IMPROVE

Framework to IMPROVE the Integration of Patient Generated Health Data to Facilitate Value Based Healthcare

D2.1: Systematic review report and updates V1

Version 1.0

Editors:

Rens van de Schoot (UU)

Frans Folkvord (PBY)

Davide Guerri (Dedalus)

Co-funded by the European Union

Document Control Sheet

Deliverable Number	D2.1
Deliverable Responsible	UU
Work Package	WP2
Lead Editors	Rens van de Schoot (UU) Frans Folkvord (PredictBy) Davide Guerri (Dedalus)
Co-authors	Felix Weijdema (UU) Kianush Monschau (UU) Elena Jalsovec (UU) Rutger Neeleman (UU) Qixiang Fang (UU) Emily Westerbeek (UU) Jim Ingebretsen Carlson (PBY) Carmela Genovese (Dedalus) Chiara Macagnano (Dedalus) Franco Chiarugi (Dedalus) Laura Pinna (Dedalus) Nadine Bol (TiU) Emiel Krahmer (TiU) Linwei He (TiU) Hans Peeters (PMS) Eva Turk (STPUAS) Martin Ernst (STPUAS) Beatriz Merino (UPM) Diego Carvajal (UPM) Manuel Ottaviano (UPM) Giuseppe Fico (UPM) Caridad Pontes (IDIBELL) Marina Ramiro Pareta (IDIBELL) Alba Jiménez Rueda (IDIBELL) Jordi Piera Jiménez (IDIBELL) Eva Podovšovnik (OBV) Laia Juan (MDT) Clàudia Navarro (MDT) Marko Ogorevc (IER)

Reviewer(s)	Clàudia Navarro (MDT), Manuel Ottaviano (UPM)

History of Changes

Date	Version/Page	Change
25-05-2024	0.1	Pre-registration (published on Prospero)
18-06-2024	0.2	Update with topic allocation
17-10-2024	0.3	Adjust Lay-out and add screening results
22-11-2024	0.4	Add FT screening results and quantitative gap analyses
26-11-2024	0.5	Add qualitative gap analyses and conclusion
27-11-2024	0.6	Finalize document for review
04-12-2024	1.0	Version ready for submission

Statement of Originality

This deliverable contains original unpublished work except where clearly indicated otherwise. Acknowledgement of previously published material and of the work of others has been made through appropriate citation, quotation or both.

Legal Disclaimer

The information in this document is provided “as is” and as it has been collected according to the inputs provided by the different partners. The above referenced consortium members shall have no liability to third parties for damage of any kind including without limitation direct, special, indirect, or consequential damages that may result from the use of these materials subject to any liability which is mandatory due to applicable law. This data management plan is a living document and will evolve with the advancement of the project.

Abstract

This report describes the protocol used and its results for conducting the umbrella review focusing on the existing state-of-the-art scientific evidence about the use of various types of Patient Generated Health Data (PGHD) in healthcare delivery, concerning PGHD in several disease areas throughout the patient journey. We focus on obtaining state-of-the-art scientific evidence on the integration of in-clinic and out-of-clinic PGHD and experiences to harness value-based healthcare (VBHC) through improving the quality, reliability and use of Patient-Reported Outcome Measures, Patient Preference Information, Patient-Reported Experience Measures to enhance healthcare enabling accelerated innovation of cost-effective and personalized patient journeys, based on accurate insight in health condition and treatment options in relation to foreseeable outcomes, patient experiences and preferences which are integrated for informed decision making by the patient, family members, and health care professionals. The protocol resulted in a database with scientific evidence on integrating PGHD. Because there is a growing and large corpus of documents published on this subject area and we do not want to limit the search to a specific subset, we first created a database limited to systematic reviews or meta-analyses published in English in the last 5 years (i.e., umbrella review). In the second year, we will update the search to individual Randomized Control Trials (RCTs) of any publication date and language. The records obtained via the systematic search in Y1 were screened by the use-case experts, and relevant data was extracted and put in a database. To deal with the enormous amount of potentially relevant records, we applied the Screenathon Review procedure where the labeling process is crowdsourced. In addition, we have conducted a gap analysis to see where the most important impact can be achieved, leading to the coordination of the use cases to be defined. The resulting database with extracted data from the relevant systematic reviews will establish a starting point for the development of the **IMPROVE** lab in WP3 and the development and execution of the data collection in WP4 and WP5, and for the guidelines and best next practices in WP7.

Keywords: Systematic Review, Gap analysis, Patient Generated Health Data; Knowledge warehouse

Abbreviations and Acronyms

AI	Artificial Intelligence
EC	European Commission
EHDS	European Health Data Space
EHR	Electronic Health Record
EU	European Union
HE	Horizon Europe
HTA	Health Technology Assessment
KPI	Key Performance Indicator
MA	Meta-Analysis
ML	Machine Learning
M	Project month
MS	Milestone
MSC	Master Scorecard
NLF	Noisy Label Filter
NLP	Natural Language Processing
OECD	Organization for Economic Cooperation and Development
PGHD	Patient Generated Health Data
PROMs	Patient-Reported Outcome Measures
PPI	Patient Preference Information
PREMs	Patient-Reported Experience Measures
RCT	Randomized Controlled Trial
RWD	Real World Data
SR	Systematic Review
TA	Title-Abstract
VBHC	Value-Based Healthcare
WP	Work Package

Table of Contents

List of Figures

List of Tables

1. Introduction

Patient Generated Health Data (PGHD) and knowledge-sharing across the European Union (EU) will make healthcare provision 'smarter' and will accelerate the development of (cost-)effective and patient-preference based new treatments and medical devices and reduce the operational costs of integrated healthcare solutions by making the patient more central in the healthcare process (Tian et al., 2019). Healthcare professionals, pharmacists, researchers, health industry, and health regulators all over the EU generate and use large numbers of essential patient-related healthcare data that are critical to the quality and effectiveness of their work. Unfortunately, there are still complex obstacles that make it difficult to reach the full potential of digital health and patient-related data. An important and highly relevant initiative activated by the European Commission (EC), the European Health Data Space (EHDS; *European Health Data Space*, 2024), is promised to overcome these obstacles. The EHDS is a sharing framework that establishes clear rules, common standards and practices, infrastructures and a governance framework for the use of electronic health data by patients and for research, innovation, policy making, patient safety, statistics or regulatory purposes. In line with this, the main aim of IMPROVE is to create an accessible, functional, transferable and (cost-)effective framework that is capable of automatically enabling and integrating the added value of PGHD integrated healthcare solutions using patient-reported outcome measures (PROMs), patient preference information (PPI), and patient-reported experience measures (PREMs) and other people-generated information, accompanied by a management structure that can meet regulatory (e.g., AI Act, Data Act), ethical, legal, statistical and data requirements to facilitate decision makers, patients, researchers, and healthcare professionals.

IMPROVE will develop an evidence-based and real-time framework to effectively leverage integrated added value of people-centred integrated healthcare solutions, using predominantly, but not limited to, PROMs, PPI, and PREMs. This information will be established in first instance by scientific evidence, subsequently complemented by Real-World Evidence and Real-World Data, in order to have a more comprehensive understanding of how patient-generated evidence can best be used to improve outcomes, support decision making, and accelerate innovation by providing tailored solutions to the industry. Developing approaches for such comprehensive data collection framework is timely in view of the challenge and ambition formulated in the EHDS for both primary and secondary data use. IMPROVE will increase the effective usage of such data enabling clinical innovation, better health outcomes, and advancing and consolidating evidence-based decision making for further acceleration of innovation and health system sustainability.

In this report we elaborate on the protocol, which has been pre-registered at Prospero (https://www.crd.york.ac.uk/prospero/display_record.php?RecordID=546427), for the extensive literature review we conducted together with the IMPROVE partners in Work Package 2 (WP2). This literature review is focused on the existing state-of-the-art scientific evidence about the use of Patient Generated Health Data PGHD in healthcare delivery, conceptual relations concerning PGHD in several disease areas throughout the patient's journey. The protocol resulted in a database with scientific

evidence on integrating PGHD. Because there is a growing and large corpus of documents published on this subject area and we do not want to limit the search to a specific subset, we first created a database limited to systematic reviews or meta-analyses published in English in the last 5 years (i.e., an umbrella review). In the second year, we will update the search to individual RCTs of any publication date and language. The records obtained via the systematic search in Y1 were screened by experts via the Screenathon Review procedure (see section 1.1). Relevant data from the included systematic reviews was extracted and put in a database. This database will establish a starting point for the development of the IMPROVE lab in WP3 and the development and execution of the data collection in WP4 and WP5, and for the guidelines and best next practices in WP7. In addition, we conducted a gap analysis to see where the most important impact can be achieved, leading to the coordination of the use cases to be defined. In what follows, we describe relevant background information about systematic reviewing and the Screenathon Review procedure, the methodological steps for obtaining the database with potentially relevant papers, followed by its results and the gap analysis.

