delphi studies. *J Clin*
570 *Epidemiol.* 2014;67: 401–409. doi:10.1016/j.jclinepi.2013.12.002
- 571 36. Belton I, MacDonald A, Wright G, Hamlin I. Improving the practical application of the Delphi method
572 in group-based judgment: A six-step prescription for a well-founded and defensible process.
573 *Technol Forecast Soc Change.* 2019;147: 72–82. doi:10.1016/j.techfore.2019.07.002
- 574 37. Weijters B, Cabooter E, Schillewaert N. The effect of rating scale format on response styles: The
575 number of response categories and response category labels. *Int J Res Mark.* 2010;27: 236–247.
576 doi:10.1016/j.ijresmar.2010.02.004
- 577 38. Lange T, Kopkow C, Lütznier J, Günther K-P, Gravius S, Scharf H-P, et al. Comparison of different
578 rating scales for the use in Delphi studies: different scales lead to different consensus and show
579 different test-retest reliability. *BMC Med Res Methodol.* 2020;20: 28. doi:10.1186/s12874-020-
580 0912-8
- 581 39. Government of Canada. Best Practices in Equity, Diversity and Inclusion in Research. In: *New*
582 *Frontiers in Research Fund* [Internet]. 22 Jun 2021. Available: [https://www.sshrc-](https://www.sshrc-crsh.gc.ca/funding-financement/nfrf-fnfr/edi-eng.aspx)
583 [crsh.gc.ca/funding-financement/nfrf-fnfr/edi-eng.aspx](https://www.sshrc-crsh.gc.ca/funding-financement/nfrf-fnfr/edi-eng.aspx)
- 584 40. Government of Canada. Introduction to GBA Plus. 31 Mar 2021 [cited 1 Jun 2023]. Available:
585 [https://women-gender-equality.canada.ca/gbaplus-course-cours-](https://women-gender-equality.canada.ca/gbaplus-course-cours-acsplus/eng/mod00/mod00_01_01.html)
586 [acsplus/eng/mod00/mod00_01_01.html](https://women-gender-equality.canada.ca/gbaplus-course-cours-acsplus/eng/mod00/mod00_01_01.html)
- 587 41. The First Nations Information Governance Centre. The First Nations Principles of OCAP®. In: *The*
588 *First Nations Information Governance Centre* [Internet]. 1 Jun 2023 [cited 1 Jun 2023]. Available:
589 <https://fnigc.ca/ocap-training/>
- 590 42. Staniszewska S, Brett J, Simera I, Seers K, Mockford C, Goodlad S, et al. GRIPP2 reporting checklists:
591 tools to improve reporting of patient and public involvement in research. *BMJ.* 2017;358: j3453.
592 doi:10.1136/bmj.j3453
- 593 43. Gattrell WT, Logullo P, Zuuren EJ van, Price A, Hughes EL, Blazey P, et al. ACCORD (ACcurate
594 COnsensus Reporting Document): A reporting guideline for consensus methods in biomedicine
595 developed via a modified Delphi. *PLOS Med.* 2024;21: e1004326.
596 doi:10.1371/journal.pmed.1004326

- 597 44. Yang L, Weston C, Cude C, Kincl L. Evaluating Oregon’s occupational public health surveillance
598 system based on the CDC updated guidelines. *Am J Ind Med.* 2020;63: 713–725.
599 doi:10.1002/ajim.23139
- 600 45. Winkler J, Moser R. Biases in future-oriented Delphi studies: A cognitive perspective. *Technol*
601 *Forecast Soc Change.* 2016;105: 63–76. doi:10.1016/j.techfore.2016.01.021
- 602 46. Goodman CM. The Delphi technique: a critique. *J Adv Nurs.* 1987;12: 729–734. doi:10.1111/j.1365-
603 2648.1987.tb01376.x
- 604

