xShare

Expanding the European EHRxF to share and effectively use health data within the EHDS

Working Paper: *Proposal* for a harmonized core data set across health care, population health and clinical research

Project title: xShare - Expanding the European EHRxF to share and effectively use health data within the EHDS.

Grant Agreement: 101136734

Call identifier: HORIZON-HLTH-2023-IND-06-02

Dissemination level: Public

This project has received funding from the European Health and Digital Executive Agency (HADEA) under grant agreement no. 101136734.

Working paper description

Number and name of working paper:	Proposal for a harmonized core data set across health care, population health and clinical research-WP5-CDISC
Publishable summary:	This This working paper (D5.1) reports on an analysis of the data elements from a number of sources to identify and to some degree align the healthcare/IPS data elements of greatest research value and proposes the data elements that could usefully be added to the existing IPS in order to maximise its value for public/population health and clinical research in addition to its primary value for healthcare. This core data element set will be subject to consultation by stakeholders, and then, aligned CDISC and HL7 FHIR implementation guides will be produced. This will set the groundwork for IPS+R and for additional health information domains, new or extended priority data categories in the EHDS.
Editors	Rebecca Baker and Rebecca Kush (CDISC), Dipak Kalra (i~HD), Catherine Chronaki (HL7 Europe)

Statement of originality

This working paper contains original unpublished work except where clearly indicated otherwise.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.

Disclaimer

Funded by the European Union. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or European Health and Digital Executive Agency (HADEA). Neither the European Union nor the granting authority can be held responsible for them.

Table of Contents

List of abbreviations

All abbreviations included here can also be found in the xShare online glossary: https://glossary.ramit.be/public/home.cfm?pid=11

Abbreviation	Term
ACT	Accrual to Clinical Trials
ADaM	Analysis Data Model

BR&R	Biomedical Research and Regulation
BRIDG	Biomedical Research Integrated Domain Group
CDASH	Clinical Data Acquisition Standards Harmonization
CDE	Common Data Element
CDISC	Clinical Data Interchange Standards Consortium
CDMH	Common Data Model Harmonization
CFAST	Coalition for Accelerating Standards and Therapies
C-Path	Critical Path Institute
EC	European Commission
EEHRxF	European electronic health record exchange Format
EHDS	European Health Data Space
EHR	electronic health record
EHR2EDC	electronic health record systems to electronic data capture project
EHR4CR	electronic health record systems for clinical research project
EVS	enterprise vocabulary services
FDA	Food and Drug Administration of the United States
FHIR	Fast Healthcare Interoperability Resources
FHIR IG	FHIR implementation guide
G7	group of seven
GDHP	Global Digital Health Partnership
HHS	US Department of Health and Human Services
HL7	Health Level Seven International
i2b2	Informatics for Integrating Biology & the Bedside
ICD	International Classification of Diseases
ICH	International Conference on Harmonisation
ICH M11	Clinical Electronic Structured Harmonised Protocol Template
IDDO	Infectious Diseases Data Observatory
IEC	International Electrotechnical Commission
IHE	Integrating the Healthcare Enterprise
IHI	Innovative Health Initiative
IMI	Innovative Medicines Initiative
IPS	International Patient Summary
IPS+R	International Patient Summary plus for Research
ISO	International Organisation for Standardization
JSON	JavaScript Object Notation
LOINC	Logical Observation Identifiers Names and Codes
MedDRA	Medical Dictionary for Regulatory Activities
MoH	Ministry of Health
NCATS	National Center for Advancing Translational Sciences
NCI	National Cancer Institute
NDC	National Drug Code
NIAID	National Institute of Allergy and Infectious Disease
NICHD	National Institute of Child Health and Human Development
NIDDK	National Institute of Diabetes and Digestive and Kidney Diseases
NIH	National Institutes of Health
NIMH	National Institute of Mental Health
NINDS	National Institute of Neurological Disorders and Stroke