1.1. Background Information on Systematic Reviewing

A systematic review (SR), with or without a formal quantitative synthesis or meta-analysis (MA), is a comprehensive overview of the existing evidence pertinent to a clearly formulated question, which uses pre-specified and standardised methods to identify and critically appraise relevant research, and to collect, report and analyse data or information from the publications (peer-reviewed articles, reports, guidelines, etc.) studies that are included in the review. The methodology for SRs was originally standardised by the Cochrane Collaboration (Chandler & Hopewell, 2013; Higgins et al., 2024) for addressing the effects of health care interventions, who developed and refined all SR relevant steps: the a priori specification of a research question; clarity on the scope of the review and which studies are eligible for inclusion; making every effort to find all relevant research and to ensure that issues of bias in included studies are accounted for; and analysing the included studies in order to draw conclusions based on all the identified research in an impartial and objective way.

Performing an SR is a very rigorous process, increasingly resource-intensive due to the ever-increasing number of scientific publications to review. Nevertheless, SRs are pivotal not only for scholars, but also for clinicians, policymakers, journalists, and, ultimately, the general public (Gough & Elbourne, 2002), and therefore, an excellent tool for the objectives of this report. Over time, SRs have extended to many other research fields than healthcare and include many other health-related questions than only intervention effects, namely also the value and accuracy of diagnostic and screening tests, unintended effects, prognostic accuracy and etiologic questions.

Developing a search strategy for an SR is an iterative process aimed at balancing recall, precision, and quality (Lefebvre et al., 2008). That is, including as many potentially relevant and ideally high-quality studies as possible (recall and sensitivity), while at the same time limiting the total number of studies to screen (precision or specificity). Critical appraisal is an essential step to focus on results that are relevant to the research question, and that can reliably support or refute its health claims or safety

issues with high-quality evidence or identify high-level research relevant to the practice. Hence, choosing an appropriate critical appraisal tool relevant to the research design is an important component of evidence-based decision-making.

In a systematic review, primary research studies rather than reports of studies are the principal unit of interest, while in an **umbrella review**, the systematic reviews are the main unit of interest. There is a growing number of institutions and research centres working on standards for conducting SRs and MAs in their respective fields. SRs should use eight key steps in the process, which are described in Section 2:

1. preparing a review protocol,
2. searching for studies,
3. selecting studies for inclusion,
4. extracting data from included studies,
5. assessing the methodological quality of the included studies,
6. qualitative or quantitative synthesising of the extracted safety or effectiveness data from the studies,
7. presenting the synthesized data and its meta-analysed results, and
8. interpreting the results and drawing conclusions.

For the present project we make use of a novel title-abstract screening procedure termed a **Screenathon Review**. This methodology is designed specifically for a consortium setting. It uses crowdsourcing to screen literature in a limited amount of time whilst accomplishing two adjacent social goals: It helps consortium members form consensus about inclusion criteria and definitions, whilst also helping to create research community through incentives and gamification. Its main innovation over traditional screening is the Screenathon event, a multi-day meeting where relevant stakeholders come together to screen a pre-determined number of records. For details about the procedure see Monschau et al. (2024) and for our implementation see Section 2.2 and for the results see Section 3.1.

1.2. Risk of Bias / Quality Assessment

In a systematic review, primary research studies rather than reports of studies are the principal unit of interest, while in an umbrella review, the systematic reviews are the main unit of interest. An important step to consider is the examination of any retraction statements and errata for information. In fact, some studies may have been found to be fraudulent or may, for other reasons, have been retracted since publication. Errata can reveal important limitations, or even fatal flaws, in included studies. All of these may potentially lead to the exclusion of a study from a review or meta-analysis. Care will be taken to ensure that this information is retrieved in all database searches by downloading the appropriate fields together with the citation data.

We only include systematic reviews and meta-analyses published in scientific journals. Risk of bias is assessed by means of quality of the data collected (e.g., RCTs, clinical trials, interviews, discussion papers), number of participants included, disease agnostic, etc.

By extracting data consistently and transparently, researchers can minimize bias. Moreover, the use of dual (or even more) extraction, with two or more reviewers extracting data independently, followed by identification of discrepancies and conflict resolution can mitigate the risk of bias. Additionally, during the data extraction, the reported risk of bias assessment will be extracted, when available, from each selected systematic review or study and will be taken into account for the evaluation of the quality of the selected systematic review or study.

An important step to consider is the examination of any retraction statements and errata for information. In fact, some studies may have been found to be fraudulent or may, for other reasons, have been retracted since publication. Errata can reveal important limitations, or even fatal flaws, in included studies. All of these may potentially lead to the exclusion of a study from a review or meta-analysis. Care will be taken to ensure that this information is retrieved in all database searches by downloading the appropriate fields together with the citation data.

1.3. Nomenclature

The nomenclature that was used for developing the inclusion criteria are listed in Table 1 below.

Table 1 Nomenclature Used to Develop Inclusion Criteria

Nomenclature	
HTA	**Health Technology Assessment**: Health Technology Assessment (HTA; *Health Technology*, n.d.) informs reimbursement and coverage decisions on how to allocate healthcare resources to different health technologies by carefully assessing the costs and benefits of health interventions, using cost-effectiveness and impact assessment as instruments.
PC	**Patient centricity**: Putting the patient first in an open and sustained engagement throughout the full process, to respectfully and compassionately achieve the best experience and outcome for that person and their family, committed to hearing,

	understanding and integrating patients' perspective in regulatory decision making as appropriate, considering '*valid scientific evidence*' when conducting benefit-risk assessment, including nonclinical and clinical investigations and patient information (such as PGHD).
PGHD	**Patient Generated Health Data**: Patient-generated health data (PGHD), created and captured from patients via wearable devices and mobile apps, are proliferating outside of clinical settings. Examples include sleep trackers, fitness trackers, continuous glucose monitors, and RFID-enabled implants, with many additional biometric (*Biometrics*, n.d.) or health surveillance applications in development or envisioned. These data are included in growing stockpiles of personal health data (PHI) being mined for insight by health economists, policy analysts, researchers, and health system organizations (Winter & Davidson, 2022).
PROMs	**Patient-Reported Outcome Measures**: Patient-reported outcome measures are questionnaires that collect health outcomes directly from the people or patients who experience the health outcomes themselves (Williams et al., 2016).
PPI	**Patient Preference Information**: Qualitative or quantitative assessments of the relative desirability or acceptability to patients, of features that differ among alternative health states, health interventions, or health services. Desirability: preferences for positive outcomes or features – Acceptability: aversion to negative outcomes (Russo et al., 2019).
PREMs	**Patient-Reported Experience Measures**: Patient-reported experience measures are psychometrically validated tools (e.g. questionnaires) used to capture patients' interactions with healthcare systems and the degree to which their needs are being met. Patient-reported experience measures are designed to determine whether patients have experienced certain care processes rather than their satisfaction with the care received (which may be subject to bias). A Patient-reported experience measure may, for instance, be used to collect information on the patient experience of hospital admission. Data derived from this could be used to inform service development and configuration (*Patient-Reported Experience Measure*, n.d.).
VBHC	**Value-Based Healthcare** (Koehring, 2015; Porter, 2010)**:** Value in health care is the measured improvement in a patient's health outcomes for the cost of achieving that improvement (Winter & Davidson, 2022). The goal of value-based care transformation is to enable the health care system to create more value for patients. Because value is created only when a person's health outcomes improve, descriptions of value-based health care that focus on cost reduction are incomplete (Teisberg et al., 2020).

2. Methods

2.1. Database Search

To fit the requirements of the Screenathon set up, the aim was to run a broad search in both more subject-specific as well as more general databases. The search queries used were based on a description of the main terminology in the field of interest: Patient-Generated-Health-Data. In total four databases (PubMed, Embase.com, CINAHL, and Scopus - search date: 04-04-2024) were searched using a data limit of publication year 2014 onwards. The search focuses on systematic literature reviews and meta-analyses on the use of PGHD. We excluded conference abstracts, letters to the editor, theses, and pre-prints. For the time frame of the review, we considered studies published within the last 10 years, as it is a period often used as a convention for SR and large enough for the scope of our review. The eligible studies should be written in English and should contain a persistent object identifier (e.g., DOI, PubMed ID) and the title plus abstract should be available.

An example of part of the search in PubMed is described in Table 2. The exact searches, the nomenclature document, full search details, and the database output (.ris files) can be found on the Open Science Framework: www.osf.io/bh7fy.

Table 2 PubMed Query.