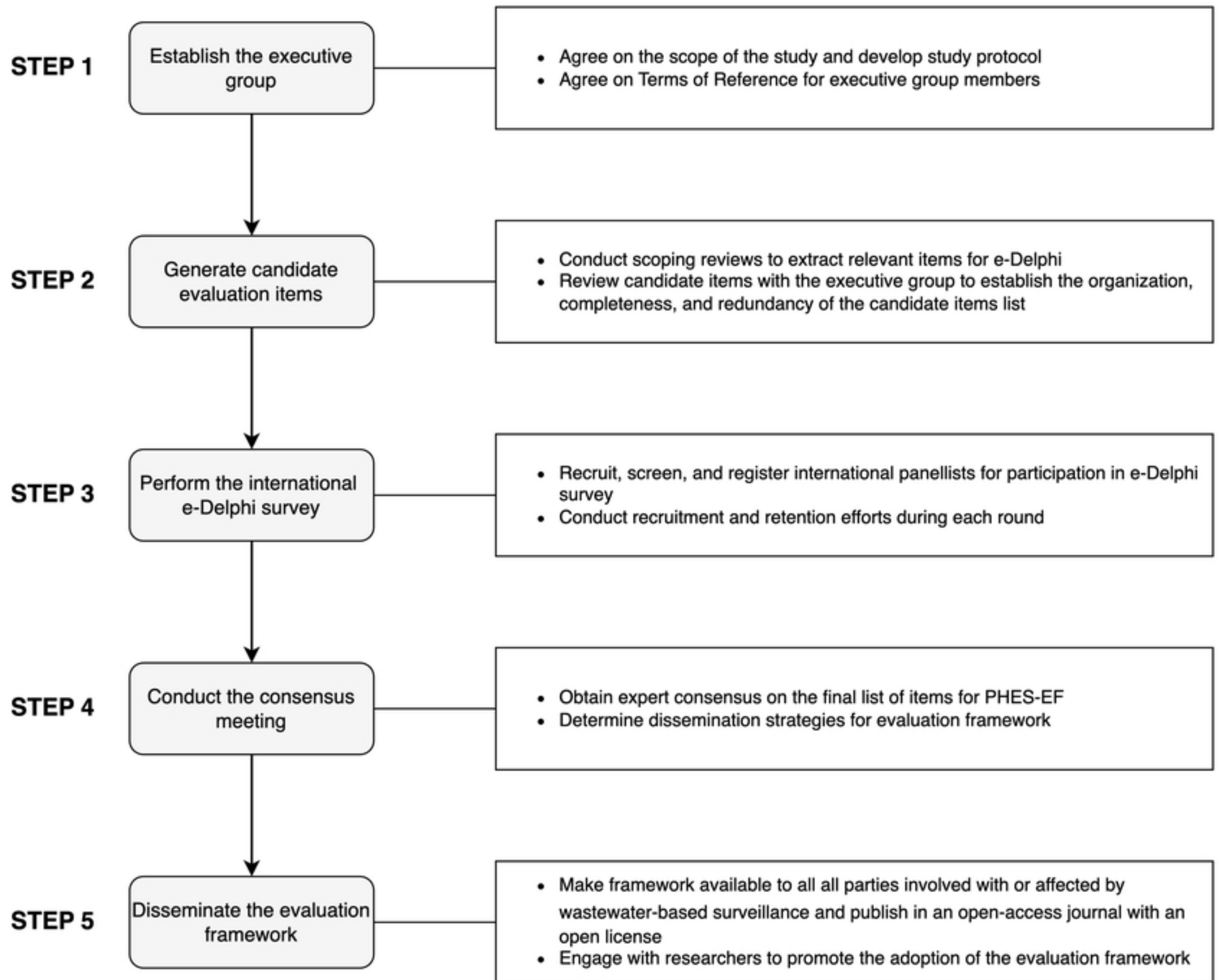


Figure 1