OHDSI	Observational Health Data Sciences and Informatics
OMOP	Observational Medical Outcomes Partnership
ONC	Office of the National Coordinator for health information technology
PCORI	(US) Patient-Centred Outcomes Research Initiative
PCORnet	Patient-Centered Clinical Research Network
PhPID	pharmaceutical product identifier
PMDA	Pharmaceuticals and Medical Devices Agency of Japan
RCRIM	HL7 Regulated Clinical Research Information Management Working Group
RWD	real world data
SDTM	Study Data Tabulation Model
SNOMED	Systematized Nomenclature Of Medicine
STU	HL7 standard for trial use
TA	Therapeutic Area
UDP	HL7 Vulcan Utilizing the Digital Protocol
UNICOM	Up-scaling the global univocal identification of medicines
USDM	Unified Study Definitions Model
WHO	World Health Organisation
WP	work package
XML	extensible markup language

List of Tables

List of Figures

Executive summary

The Regulation on a European Health Data Space (EHDS), which has recently (24 April 2024) been approved by the European Parliament and the Council[1], will change the status of the European electronic health record exchange format (EEHRxF) to mandatory for adoption by all Member States. Member States will need to establish ways of consolidating this data at a national level, per patient, to break down the data fragmentation that presently exists, and to enable rapid communication of this patient summary data between Member States in the case of a cross-border healthcare need. The EEHRxF will be the mandated standard, although the detailed specifications for the data representations and semantic interoperability layers will be specified in forthcoming secondary legislation (Implementing Acts) over the next couple of years. It is widely expected that the 10-year-old European patient summary guidelines implementation will be gradually aligned with the ISO International Patient Summary (IPS), with the added benefit that this aspect of the EEHRxF will provide input to international standards. Moreover, the decision was taken that future domains in myHealth@EU will follow the HL7 FHIR standard.

Clinical research and public health are already international in scope and have been exploring ways to leverage healthcare/EHR data (also called Real World Data) for their purposes for the past several decades. The xShare project is positioned keenly to lead the necessary initiatives to develop a core global set of EHR data elements that can streamline research and public health data streams, reducing resource needs and costs while increasing quality and timeliness of information.
This working paper (xShare D5.1) reports on an analysis of the data elements from a number of sources to identify and to some degree align the healthcare/IPS data elements of greatest research value and proposes the data elements that could usefully be added to the existing IPS to maximise its public/population health and clinical research value. The resulting core set will pave the way towards the aligned International Patient Summary plus for Research (IPS+R). This document presents the methodology adopted for a thorough analysis, including descriptions of the various data element sources and a summary of related background research projects, initiatives and standards that have been synthesised to compile the proposed research-relevant data element set. Also included are considerations for the recommended levels of granularity down to the data element concept level. This working paper stops short of assigning recommended terminologies/vocabularies/value sets or codelists: these are essential for semantic interoperability, but it is recommended that consensus on the core data elements be reached first. This document includes a table with the recommended core data elements and definitions and a set of depictions of the recommended elements sorted by IPS Category and the aligned Research Domain.

It is recommended that the core IPS data elements (approximately 90) be extended by a limited number of elements (approximately 10) that would be critical for research, including Research Subject Identifier, which would indicate that the patient is in a clinical research study. It is also proposed that information on Adverse Events be considered; this would add another set of data elements. Depending on the use case and implementation process for data collection of adverse event data, it is possible that the burden on the healthcare clinician could be minimized in this case. The next steps within xShare should include extensive reviews of the IPS+R core data element set by xShare partners and external experts to iterate and build consensus. At the same time, we will explore how these core elements are to be represented in HL7 FHIR resources. After this, there will

[1] The European Commission. European Health Data Space. 2024. Available from https://health.ec.europa.eu/ehealth-digital-health-and-care/european-health-data-space_en

need to be an effort to include requirements for controlled terminology, codelists and/or valuesets to make it possible to exchange the key data while preserving the meaning (semantic interoperability) and ensuring quality. A WP5 working paper in the coming months will be a more detailed specification for these elements, including terminology, valuesets or codelists, as appropriate. In addition, many of the 'optional' elements in IPS should be 'required' to meet the needs of the research and public health communities. The finally agreed inventory of data elements needed for research or public health that are not already required for the IPS will be proposed as IPS extensions.

Eventually, the success of xShare will provide the EEHRxF and IPS with standard data elements and terminologies to support semantic interoperability for healthcare, research, and public health. The xShare core data element set (IPS+R) will be designed to support research and public health; thus, it should also represent a global standard. Clinical research and the associated data standards are globally harmonized and have been so for many years. Public health also must act internationally, especially in cases of outbreaks, as should healthcare. When someone travels to a different country, they will benefit if their healthcare summary can be readily available. In addition, in a true learning health system, every person is essentially a research subject.