("Patient centri"[tiab] OR "Patient centered"[tiab] OR "Patient focus*"[tiab]) OR ("Patient-Centered Care"[Mesh]) ("Patient Reported Outcome Measures"[Mesh]) OR ("Patient reported outcome*"[tiab] OR "Patient assessed outcome*"[tiab] OR "PROMs"[tiab] OR "PROM"[tiab])* *("Patient Preference"[Mesh]) OR ("Patient preference info*"[tiab] OR "PPI"[tiab] OR "PPIs"[tiab])* *("HTA"[tiab] OR "Health technology assess*"[tiab] OR "Biomedical technology assess*"[tiab]) OR ("Technology Assessment, Biomedical"[Mesh]) ("Patient Generated Health Data"[Mesh]) OR ("PGHD"[tiab] OR "Patient generated health data"[tiab] OR "Patient generated data"[tiab] OR "Patient reported health data"[tiab] OR "Patient reported data"[tiab] OR "Self-assessed health data"[tiab])* *("PREM"[tiab] OR "PREMs"[tiab] OR "Patient reported experience measur*"[tiab])* *("Value-Based Health Care"[Mesh]) OR ("VBHC"[tiab] OR "Value-based health care"[tiab] OR "Value health care"[tiab]) "patient centri*"[Title/Abstract] OR "Patient centered"[Title/Abstract] OR "patient focus*"[Title/Abstract] OR "Patient-Centered Care"[MeSH Terms] OR "Patient Reported Outcome Measures"[MeSH Terms] OR "patient reported outcome*"[Title/Abstract] OR "patient assessed outcome*"[Title/Abstract] OR "PROMs"[Title/Abstract] OR "PROM"[Title/Abstract] OR "Patient Preference"[MeSH Terms] OR "patient preference info*"[Title/Abstract] OR "PPI"[Title/Abstract] OR "PPIs"[Title/Abstract] OR "HTA"[Title/Abstract] OR "health technology assess*"[Title/Abstract] OR "biomedical technology assess*"[Title/Abstract] OR "technology assessment, biomedical"[MeSH Terms] OR "patient generated health data"[MeSH Terms] OR "PGHD"[Title/Abstract] OR "patient generated health data"[Title/Abstract] OR "Patient generated data"[Title/Abstract] OR "Patient reported health data"[Title/Abstract] OR "Patient reported data"[Title/Abstract] OR "Self-assessed health data"[Title/Abstract] OR "PREM"[Title/Abstract] OR "PREMs"[Title/Abstract] OR "patient reported experience measur*"[Title/Abstract] OR "value based health care"[MeSH Terms] OR "VBHC"[Title/Abstract] OR "value based health care"[Title/Abstract] OR "Value health care"[Title/Abstract]* *AND "systematic review*"[Title/Abstract] OR "meta analys*"[All Fields] OR "systematized review*"[Title/Abstract] OR "scoping review*"[Title/Abstract] OR "literature review*"[Title/Abstract] OR "umbrella review*"[Title/Abstract] OR "Systematic Review"[Publication Type]* *AND 2014/01/01:3000/12/31[Date - Entry]* *NOT (letter[PT] OR comment*[PT] OR editorial[pt] OR preprint[pt])* *AND (english[Filter])* *20240408_PbMdpart1_20142020_5984.nbib* *20240408_PbMdpart2_2020-2024_5003.nbib*

The data was managed using Endnote stored in Microsoft's SharePoint for easy access. The search hits (including publication title, authors, abstract and DOI) were downloaded in RIS file format. A file containing all the hits for each search was stored in Microsoft's SharePoint for easy access were

merged and duplications were removed. Duplicates were removed while importing the database output using EndNote X20 (default settings). The combined deduplicated results of the four databases were the starting point for creating the Screenathon dataset.

The dataset for the Screenathon had some additional requirements compared to traditional screening. Based on the number of participants and available hours, we tried to create a large dataset that could be fully screened. This led to an additional filter on publication year: records from 2019 and onwards were included. To facilitate automatic processing and accurate screening of these records during the entire Screenathon, records missing a DOI a Title, or Abstract were excluded from the Screenathon dataset. In the result section we provide all the details about how many records were included, and how many participants from what type of different institutions participated in the Screenathon.

2.2. Eligibility Criteria

The output of the searches was merged into a unified dataset and de-duplicated, via EndNote. The inclusion and exclusion criteria for the studies eligible for this review can be summarized as:

1. Study type: Literature reviews (e.g., systematic, narrative, scoping) or meta-analyses.
2. Content: It should contain patient-generated health data (PGHD). PGHD Is defined as data created and captured from patients via wearable devices, mobile apps or surveys, which are proliferating outside of clinical settings. Examples include sleep trackers, fitness trackers, continuous glucose monitors, and RFID-enabled implants, with many additional biometric or health surveillance applications in development or envisioned.
3. Population: The population of interest of the studies under review is restricted to adult human patients that are, have been or will be under treatment for a certain condition. For this search, the interventions considered are the studies assessing factors influencing treatment adherence, with the identification of the effect on adherence of one or more factors as an outcome of these studies. Studies considering adult human subjects (≥ 16 years old). For reviews and overviews, only those including ≥ 80% of included studies analysing adult population.
4. Condition Type: Both chronic and acute physical conditions.
5. Treatment: The studies eligible for this review are those that analyse PGHD to any kind of treatment or medical recommendation, meaning not only medication taking, but also other health behaviours such as attending follow-up appointments, implementing lifestyle changes (e.g., avoiding certain foods, engaging in specific exercise), using medical devices, among others
6. Data (for the full-text): Studies that for the factors analysed report at least the direction of the effect accompanied by its statistical significance and its uncertainty estimates.

2.3. Title-Abstract Selection Process

The selection process began with a training procedure consisting of two steps: A calibration session and an individual screening. The calibration session was aimed at standardizing evaluation criteria among participants. It involved all members assessing a predetermined set of 10 records to harmonize their understanding of the inclusion criteria. These records and their labels were discussed in a group setting, which enhances the prospective reviewers' understanding of both the criteria and the screening process. This is also a setting in which criteria can be questioned and critically evaluated. Following this, the individual screening component required each reviewer to evaluate a uniform set of 20 records annotated with known outcomes which were unknown to the screeners. To increase the transparency of screening decisions made by reviewers, they were also required to indicate based on which exact criteria a paper was deemed (ir)relevant. In the third phase, we divided the remaining records into batches of 100 records each, and the participants screened as many batches as they can in the two days of the Screenathon. To counteract screening fatigue, attendees were free to flexibly decide how many records they would like to screen. The event included a program of educational activities and venues which attendees were free to attend. A support team was also on-site to facilitate the screening. Attendees were encouraged to mark papers that do not fit the criteria but are deemed interesting for the IMPROVE project nonetheless and were free to discuss the screening process amongst each other. The total number of records screened, and the percentage included, can be found in the results section.

After the Screenathon, the excluded records were systematically reviewed by a screener using ASReview, an open-source tool designed to implement active learning for screening prioritization. In this process, a machine learning model was trained using the records included during the Screenathon as relevant examples and a subset of irrelevant records to distinguish between the two categories. This model predicts the likelihood of relevance for each excluded record. The screener then prioritized reviewing the records that the model predicted as most likely relevant but were initially excluded during the Screenathon.

To ensure a balance between thoroughness and efficiency, we applied the SAFE procedure (Boetje & Van De Schoot, 2024) as a stopping rule. The SAFE procedure provides a statistically supported framework for determining when the likelihood of additional relevant records being missed becomes acceptably low. This approach minimizes the risk of false negatives (relevant records being overlooked) while also reducing unnecessary screening efforts. Any newly identified relevant records were subsequently added to the dataset, ensuring the comprehensiveness of the review.

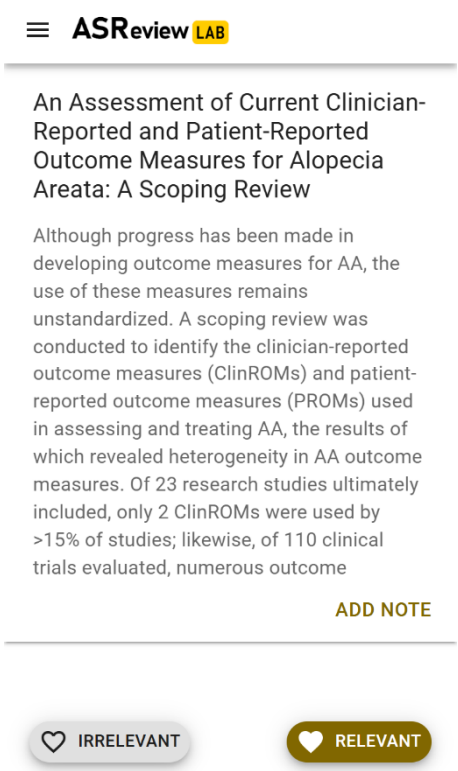

Figure 1 Screenshot of the review screen on a mobile device for one of the papers screened during the Screenathon.

2.4. Topic Allocation

The abstracts from the TA screening phase covered a broad range of disease topics. However, at this point of the project, only the papers that match our specified scope, in terms of disease topics, are of interest to us. The IMPROVE project focuses on the following 11 disease topics in 5 therapeutic areas:

1. Oncology
    - Prostate cancer
    - Cervical cancer
    - Neck and head cancer
    - Breast cancer
2. Ophthalmology
    - Macular degeneration
3. Cardiovascular disease
    - Heart failure
    - Coronary artery diseases
    - Atrial fibrillation
    - Severe aortic stenosis
4. Neurology:
    - Multiple sclerosis
5. Chronic inflammation:
    - Chronic rhinosinusitis

To filter out the irrelevant papers, we made use of OpenAlex's topic system (*Topics | OpenAlex Technical Documentation*, 2024), where OpenAlex uses a trained classifier to assign up to 3 most likely disease topics to every paper hosted on its platform. By using this topic system, we can quickly identify papers that are relevant for us per our defined scope of disease topics, without the need for manual screening or training a classification model on our own.