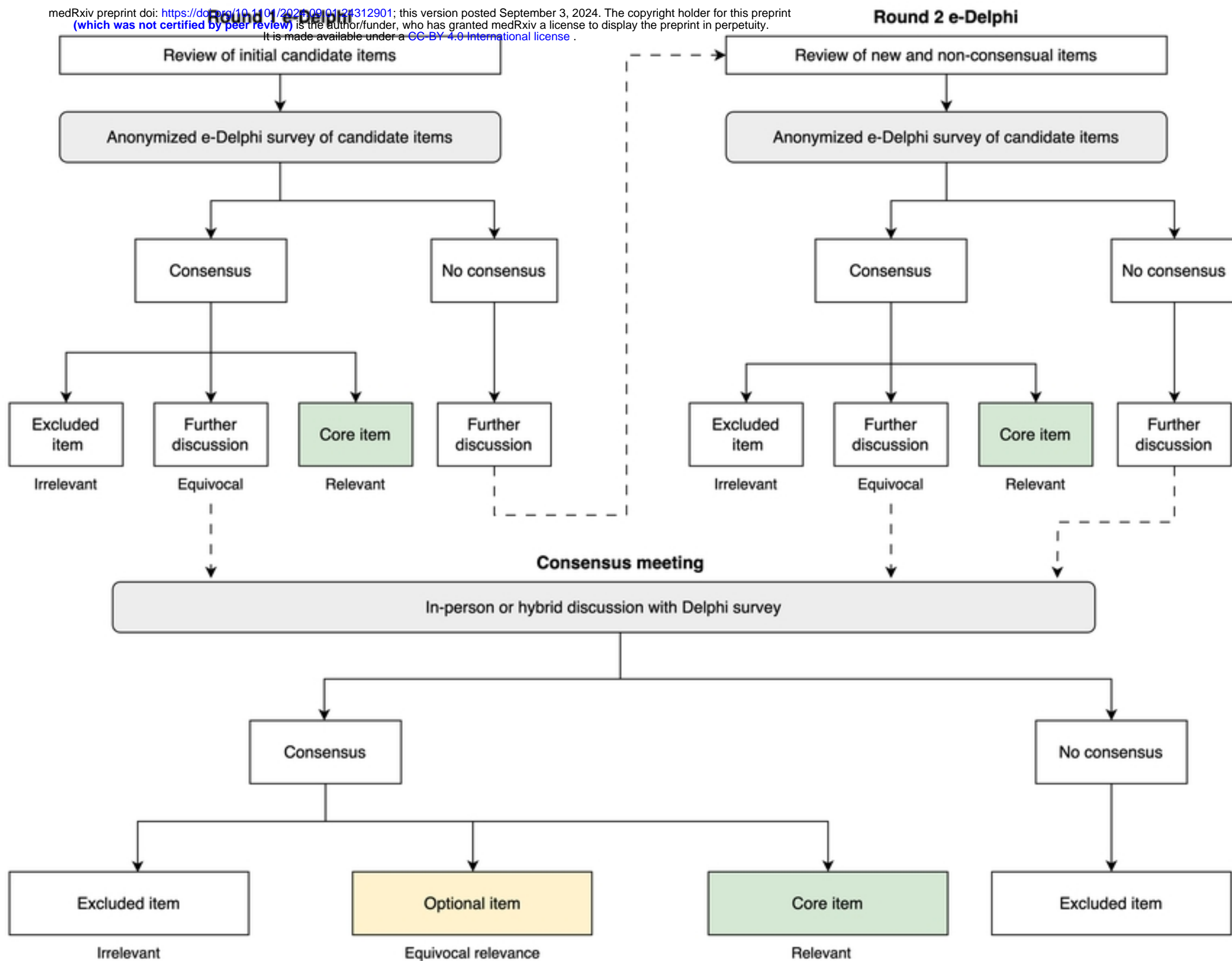


Figure 2