At some point, it is hoped that, as with clinical research data standards, the core xShare data element set will set a global standard and that it will eventually be extended to support paediatrics as well as numerous therapeutic areas. Mapping it to HL7 FHIR and ensuring a clear path to adoption in public/population health data sets would be instrumental in increasing the value for EHRxF and the EHDS as a whole, creating valuable bridges between health and care, clinical research and public health and population health.

1 Introduction

1.1 The objective and structure of this working paper

Electronic health records (EHRs) have important value to clinical research as well as direct patient care. There are now extensive findings that indicate which data elements are most relevant and useful to key research use cases, independently of the specific therapeutic area of intended research (i.e. cross-cutting high value data elements)[2]. The availability and quality of these data elements within EHRs varies considerably.

National health systems and transnational bodies like the European Commission are increasingly making use of patient health summaries as a way of prioritising the quality and interoperability of key data elements that can support continuity of care, especially in situations where a patient is needing healthcare in locations that do not have access to their complete EHR, such as urgent unplanned care including cross-border or cross-jurisdictional situations. The latest specification for this kind of patient summary is being published by ISO as the International Patient Summary (IPS). The latest version is internal to ISO members and has not yet been published. There are also IPS specifications for HL7 FHIR IPS, IHE IPS, and European Patient Summary used in the cross-border exchange of patient summaries in myhealth@EU using HL7 CDA. Members of the Joint Initiative Council for Global Health Informatics Standardization have invested significant effort in their alignment. As these specifications are progressively adopted by national health systems and aligned by their exchange in transnational data ecosystems such as the European Health Data Space, investments in the quality and interoperability of this dataset are expected to grow as a priority over the whole EHR data. There is therefore a potential advantage for clinical research and population health in aligning its interests in reusing EHR data with the content of the IPS. Reciprocally, it is in the interest of clinical research and population health to compare the data elements of greatest value for research with the IPS, in order to identify additional data elements that it could propose to ISO and to adopting countries, that would extend the IPS and greatly increase the value of national repositories of IPS-conformant summaries to clinical research.

This working paper reports on this analysis of the data elements of greatest research and population health value and carries out a gap analysis proposing the data elements that could usefully be added to the IPS to maximise its research value.

The rest of this chapter summarises the case for extending the IPS to better support clinical research and public/population health. Chapter 2 presents the methodology adopted for the data element analysis, including a summary of the background research projects, initiatives and standards that have been synthesised to compile the proposed core data set. Chapter 3 lists the core data elements and discusses some design considerations that were considered. Chapter 4 proposes *the core data element set* and shows the elements in a table and by category/domain. A gap analysis between IPS and the core data element set that includes additional elements useful for research is included in Chapter 5. It outlines the next steps in the work to be undertaken by xShare in relation to establishing consensus on core data element set across healthcare, public health and clinical research. This work will provide input to Task 5.3 *Harmonisation of data element requirements to*

[2] Leavy MB, Swenson A. Data Sources. In: Gliklich RE, Leavy MB, Dreyer NA, editors. Tools and Technologies for Registry Interoperability, Registries for Evaluating Patient Outcomes: A User's Guide, 3rd Edition, Addendum 2 [Internet]. Rockville (MD): Agency for Healthcare Research and Quality (US); 2019 Oct. Chapter 2. Available from: https://www.ncbi.nlm.nih.gov/books/NBK551884/

establish IPS+R, an extension of IPS for research and its planned interactions with other work packages. Chapter 6 concludes the reported work. A bibliography is given in Chapter 7.