To start with, we identified paper records on OpenAlex that match our included papers, based on their DOIs and paper titles. Then, we matched our defined list of disease topics to those of OpenAlex's. For every specific topic that we defined, we counted the number of papers that have been assigned a related OpenAlex topic. This means that a paper can also be associated with different topics and hence gets counted more than once.

The papers matching the topics were moved to the full-text screening phase.

2.5. Full-text Review

After the TA screening and topic allocation phases, the included papers were screened for eligibility by two independent screeners for each of the five use cases, considering the eligibility criteria described above. For each publication assigned, the reviewer checked each criterion and assessed the inclusion of only those publications meeting the full criteria. Each of the publications was reviewed by a second independent reviewer from PredictBy.

The results from all screening phases were then combined into one excel file. This file contains a row for every paper screened during the Screenathon. There are binary columns to indicate inclusion in the screening steps, along with various metadata and more elaborate screening decision data.

2.6. Data Extraction

Data extraction is the process of systematically identifying relevant characteristics of (systematic) reviews based on the information available in the selected publications. It provides the basis for meta-analysis and is a necessary step that precedes the assessment of the risk of bias in authors' synthesis of their findings or in reviewers' interpretation.

Data extraction mainly followed the PICO framework (patient (problem or population), intervention, comparison (control or comparator), outcome), but other data like T (timing) and S (type of study) was extracted when available. Meta-data of the articles, such as DOI, authors, title, journal, publication year and open access availability were extracted via OpenAlex.

Reviewers extracted the fields listed below from the systematic reviews and the individual studies as reported in the reviews (note that the individual studies themselves will be inspected in the update in Year 2). For the extraction a custom template was used in Covidence, created for this project. This streamlined the extraction, making it easier and ensuring that the format was the same for all papers.

These fields focus on gap analysis and apply to systematic reviews:

1. Type of review
2. Research questions/objectives
3. Inclusion/exclusion criteria
4. Reported risk of bias assessment of the systematic reviews (Y/N)
5. Type of PGHD (Patient-Generated Health Data) involved
    a. PROMs
    b. PREMs
    c. PPIs
    d. RWE
    e. RWD
    f. Other

6. Main findings/conclusions (outcomes, costs, healthcare delivery, sustainability, etc.)

These fields are specific to the individual studies reported within the systematic reviews and were only extracted if reported in the systematic review, the remainder will be extracted in Y2 in the update when we process the individual studies:

1. Research questions/objectives
2. Reported risk of bias assessment of RCT
3. Type of PGHD involved
4. Methodology for the collection of PGHD
5. Main findings/conclusions (outcomes, costs, healthcare delivery, sustainability, etc.)
6. Country of the study
7. gender dimension
8. Type of study (RCT, cohort studies, case-control studies, cross-sectional studies, case series/report)
9. Timing (years and duration)
10. Subjects included (number, main demographic characteristics (age, sex))
11. Disease area (including comorbidities)
12. Type of intervention evaluated
13. Factors affecting adherence considered
14. Study design/methodology
15. Type of experimental or non-experimental design
16. Values of effect sizes
17. Outcomes costs

2.7. Gap Analysis

After data extraction and data synthesis which include also, if enough data is available, a subgroup analysis focused on the different diseases foreseen in the project, a gap analysis was performed to check weaknesses in the use/exploitation of the PGHD among the relevant disease areas and additional disease areas. The gap analysis will be conducted to see where historical data needs to feed the framework (WP4) and how we can design the use cases (WP5) to complement existing knowledge. Weaknesses were detected in terms of missing data, poorly reliable data, too complex data collection mechanism and any other issue that can lead to poor clinic evidence. This is an input for the use cases in order to promote a data collection able to fill these gaps. Considering that we decided to do a very broad search for how PGHD is used in healthcare research, including all disease areas that are mentioned in the scientific research, with the consequence that we have collected a large amount of scientific studies to be analysed, we were not able to analyse all papers for this first deliverable. Taking into account that this report is a continuous process and the data is still under analyses, we will complete the excel over the upcoming months and provide an updated report in V2 of this deliverable.

2.8. Desk Search

In IMPROVE we complement the scientific research with an online desk research in order to get a more comprehensive understanding of the state of the art led by PBY. Desk research for patient-generated health data involves the collection, review, and analysis of information that is gathered through analyzing grey literature, white papers, European research projects deliverables, websites, etc., focusing predominantly on the 5 disease areas of IMPROVE. Unlike traditional primary research, which relies on direct data collection through experiments or interviews, desk research leverages existing datasets, reports, academic studies, and publicly available data to gain a comprehensive understanding of patient experiences without engaging directly with participants.

Patient-generated health data offers an intimate glimpse into patients' day-to-day lives, capturing nuances of their health that would often go unnoticed during clinical encounters. It encompasses data on symptoms, physical activity, medication adherence, diet, sleep patterns, and even mood, allowing researchers to see health in the context of a patient's real-life circumstances rather than the more controlled clinical settings. By engaging in desk research, analysts can systematically review the troves of secondary data to derive insights into the broader patient experience, exploring areas such as chronic disease management, treatment effectiveness, behavioural trends, and quality of life.

This type of research also has significant implications for understanding the impact of lifestyle changes, digital health tools, and self-care behaviours. Researchers can analyse these insights to identify correlations and common behaviours, enhancing understanding of patient engagement, technology usability, and intervention efficacy. By examining multiple sources of existing patient-generated data, researchers can also highlight gaps in care or unmet patient needs, offering valuable perspectives for healthcare providers and policymakers to improve services.

Moreover, desk research is particularly cost-effective and time-efficient when compared to traditional primary research. It allows researchers to examine data across a wide patient population without needing direct, labour-intensive recruitment efforts. This makes it highly scalable, offering the ability to explore a range of health topics that affect different demographic groups. The insights gained can inform the development of targeted interventions and the design of more patient-centred healthcare solutions. For example, understanding trends in wearable device data can help health systems develop proactive approaches to managing chronic conditions or identify early warning signs of health deterioration.

Another advantage of desk research with patient-generated health data is that it offers a diversity of perspectives. The data often reflects patient experiences across different settings, geographies, and contexts, which helps mitigate some of the biases inherent in traditional clinical studies. This diversity enables a broader understanding of health behaviours and outcomes and supports the tailoring of care to diverse populations. The real-world context of patient-generated data also allows for more dynamic insights, as patients are documenting their health conditions in natural environments without the constraints of a medical setting, leading to a more authentic representation of their health journey.

Overall, desk research for patient-generated health data represents a powerful approach to understanding patient behaviour and experience on a large scale. It offers a window into how people manage their health autonomously, revealing patterns that can guide personalized treatment strategies, influence the development of digital health tools, and improve overall health outcomes. By synthesizing and contextualizing these insights from existing data sources, coming from the screenathon for example, researchers and healthcare professionals can help shape a more patient-centred and proactive healthcare ecosystem.

3. Results

3.1. Screening

For an overview of the literature search and screening process reference Figure 2.

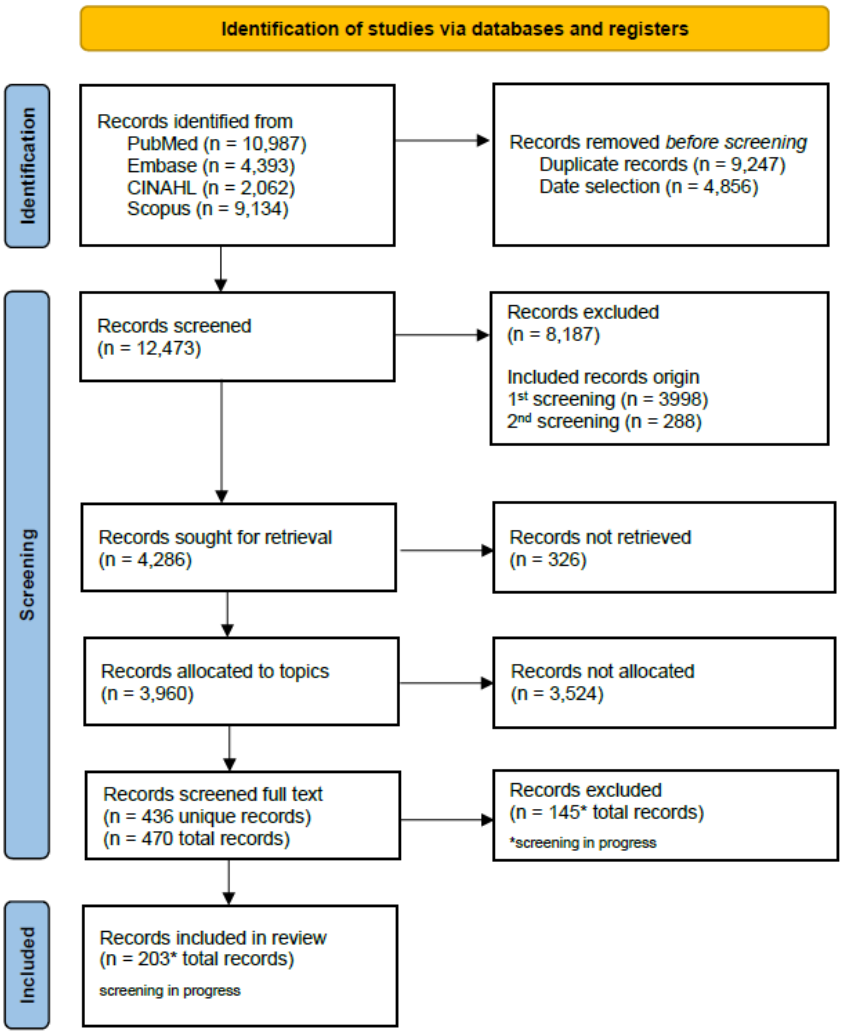

Figure 2 PRISMA Flowchart Summarizing Literature Search, Screening and Final Inclusions.