1.2 The case for reusing EHRs for clinical research

EHRs are the most important, rich sources of information about patients, collected and curated by healthcare provider organisations and increasingly shared between them and with national eHealth platforms. There is now substantial evidence that this routinely collected data (also known as Real World Data, RWD) within EHRs can play a valuable role in the design and initiation of clinical research studies, including (i) optimising eligibility criteria[3] [4], (ii) identifying sites that have viable patient numbers matching those criteria, (iii) using the site EHR data to accelerate recruitment of patients into clinical trials,[5] [6] while (iv) minimising duplication of data collection efforts between healthcare and research[7]. Interest is also growing in the use of EHR data to generate virtual cohorts of patients in particular disease areas, to further accelerate patient recruitment and to enable the collection of observational data for longitudinal natural history studies and potentially to serve as patients in a real-world comparator arm[8]. This will improve the efficiency, scale, and equity of clinical research using routinely collected EHR information from hospitals that treat most patients. Randomized controlled trials (RCTs) with narrow inclusion/exclusion criteria, potential enrolment biases, and highly protocolized care are often not generalizable to large segments of the population across diverse geographies, sociodemographic categories, or subtypes of disease manifestation and progression. Today, there is a consensus from academia, industry, and regulators that RWD should also play a role in determining safety and efficacy for drug approvals, as evidenced for example by the recent investment by the European Medicines Agency in DARWIN EU[9].

For many of the potential uses of EHR data for clinical research listed above and elaborated in a forthcoming xShare working paper 5.2, there is also a growing consensus on the data elements that are of greatest value to be able to analyse. These data elements include, for example, the most frequently occurring data elements across therapeutic areas that are used as eligibility criteria[10]. When examining data sets required to support these different use cases, there is a substantial overlap between commonly occurring core information, such as demographics, conditions, labs, and procedures. Although each disease area adds several specialised data items, the evidence points to a common set of data elements, which map to a study population to significantly reduce the inefficiencies/delays in evaluating whether a patient would be eligible for that protocol, reducing to a

[3] De Moor G, Sundgren M, Kalra D et al. Using electronic health records for clinical research: The case of the EHR4CR project. J Biomed Inform 2015;53:162-173. https://doi.org/10.1016/j.jbi.2014.10.006

[4] Claerhout, B., Kalra, D., Mueller, C., Singh, G., Ammour, N., Meloni, L., Blomster, J., Hopley, M., Kafatos, G., Garvey, A., Kuhn, P., Lewi, M., Vannieuwenhuyse, B., Marchal, B., Patel, K., Schindler, C., & Sundgren, M. Federated electronic health records research technology to support clinical trial protocol optimization: Evidence from EHR4CR and the InSite platform. Journal of biomedical informatics, 2019, 90, 103090. https://doi.org/10.1016/j.jbi.2018.12.004

[5] Prokosch HU, Ganslandt T. Perspectives for medical informatics. Reusing the electronic medical record for clinical research. Methods Inf Med, 2009;48 (1), pp. 38-44

[6] Dugas M, Lange M, Muller-Tidow C, Kirchhof P, Prokosch HU. Routine data from hospital information systems can support patient recruitment for clinical studies. Clin Trials 2010;7(2):183–189. https://doi.org/10.1177/1740774510363013

[7] The EHR2EDC project. Available at https://www.i-hd.eu/rd-and-collaborative-projects/ehr2edc/

[8] Lombardo, G., Couvert, C., Kose, M., Begum, A., Spiertz, C., Worrell, C., Hasselbaink, D., Didden, E-M., Sforzini, L., Todorovic, M., Lewi, M., Brown, M., Vaterkowski, M., Gullet, N., Amasi- Hartoonian, N., Griffon, N., Pais, R., Rodriguez Navarro, S., Kremer, A., Maes, C., Hooi Tan, E., Moinat, M., Genescà Ferrer, J., Pariante, C.M., Kalra, D., Ammour, N., Kalko, S., Electronic health records (EHRs) in clinical research and platform trials: application of the innovative EHR-based methods developed by EU-PEARL, Journal of Biomedical Informatics, 2023; https://doi.org/10.1016/j.jbi.2023.104553

[9] European Medicines Agency. Data Analysis and Real World Interrogation Network (DARWIN EU®). Available from https://www.ema.europa.eu/en/about-us/how-we-work/big-data/data-analysis-and-real-world-interrogation-network-darwin-eu.

[10] Doods J, Botteri F, Dugas M, Fritz F. A European inventory of common electronic health record data elements for clinical trial feasibility. Trials 2014;15(1):18. https://doi.org/10.1186/1745-6215-15-18

very practical level the number of patients who require clinician final review to identify the right patients.