In total, 26 participants from 12 different organizations (universities as well as industry partners) took part in the Screenathon event. On day two, 53 batches were screened by 21 participants, and on day three, 72 batches were finished by 24 participants. At the end of day three, 125 batches had been completed by 26 participants, which meant our goal was met: we screened 12,473 records on TA in two days. Of these records, 8,475 (67.9%) were labelled 'irrelevant', and 3998 (32.1%) were labelled as 'relevant'.

The first quality check entailed a check for additional relevant records in the excluded records during the Screenathon event. This was carried out by a colleague using the noisy label filter (NLF) Procedure (Neeleman et al., 2024), which included the use of the screening-prioritization tool ASReview (van de Schoot et al., 2021). 456 records (144 relevant and 312 irrelevant records), all screened by a content expert participant during the Screenathon event, were used as prior information. Using feature extractor TF-IDF, classifier Naive Bayes, Query Strategy Maximum, and Dynamic Resampling as Balancing Strategy, another colleague started screening the records that were labelled as irrelevant before by all participants. However, following the NLF Procedure, they were now ranked based on relevance, so she would first screen the most likely relevant articles within the irrelevantly labelled articles. The colleague would stop screening after reaching 50 irrelevant TA labels in a row.

The second quality check was about the question if all cornerstone records were in the relevantly labelled data. These 12 cornerstone records were sent to us by content experts, who sent us lists of what they considered to be cornerstone records in the field. During quality check 1, 1,158 records were labelled in ASReview. 284 additional relevant records were found and were added to the dataset of relevant records. During quality check 2, we found that the 12 cornerstone records matched the dataset of relevantly labelled articles during the Screenathon event. No additional relevant articles had to be added to our dataset.

The 3,739 included papers from the Screenathon cover a broad range of disease topics that are outside the scope as defined above. That is, at this point of the project, only the papers that match our specified scope, in terms of disease topics, are of interest to us. To filter out the irrelevant papers, we used of OpenAlex's topic system (*Topics|OpenAlex Technical Documentation*, 2024), where OpenAlex uses a trained classifier to assign up to 3 most likely disease topics to every paper hosted on its platform. By using this topic system, we can quickly identify papers that are relevant for us per our defined scope of disease topics, without the need for manual screening or training a classification model on our own. To start with, we identified paper records on OpenAlex that match our included papers, based on their DOIs and paper titles. In total, we were able to identify 3,735 papers (of 3,739).

Then, we matched our defined list of disease topics to those of OpenAlex's. For every specific topic that we defined, we counted the number of papers that have been assigned a related OpenAlex topic. This means that a paper can also be associated with different topics and hence gets counted more than once. Among these papers, we identified in total 283 unique ones. These served as the basis for the full text review in the next step. An overview of the unique papers identified per disease topic can be found in Table 3.

Table 3 Overview of Papers to Disease Topic Matching.

Disease area	Topic	OpenAlex Topic(s)	# of Matches
Oncology	Prostate Cancer	Prostate Cancer Research and Treatment, Advancements in Prostate Cancer Research	55
	Cervical Cancer	Human Papillomavirus and Cervical Cancer Epidemiology	14
	Neck and Head Cancer	Epidemiology and Treatment of Head and Neck Cancer	56
	Breast Cancer	Molecular Research on Breast Cancer, HER2 Signaling in Breast Cancer Treatment, Cyclin-Dependent Kinase 4/6 Inhibitors in Breast Cancer, Male Breast Cancer and Gynecomastia Research, Breast Cancer Screening Technology	118
Cardiovascular Disease	Heart Failure	Diagnosis and Treatment of Heart Failure, Cardiac Resynchronization Therapy in Heart Failure	67
	Coronary artery diseases	Vascular Access for Coronary Procedures and Trauma Management, Clinical Studies on Coronary Stents and Revascularization	31
	Atrial Fibrillation	Atrial Fibrillation	24
	Severe Aortic Stenosis	Diagnosis and Management of Aortic Disease, Management and Pathophysiology of Abdominal Aortic Aneurysms, Diagnosis and Treatment of Renal Artery Stenosis	32
Ophthalmology	Macular Degeneration	Age-Related Macular Degeneration Research	13
Neurology	Multiple Sclerosis	Diagnosis and Pathogenesis of Multiple Sclerosis	32
Chronic Inflammation	Chronic Rhinosinusitis	Chronic Rhinosinusitis and Nasal Polyps	28

In Table 4 we show the total number of full manuscripts that have been analysed, the included articles that have been selected, and the number of individual studies that have been included in these literature reviews. This has been divided by each of the disease areas for the use cases to inform them and WP5 in more detail about the state of the art in scientific research.

Table 4 Full-text Screening Decision Frequencies per Topic.

Disease area	Use case related	Screened	Included	Nr. Individual studies
Oncology	Breast	118	70	2297
Oncology	Cervical	14	3	101
Oncology	Neck and head	56	34	777
Oncology	Prostate	55	41	TBA
Ophthalmology	Macular Degeneration	13	8	TBA
Neurology	Multiple Sclerosis	32	23	TBA
Cardiovascular	Atrial Fibrillation	24	20	TBA
Cardiovascular	Coronary Artery	31	15	TBA
Cardiovascular	Hearth Failure	67	45	TBA
Cardiovascular	Aortic Stenosis	32	5	386
Chronic Inflammation	chronic rhinosinusitis	28	19	364

TBA = To Be Analyzed

Subsequently we analysed the amount of studies that were included from each country in the studies that we have analysed; see Table 5. As you can see, the majority of the studies conducted were done in Western countries, like United States, Canada, Netherlands, Germany, United Kingdom, France, Italy and others. For Asian countries, most studies were included for China, were we found that a maximum of 10 studies were reported, and for Taiwan, Japan, Korea, and India much less. This means that there were almost no studies included for African countries or Latin America, which creates a significant bias in the interpretation of the findings. The same is found for Eastern European countries, that are mostly missing in the scientific databases.

Table 5 Full-text Screening Decision Frequencies per Topic for three of the use-cases.

Countries	Total	Oncology*	Chronic Inflammation	Cardiovascular disease*
United States	48	45	0	3
Canada	29	25	1	3
Australia	23	23	0	0
United Kingdom	20	20	0	0
Netherlands	18	16	0	2
Germany	16	14	0	2
China	14	13	1	0
Sweden	11	8	0	3
Italy	11	11	0	0
Taiwan	10	10	0	0
France	9	7	0	2
Korea	8	8	0	0
Brazil	7	7	0	0
Denmark	7	7	0	0
Norway	6	3	0	3
Japan	6	6	0	0
Remaining	53	50	0	3

Note: The numbers in this table give estimates. Data was not always provided or extracted in the same way. Additionally, this is the number of reviews the country appeared in; it might be that within the review the distribution among countries is very uneven.

3.2. Desk Research

The desk research on the use of PGHD provided an overview of how this data type is increasingly integrated into healthcare, touching on the benefits, challenges, and practical applications within clinical settings. PGHD, which includes data from devices like wearable fitness trackers, mobile health apps, and patient-reported outcomes, holds promise for improving health monitoring, personalized care, and preventive health measures. We will here report the outcomes of a first version of the desk research, analyzing non-scientific articles online that report the use of PGHD in healthcare,

Subsequently, we have established the most relevant European projects that work on PGHD that will be analyzed in more detail in D2.2.

3.2.1. Main Outcomes of the Desk Research on PGHD

3.2.1.1. Enhanced Personalization of Care

PGHD allows healthcare providers to track patient behaviours, symptoms, and outcomes beyond clinical visits that could facilitate better healthcare provision and more tailored and personalized care for patients that need it the most. This enables more personalized treatment plans, especially for chronic diseases such as diabetes, hypertension, and mental health conditions. In non-academic sources like healthcare technology blogs and industry reports (e.g., HIMSS, Health IT News), PGHD is frequently discussed as a tool for creating patient-centred care models that better reflect individual lifestyle and environmental factors. This could subsequently lead to better and more effective promotion of preventive care and self-management (Boumenot, 2024), empowering patients to take an active role in managing their health, as it provides real-time feedback on lifestyle choices, medication adherence, and symptom tracking. Technology reviews and news outlets, such as TechCrunch (Joshi, 2020) and Forbes Health, highlight the rapid growth of wearable health tech (e.g., smartwatches) and apps that facilitate self-monitoring, which has been shown to encourage healthier behaviours and prevent complications in chronic conditions.

3.2.1.2. Improved Health Outcomes and Cost Reduction

Studies indicate that continuous monitoring using PGHD can reduce hospital readmissions, emergency room visits, and healthcare costs, especially for at-risk populations. Reports from consulting firms, like McKinsey (Evers et al., 2022) and Deloitte (Siegel, n.d.), emphasize the potential for PGHD to support health system sustainability by reducing resource burdens through remote patient monitoring. Most importantly, they mention that better integration with Electronic Health Records (EHR) and Data Interoperability Challenges are considering to be essential for PGHD to be useful in clinical decision-making. However, technical and privacy barriers persist, including the lack of standardized formats and interoperable systems. In addition, there are several privacy, security, and ethical concerns when it comes to the collection, storage, and usage of PGHD. Patients and providers alike express concerns about privacy, data ownership, and the potential for misuse of sensitive health data. Regulatory frameworks, such as GDPR in Europe and HIPAA in the U.S., play a crucial role in shaping the ethical use of PGHD. Publications like Wired (*How a Data-Enabled Future Will Revolutionize Patient Experiences*, n.d.) and The Verge (Shakir, 2024) have examined the ethical dimensions of health data sharing, often emphasizing the need for transparent consent mechanisms and robust cybersecurity measures to maintain patient trust.

3.2.1.3. Data Quality and Reliability Issues

PGHD's effectiveness depends on the data quality, as inconsistencies in self-reported data and technical limitations in consumer health devices can undermine clinical utility and the use of the data. Subsequently, if the information is not so useful and effective, the motivation of patients and healthcare professionals will also decrease to continue collecting the data and use of it. In the end it

will lead to an increased provider workload and workflow integration, whereby many providers report that incorporating PGHD into routine care adds to their workload, as they must analyse and interpret additional data streams. Effective integration of PGHD into clinical workflows requires streamlining data processing, often through AI or other automation solutions. Health informatics platforms, such as HealthIT.gov, and industry-focused publications as have been mentioned before have explored this challenge, stressing the importance of training and supportive technology for providers to manage PGHD effectively. The research also underscores practical implications for healthcare systems. Policymakers and healthcare leaders are encouraged to invest in infrastructure that supports data integration, protects patient rights around data usage, and incentivizes technology development that aligns with clinical needs.

In sum, while PGHD offers significant potential benefits for improving healthcare delivery and patient outcomes, it also presents challenges related to data integration, security, and clinical usability that need addressing to realize its full potential. Non-academic sources provide a valuable perspective on the practical realities and ongoing developments in health tech, policy discussions, and consumer adoption of PGHD technologies.

Subsequently, for the desk research we have identified a large number of initiatives and research projects that focus on the use of PGHD in different disease areas and different steps of the treatments and diagnosis (N=259) across countries and regions. As an outcome of the desk research we have found a large number of European (e.g., H2020s, IHI, Horizon Europe's) and national projects and partners have already been identified, see Appendix A for the overview of recent EU-projects and beyond, associations that are linked to IMPROVE and other organizations that work in similar areas. In *Deliverable 2.2 Practice report and updates V1* we have analyzed a first selection of five of these projects in more detail to use the most important outcomes in order to develop a knowledge base of the existing practices that are conducted to develop methods or frameworks for collecting and using patient-reported outcomes. Subsequently, data gathering will be done in existing repositories of good practices in different fields. In the upcoming months, we will approach these organizations and projects to start collaborations and improve the way we work by having open conversations.

3.3. Gap Analysis

In the end, the IMPROVE project is dedicated to harnessing the potential of PGHD through the use of m-health and e-health technologies. This project aims to bridge the current gaps in data utility and fragmentation by integrating and enhancing insights into the daily lives and challenges of patients across all ages who suffer from complex, chronic diseases and comorbidities. By doing so, IMPROVE seeks to extend the capabilities of existing platforms and approaches to Patient-Centred Outcome Measures, enriching them with real-world data that reflect true patient experiences and preferences.

At the core of IMPROVE is the development of a robust platform designed to enable the intelligent use of patient input and generated evidence. This platform will facilitate three key advancements:

- **Enhancing treatment selection**: By advancing the role of patient preferences and experiences in choosing treatments, thereby personalizing healthcare to meet individual needs more effectively.
- **Medical device design improvement**: By incorporating patient feedback directly into the design process, ensuring that new medical devices are more aligned with user expectations and experiences.
- **Accelerating market entry**: By speeding up the introduction of patient-centric and cost-effective advanced integrated care solutions, thus enhancing the accessibility of innovative treatments.

Considering the project aims to demonstrate the improved clinical adoption of Value-Based Health Care (VBHC) and a higher return on research and innovation investments across various European care settings it is important to analyse whether the scientific literature and evidence generated until now accommodates this. In the literature we have mostly found studies that focus on the improvement of treatment selections and giving patients a more comprehensive understanding of the outcomes of treatments on patients.

Enhancing treatment selection

The integration of PROMs, PREMs, and PPIs into healthcare decision-making has the potential to significantly improve treatment selection by ensuring that patient-centred care is prioritized. These tools and frameworks provide a deeper understanding of patient needs, preferences, and experiences, which are critical for tailoring treatments to achieve optimal outcomes.

For example, PROMs capture a patient's self-reported health status, symptoms, and quality of life before, during, and after treatment. This data allows clinicians to better understand how different treatments impact patient well-being and functional outcomes. By incorporating PROMs into decision-making, healthcare providers can select treatments that are not only clinically effective but also align with the patient's specific health goals and priorities. For instance, if a patient values mobility and pain management over other outcomes, treatments that excel in those areas can be prioritized. PREMs provide insights into the patient's experience with healthcare delivery, including access, communication, and overall satisfaction. These measures highlight systemic and interpersonal factors that may influence treatment adherence and effectiveness. For example, if patients consistently report

difficulties with a particular mode of treatment delivery, alternative approaches may be considered to improve both their experience and outcomes. Integrating PREMs ensures that treatment plans are designed with an understanding of patient preferences and barriers, fostering better engagement and satisfaction. PPIs will provide more information about a certain preference of patients, but this assumes patients have all information, understand this information, and can also make a decision based on this information, presuming they are rational beings that calculate costs and benefits.

The combination of PROMs, PREMs, and PPI creates a holistic framework for treatment selection. While PROMs provide objective data on health outcomes, PREMs highlight subjective experiences, and PPI ensures that the patient's preferences are central to decision-making. Together, they enable a personalized, evidence-based approach that improves the alignment of treatments with individual needs, enhances satisfaction, and fosters better long-term health outcomes. Incorporating these patient-centred measures into routine clinical practice not only refines treatment selection but also advances the principles of shared decision-making. This ensures that care is truly tailored, equitable, and responsive to the complexities of patient lives. In general, through active collaboration, patients can voice their priorities and provide insights into how treatments affect their lives in ways that might not be captured by clinical metrics alone. This input is invaluable in the selection of treatments that respect patient values, improve adherence, and enhance the overall relevance of care. For example, involving patients in the design of clinical pathways can help identify treatments that are both feasible and acceptable within their specific contexts. During the analyses of the literature review we have not found sufficient evidence how this can be done over all the different diseases and what is the most effective framework to do so. Most research that was focused on improving the treatment selection did not combine different measurements such as PROMs, PREMs and PPIs or other forms of PGHD, and also did not follow patients longitudinal whether they made different selections or preferences after gather insights from other patients that had the treatment before them.

Medical device design improvement:
Considering the scientific evidence, we conclude that not so much is reported about how to improved medical devices or the design of them. In the literature we focused mainly on the systematic literature reviews, whereby this was not investigated in the studies we selected. In the 2nd year of this project, we will focus on these elements for some disease areas to ensure we have a good understanding on the scientific literature on this aspect. It is important to incorporate patient feedback directly into the design process, thereby ensuring that new medical devices are more aligned with user expectations and experiences and also providing opportunities for accelerating market entry.

Accelerating market entry:
For the faster acceleration of the market entry, we also did not find any relevant scientific papers studying how PGHD can be used to accelerate the market entry of cost-effective solutions. The more efficient PGHD can be collected, such as automatized data collection from patients using digital tools, preferably automatically generated, this will provide a faster acceleration of (cost-)effective solutions to bring them on the market. This will be under more thoroughly investigation for the next

3.4. Quantitative analyses

In Table 6 we analyse the scientific outputs from the systematic literature review to see if the state of the art provides the IMPROVE with a better understanding of the implementation. As the results show, there are some difficulties on the reliability of extracting knowledge of dimensions like PPI and PREM. PREMs and PPIs are a very relevant dimensions because in IMPROVE we would like to follow the patient-centred approach, which has not been investigated in previous studies.

Table 6 Scientific Outputs from the Systematic Literature Review.