Unfortunately, because EHR data is generated from routine clinical practice, there are significant challenges in reusing these data for research. These challenges include missing or incomplete data, bias, variability, and semantic and syntactic heterogeneity, all of which can make analysis of EHR information difficult and resource consuming. Since the adoption of interoperability standards to communicate EHR data is also patchy within and between countries, the data that does exist is often still siloed within healthcare provider EHR systems. The common challenge facing many use cases for utilizing EHR data is therefore access to a complete high-quality core data set to scale up across healthcare sites and many patients.

1.3 The opportunity presented by the International Patient Summary, in Europe

Given the vast diversity of types of health data that are captured and used during healthcare delivery, and the limited uses made today within healthcare settings for the computability of this data, it is perhaps not surprising that formalising the consistency of clinical data elements and their containing data structures, prioritising the design of EHR systems and a culture of clinical documentation to maximise the quality of this data, and investing in its semantic interoperability, have all been limited to date. As a consequence, the data within EHR systems is of variable quality, a high proportion of the data still being captured in free text. Many countries have invested in initiatives to improve the interoperability, and sometimes the underlying quality, of key data categories such as prescriptions, laboratory test results and data items used for reimbursement and public health surveillance purposes. Amongst those key data categories has been some form of a patient summary, primarily the health conditions, allergies, medications and other noteworthy medical facts that may support safe continuity of care between healthcare providers. These investments have primarily been for within-country communication, and the summary specifications adopted have been similar but not identical across countries.

In Europe, over a decade ago the eHealth Network published a specification, as a guideline, for a European patient summary[11]. This consolidated the empirical learning from the European Commission funded European Patient Smart Open Services project (epSOS), which piloted the cross-border exchange of a patient summary, as well as prescription and dispensing information, between several Member States[12]. The European patient summary has informed but not dictated the national summaries adopted by European Member States, which have varied in the extent to which they have invested in adoption and achieved penetration across their healthcare provider systems. In 2019 the European Commission formalised the adoption of the European patient summary as well as hospital discharge reports, laboratory and radiology reports, prescriptions and dispensing as the European Electronic Health Record Exchange Format (EEHRxF)[13]. This has influenced but also not dictated Member State adoption practices.

[11] Guideline on the electronic exchange of health data under Cross-Border Directive 2011/24/EU: Patient Summary https://health.ec.europa.eu/document/download/e020f311-c35b-45ae-ba3d-03212b57fa65_en?filename=ehn_guidelines_patientsummary_en.pdf

[12] The European Commission. Cross-border health project epSOS: What has it achieved? 2014. Available from https://digital-strategy.ec.europa.eu/en/news/cross-border-health-project-epsos-what-has-it-achieved

[13] The European Commission. Recommendation on a European Electronic Health Record exchange format. 2019. Available from https://digital-strategy.ec.europa.eu/en/library/recommendation-european-electronic-health-record-exchange-format

The Regulation on a European Health Data Space, which has just recently (24 April 2024) been approved by the European Parliament and the Council[14], will change the status of the EEHRxF to mandatory for adoption by all Member States. A combination of European Commission and Member State budgets will therefore now be allocated to ensuring the interoperability of EEHRxF format data (including the patient summary) within and in between countries. Member States will need to establish ways of consolidating this data at a national level, per patient, to break down the data fragmentation that presently exists, and to enable rapid communication of this patient summary data between Member States in the case of a cross-border healthcare need. It is therefore also likely that they will establish national or regional repositories of patient summaries that are regularly updated. The EEHRxF will be the mandated standard, although the detailed specifications for the data representations and semantic interoperability layers will be specified in forthcoming secondary legislation (Implementing Acts) over the next couple of years. It is widely expected that European patient summary guidelines will be gradually aligned with the ISO International Patient Summary, with the added benefit that this aspect of the EEHRxF will also be international. Member States will almost universally now scale up investments in the capability of their deployed EHR systems to capture, process and communicate the IPS, and in the digital and data literacy of health and care professionals to ensure the quality of this data set.

This improved availability and quality will be an important opportunity for the clinical research community to analyse IPS data repositories for the clinical research use cases mentioned above and others as those detailed in working paper 5.2. It is therefore timely for clinical research to critically examine the IPS from the perspective of its RWD use cases, and to consider proposing a modest (i.e. likely to be acceptable and achievable) extension to it, to introduce the missing data elements that have maximum research value. This working paper proposes this core data element set.