Main topic	Subtopic	PREMs	PROMs	PPIs	Pre	During	Post
Cancer	Breast	1	66	0	14	18	64
Cancer	Cervical	0	3	1	3	3	3
Cancer	Neck and head	2	31	2	20	21	31
Cancer	Prostate						
Ophthalmology							
Neurology							
Cardiovascular	Atrial Fibrillation						
Cardiovascular	Coronary Artery						
Cardiovascular	Hearth Failure						
Cardiovascular	Aortic Stenosis	1	1	2	1	0	1
Chronic Inflammation					6	1	15

Note: The numbers in this table give estimates. Extraction was done by different people and can contain different definitions. Additionally, data was not always provided. Empty cells: extraction in progress.

4. Conclusion and Next Steps

The systematic umbrella review conducted in this study serves as a pivotal resource for bridging existing evidence gaps, refining the relevance and precision of insights, and accelerating the development of actionable use cases. This comprehensive approach underscores the critical importance of methodical and evidence-based strategies in fostering innovation and addressing the multifaceted challenges inherent in healthcare delivery, especially considering only a very limited amount of studies have focused on PREMs and PROMs. In addition, not so many studies have been conducted on the use of PGHD on medical device design improvement and accelerating market entry. By systematically identifying and analysing cornerstone systematic reviews and meta-analyses, we have been able to construct a detailed and holistic overview of the current evidence regarding the use of various types of Patient Generated Health Data (PGHD). This spans their integration across multiple disease areas, the stages of the patient journey, and their implications for value-based healthcare. These findings lay the groundwork for subsequent steps, including the development of the IMPROVE lab, coordination of data collection in WP4 and WP5, and the creation of guidelines and best practices in WP7.

There are five important next steps, see Figure 3.

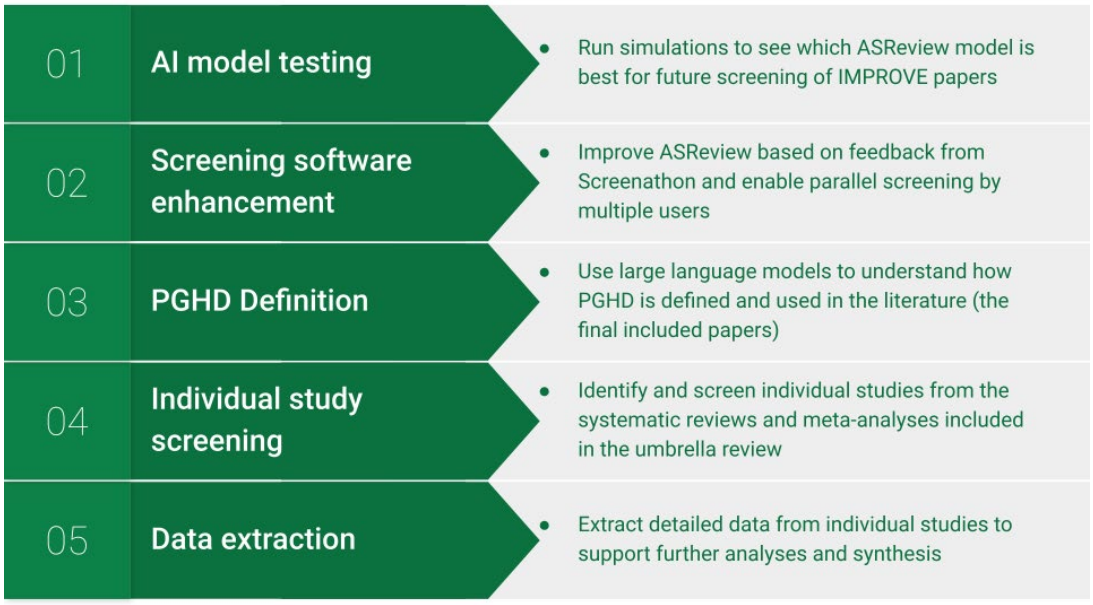

Figure 3 Next steps for WP2.

First, to automate the yearly update process, UU and Dedalus will leverage the human-labelled data in the database to test and evaluate different AI models. Using ASReview's simulation functionality, we will simulate the screening process by comparing the performance of various machine learning models on our labelled dataset. This will allow us to identify the model that is most effective at prioritizing relevant papers while minimizing screening effort. ASReview's simulation capabilities enable us to mimic the real-world screening process, providing insights into how the selected model would perform with unseen data. We will utilize Makita, a workflow manager integrated into ASReview, to streamline and automate these simulations. Makita simplifies the process of setting up, running, and analysing

multiple simulation experiments by offering reproducible pipelines. This ensures that our evaluation of AI models is not only robust but also scalable for future updates. Once the optimal model is identified, it will be implemented in the open-source ASReview software to optimize the screening effort in Year 2, allowing us to efficiently identify and include the most relevant new evidence as it becomes available.

Second, UU will enhance the software to provide a more seamless and efficient experience for crowdsourcing the screening process in the upcoming year. One key improvement will be the automation of database management and the strategic assignment of papers to individual screeners. By intelligently distributing papers for review, we aim to ensure that all screeners contribute to training the same AI model collaboratively. This approach will help maintain consistency in labelling and optimize the training data for the AI model and by integrating these features into the platform, we can significantly reduce screening time while maintaining high-quality results for the yearly update process.

Third, leveraging the systematic reviews identified in the current study, UU and Dedalus aim to harness Large Language Models (LLMs) trained on the included studies to develop a data-driven understanding of how to define Patient Generated Health Data (PGHD). This process will involve several steps:

1. **Named Entity Recognition (NER):** Using NER techniques, we will identify key terms, concepts, and stakeholders mentioned in the included studies. This will provide a structured foundation for analysing the diverse ways PGHD is described and contextualised in the literature.
2. **Extracting Relationships Among Entities:** By analysing the identified entities, we will map the relationships between them. This step is essential for uncovering the interplay between different components of PGHD, such as data sources, applications, and stakeholders.
3. **Determining Related Terms of PGHD:** Through co-occurrence analysis and semantic similarity techniques, we will identify terms closely related to PGHD. These terms will provide insights into its broader context and nuances across different domains and applications.
4. **Extracting Contexts of Key Terms:** Using text mining methods, we will extract the contexts in which PGHD-related terms are used. This step will reveal patterns and variations in how these terms are applied, enabling a richer understanding of their meanings.
5. **Applying Topic Modelling or Clustering Techniques:** We will use advanced clustering methods, such as topic modelling, to group related contexts and identify thematic trends. This will help in categorizing the diverse definitions and uses of PGHD found in the literature.
6. **Synthesizing Definitions and Uses of PGHD:** Finally, we will employ LLMs to synthesize definitions and uses of PGHD across the studies. By analysing the extracted contexts and thematic clusters, the LLMs will generate a consolidated understanding of PGHD that captures its multifaceted nature.

This systematic, data-driven approach will not only provide a comprehensive definition of PGHD but also offer insights into its various applications and implications, facilitating its integration into value-based healthcare frameworks.

Fourth, to complement the existing gap analyses, future systematic reviews conducted by the entire WP2 team (led by the UU), should prioritize the individual randomized controlled trials (RCTs) and other experimental studies specific to each use case. If these are not available observational studies or other studies should be analysed to have a better understanding about the use of PGHD in these disease areas. These methodologies provide robust, high-quality evidence that can directly inform use case development by addressing specific research questions with precision and rigor. Systematic reviews focusing on the individual RCTs enable a deeper understanding of causality and efficacy within the context of each use case.

Fifth, by extracting and synthesizing findings from well-designed experimental studies, such reviews can highlight the most effective interventions, technologies, or methodologies tailored to the unique requirements of each application, including the use of PGHD and how they are evaluated. Moreover, targeted reviews of experimental studies contribute to refining the theoretical frameworks and practical applications within the use case. They identify not only successes but also limitations and contextual factors that influence outcomes. This specificity ensures that recommendations are not only evidence-based but also contextually relevant and adaptable to real-world conditions. By focusing systematically on experimental research, these reviews also address existing knowledge gaps that may have been identified in earlier scoping reviews or exploratory analyses. They ensure a progressive alignment between empirical evidence and the strategic development of use cases, guiding stakeholders in making informed decisions about policy, design, and implementation.

In summary, these five steps will serve as critical complements to the outcomes of the gap analyses, we will provide a better answer on how PGHD can inform enhancing treatment selection (advancing the role of patient preferences and experiences in choosing treatments, thereby personalizing healthcare to meet individual needs more effectively), the medical device design improvement (by incorporating patient feedback directly into the design process, ensuring that new medical devices are more aligned with user expectations and experiences) and accelerating market entry (by speeding up the introduction of patient-centric and cost-effective advanced integrated care solutions, thus enhancing the accessibility of innovative treatments), especially in the disease areas that are central to IMPROVE.