1.4 The challenge of heterogenous research data models and standards

An important issue concerning research data standards and their underlying models is that there are numerous models in use by various research networks. These include, but are not limited to, Observational Medical Outcomes Partnership (OMOP) Common Data Model (CDM)[15], The National Patient-Centered Clinical Research Network (PCORNet) Common Data Model (CDM)[16], Sentinel Common Data Model (CDM)[17], and Informatics for Integrating Biology and the Bedside (i2b2)/Accrual to Clinical Trials (ACT) Network[18]. In a complementary project (Common Data Model Harmonization), these four data models have been mapped to the Biomedical Research Integrated Domain Group (BRIDG) Model, which is an ISO, HL7 and CDISC model that represents clinical research and its link to healthcare[19]. However, BRIDG is an information model without specifications for controlled terminology, and mapping across data models not only contributes to loss of meaning but is also limited in that each time a model is updated to a new version, the prior mappings are outdated. Hence, the CDMH Phase 3 led by NIH/NCATS, NIH/NCI, FDA and ONC is in the process of registering mappings by element and ensuring semantic equivalency to automate conversions of data between different models. This project includes OMOP and PCORNet linkages with the CDISC Study Data

[14] The European Commission. European Health Data Space. 2024. Available from https://health.ec.europa.eu/ehealth-digital-health-and-care/european-health-data-space_en
[15] Observational Health Data Sciences and Informatics. Standardized Data: The OMOP Common Data Model. 2024. Available from https://www.ohdsi.org/data-standardization
[16] PCORnet. Accessed 24 May 2024. Available from https://pcornet.org/data/
[17] FDA. Sentinel Common Data Model. Accessed 24 May 2024. Available from https://www.sentinelinitiative.org/methods-data-tools/sentinel-common-data-model
[18] Please see https://community.i2b2.org/wiki/display/ACT
[19] NIH. National Cancer Institute Biomedical Research Integrated Domain Group (BRIDG) Model. 23 April 2023. https://bridgmodel.nci.nih.gov/

Tabulation Model (a global standard for aggregated data across patients in clinical research). The core data element sets leverage CDASH (the CDISC by-patient data collection standard) for research and the USCDI developed through the HHS/ONC for healthcare (see Appendix II). The USCDI is in the process of adding elements to support research, public health and other use cases. Along with the IPS, this CDMH work will inform xShare core data elements. In addition, the IMI projects EHR4CR and EHR2EDC have identified commonly used elements as have groups in the population/public health space.

For xShare, specifically, bridging to health care requires harmonizing currently adopted global clinical research standards (CDISC) with the European Patient Summary (HL7 CDA IPS) which is a core service of MyHealth@EU, one of the HIDs in the EHDS draft regulation, and the topic of a G7 report[20], as well as and the more recent HL7 FHIR IPS, and IHE IPS. The GDHP, a consortium of 33 countries, is committed to adoption of the IPS bringing the topic to the G20 working groups in Brazil this year. MoH-NL leading the Regulators forum in the European EHRxF Standards and Policy Hub is the current chair of GDHP interoperability stream. With this high profile, harmonization of IPS formats is a very important issue, particularly as it relates to clinical research and returning data to patients in a format they can understand, and share should they choose to do so.

This harmonisation is partly reflected in the methodology reported in the next chapter of this working paper, to consolidate core clinical research data sets and data specifications. Harmonisation will be further pursued in subsequent work and a later working paper from xShare WP5 that will present a formal specification for the IPS+R and will be taken up and hopefully sustained by the EHRxF Standards and Policy Hub.

[20] https://assets.publishing.service.gov.uk/media/61d82fbd8fa8f505893f1c93/G7-international-patient-summary-roadmap.pdf

2 Methodology for Identifying Harmonised Core Data Elements

The focus of WP5 in xShare is and was the application of the European EHRxF to clinical research. This working paper aimed to establish a harmonized core set of data elements across clinical research and continuity of care, aligning parts (domains) of the CDISC Clinical Data Acquisition Standards Harmonization (CDASH) standard with parts (categories) of the IPS, and engaging with the public health workstream via WP4, since Population Health Information Research Infrastructure (PHIRI) has already established a set of common data models for public health. Results from the Electronic Health Records for Clinical Research (EHR4CR) and Electronic Health Records to Electronic Data Capture Systems (EHR2EDC) projects which were part of the Innovative Medicines Initiative (IMI)/ Innovative Health Initiative (IHI) were also used to establish an extended version of IPS for research called IPS+R. Review of other relevant initiatives such as the USCDI v4.0, further validated the core set of data elements. The relevant comparisons are available as supplementary materials.