References

Biometrics—An overview | ScienceDirect Topics. (n.d.). Retrieved November 21, 2024, from https://www.sciencedirect.com/topics/computer-science/biometrics

Boetje, J., & Van De Schoot, R. (2024). The SAFE procedure: A practical stopping heuristic for active learning-based screening in systematic reviews and meta-analyses. *Systematic Reviews*, *13*(1), 81. https://doi.org/10.1186/s13643-024-02502-7

Boumenot, P. (2024). *Why it's crucial to empower patients with data and digital access*. Health Data Management. https://www.healthdatamanagement.com/articles/[object Object]

Chandler, J., & Hopewell, S. (2013). Cochrane methods—Twenty years experience in developing systematic review methods. *Systematic Reviews*, *2*(1), 76. https://doi.org/10.1186/2046-4053-2-76

European Health Data Space—European Commission. (2024, April 24). https://health.ec.europa.eu/ehealth-digital-health-and-care/european-health-data-space_en

Evers, M., Pérez, L., Robke, L., & Smietana, K. (2022). *Better therapies need better data: The case for building health-data platforms | McKinsey*. https://www.mckinsey.com/industries/life-sciences/our-insights/better-data-for-better-therapies-the-case-for-building-health-data-platforms

Gough, D., & Elbourne, D. (2002). Systematic Research Synthesis to Inform Policy, Practice and Democratic Debate. *Social Policy and Society*, *1*(3), 225–236. https://doi.org/10.1017/S147474640200307X

Health Technology—An overview | ScienceDirect Topics. (n.d.). Retrieved November 21, 2024, from https://www.sciencedirect.com/topics/medicine-and-dentistry/health-technology

Higgins, J., Thomas, J., Chandler, J., Cumpston, M., Li, T., Page, M., & Welch, V. (Eds.). (2024). *Cochrane Handbook for Systematic Reviews of Interventions*. Cochrane. www.training.cochrane.org/handbook

How a Data-Enabled Future Will Revolutionize Patient Experiences in Healthcare. (n.d.). *Wired*. Retrieved November 21, 2024, from https://www.wired.com/sponsored/story/patient-experiences-healthcare/

Joshi, R. (2020, May 7). Health APIs usher in the patient revolution we have been waiting for. *TechCrunch*. https://techcrunch.com/2020/05/07/health-apis-usher-in-the-patient-revolution-we-have-been-waiting-for/

Koehring, M. (2015). *An introduction to value-based healthcare in Europe*. Economist Impact - Perspectives. https://impact.economist.com/perspectives/health/introduction-value-based-healthcare-europe

Lefebvre, C., Eisinga, A., McDonald, S., & Paul, N. (2008). Enhancing access to reports of randomized trials published world-wide – the contribution of EMBASE records to the Cochrane Central Register of Controlled Trials (CENTRAL) in The Cochrane Library. *Emerging Themes in Epidemiology*, *5*(1), 13. https://doi.org/10.1186/1742-7622-5-13

Monschau, K., Neeleman, R., Jalsovec, E., de Bruin, L., de Bruin, J., Weijdema, F., Folkvord, F., Peschke, L., Westerbeek, E., & Van De Schoot, R. (2024). *Introducing the Screenathon Review: Crowdsourcing Literature Screening in Large-Scale Research Collaborations*. PsyArXiv. https://doi.org/10.31234/osf.io/vxg6d

Neeleman, R., Leenaars, C. H. C., Oud, M., Weijdema, F., & Van De Schoot, R. (2024). Addressing the challenges of reconstructing systematic reviews datasets: A case study and a noisy label filter procedure. *Systematic Reviews*, *13*(1), 69. https://doi.org/10.1186/s13643-024-02472-w

Patient-Reported Experience Measure (PREM). (n.d.). YHEC - York Health Economics Consortium. Retrieved November 21, 2024, from http://yhec.co.uk/glossary/patient-reported-experience-measure-prem/

Porter, M. E. (2010). What Is Value in Health Care? *New England Journal of Medicine*, *363*(26), 2477–2481. https://doi.org/10.1056/NEJMp1011024

Russo, S., Jongerius, C., Faccio, F., Pizzoli, S. F. M., Pinto, C. A., Veldwijk, J., Janssens, R., Simons, G., Falahee, M., De Bekker-Grob, E., Huys, I., Postmus, D., Kihlbom, U., & Pravettoni, G. (2019). Understanding Patients' Preferences: A Systematic Review of Psychological Instruments Used in Patients' Preference and Decision Studies. *Value in Health*, *22*(4), 491–501. https://doi.org/10.1016/j.jval.2018.12.007

Shakir, U. (2024, October 25). *UnitedHealth data breach leaked info on over 100 million people*. The Verge. https://www.theverge.com/2024/10/25/24279288/unitedhealth-change-breach-100-million-leak

Siegel. (n.d.). *2024 Global Health Care Sector Outlook*. Retrieved November 21, 2024, from https://www.deloitte.com/global/en/Industries/life-sciences-health-care/analysis/global-health-care-outlook.html

Teisberg, E., Wallace, S., & O'Hara, S. (2020). Defining and Implementing Value-Based Health Care: A Strategic Framework. *Academic Medicine: Journal of the Association of American Medical Colleges*, *95*(5), 682–685. https://doi.org/10.1097/ACM.0000000000003122

Tian, S., Yang, W., Grange, J. M. L., Wang, P., Huang, W., & Ye, Z. (2019). Smart healthcare: Making medical care more intelligent. *Global Health Journal*, *3*(3), 62–65. https://doi.org/10.1016/j.glohj.2019.07.001

Topics | OpenAlex technical documentation. (2024, May 9). https://docs.openalex.org/api-entities/topics

van de Schoot, R., de Bruin, J., Schram, R., Zahedi, P., de Boer, J., Weijdema, F., Kramer, B., Huijts, M., Hoogerwerf, M., Ferdinands, G., Harkema, A., Willemsen, J., Ma, Y., Fang, Q., Hindriks, S., Tummers, L., & Oberski, D. L. (2021). An open source machine learning framework for efficient and transparent systematic reviews. *Nature Machine Intelligence*, *3*(2), 125–133. https://doi.org/10.1038/s42256-020-00287-7

Williams, K., Sansoni, J., Darcy, M., Grootemaat, P., & Thompson, C. (2016). *Patient-reported outcome measures*. Australian Commission on Safety and Quality in Health Care.

Winter, J. S., & Davidson, E. (2022). Harmonizing regulatory regimes for the governance of patient-generated health data. *Telecommunications Policy*, *46*(5), 102285. https://doi.org/10.1016/j.telpol.2021.102285

About IMPROVE

IMPROVE aims to be a dynamic, ready-to-use framework for seamlessly integrating patient-reported information. This adaptable system constantly evolves with the latest evidence, using PGHD and health system data to provide cost-effective solutions for diverse treatment conditions in real settings. The project follows Ontology, Epistemology, and Methodology principles. Ontology defines structures in patient-reported outcomes; Epistemology ensures valid knowledge; Methodology links techniques to outcomes, systematically addressed in its work.

IMPROVE optimizes patient-reported information in real settings, offering a deep understanding of patient behaviors. The project sets up ontology, epistemology, and methodology to minimize the burden on stakeholders cost-effectively. It adopts a scalable, data-driven approach with NLP-driven knowledge extraction. Real World Data is integrated into the Federated Causal Evidence module for comprehensive understanding. Evidence collected enables visualizing attributes affecting patient-reported outcomes through IMPROVE Engagement Factors and Indicators Knowledge Graphs.

IMPROVE's toolkit includes resources for decision-makers, featuring plausible scenarios via the Copenhagen Method. Patient engagement via the MULTI-ACT model ensures sustainable healthcare aligned with patient priorities. This project delivers a modular, open access strategy, providing a trustworthy ecosystem of evidence-based applications. Patient engagement and co-creation scenarios solidify its role in transforming healthcare research and care.

Funding Acknowledgement

This project is supported by the Innovative Health Initiative Joint Undertaking (IHI JU) under grant agreement No. 101132847. The JU receives support from the European Union's Horizon Europe research and innovation program and COCIR, EFPIA, EuropaBio, MedTech Europe, Vaccines Europe, and the contributing partners Universidad Politecnica de Madrid (UPM), PredictBy (PBY), Danish Medicine Agency (DKMA), Roche (ROCHE), Institute for Economic Research (IER), Copenhagen Institute for Futures Studies (CIFS), Fundació Institut d'Investigació Biomèdica de Bellvitge (IDIBELL), Philips Medical System Nederland BV (PMSN), Heinrich-Heine-Universitaet Duesseldorf (UDUS), Tilburg University (TiU), Dedalus (DEDA), Fondazione Italiana Sclerosi Multipla Fism Onlus (FISM), AReSS Puglia (ARSS), MultiMed (MM), iserundschmidt GmbH (ius), Better (BET), The Netherlands Cancer Institute (NKI), University of Applied Sciences St. Pölten (STPUAS), Eye Hospital, University Medical Centre Ljubljana (EYE), Utrecht University (UU), UDG Alliance (UDGA), Medtronic Iberica SA (MDT), Fundacio Hospital Universitari Vall D'Hebron – Institut de Recerca (VHIR), Splosna Bolnisnica Celje (SBC), ORTOPEDSKA BOLNIŠNICA VALDOLTRA (OBV), ETHNIKO KENTRO EREVNAS KAI TECHNOLOGIKIS ANAPTYXIS (CERTH).

www.ihi.europa.eu

Supporters of the Innovative Health Initiative Joint Undertaking:

Project partners:

Coordinator

Associated Partner