2.1 Harmonised Core Data Set Across Health Care, Population Health and Clinical Research

Consider data elements as a microcosm of the greater health care, population health and clinical research ecosystem. The methodology to determine the core data element set consisted of a stepwise approach. The first step started with the Health Level Seven International Patient Summary Implementation Guide v1.1.0 STU 1 (HL7 IPS IG); the second step leveraged the Clinical Data Interchange Standards Consortium (CDISC) Clinical Data Acquisition Standards Implementation Guide (CDASH IG); the third step consisted of the IMI/IHI EHR4CR, and EHR2EDC clinical research data elements and the public health, PHIRI, data elements; the fourth step incorporated a review of the data elements described in the IHE IPS document as well as the ISO IPS documents. The data elements were informed through mapping from CDISC CDASH to the HL7 IPS IG, which included an initial review of the value lists (otherwise known as codelists or valuesets) in the various terminologies. Figure 1 depicts a 30k view of the harmonisation process.

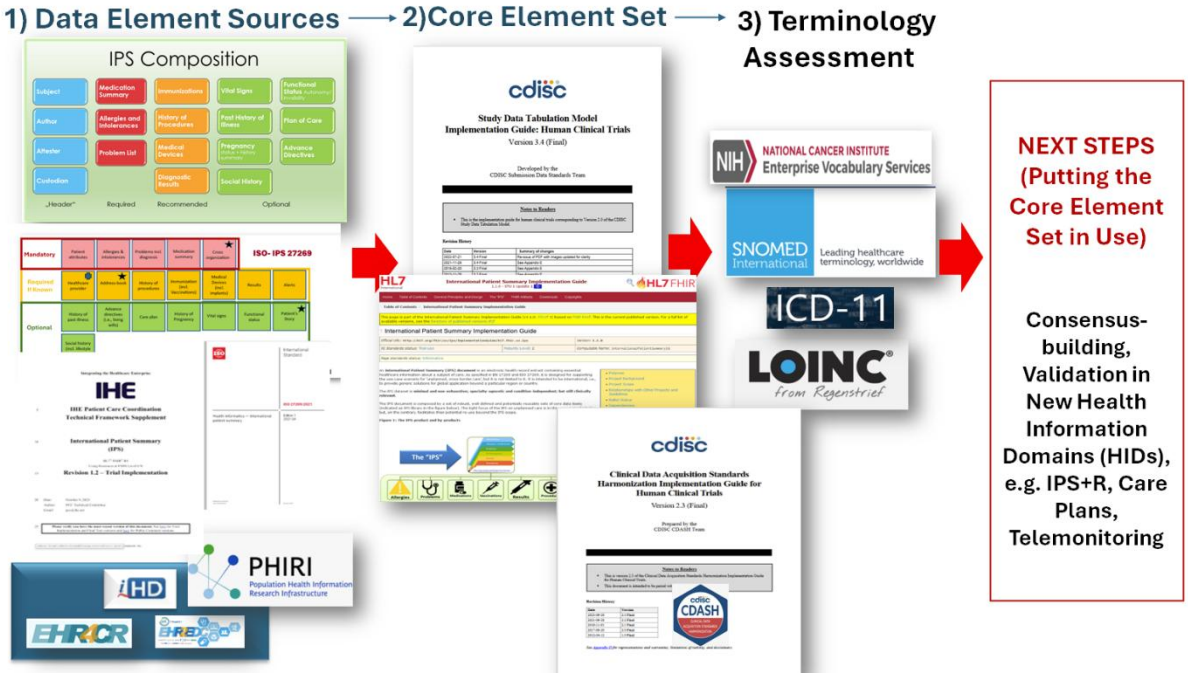

Figure 1: Core Data Element Set Development and Evaluation Process

2.2 Analysis of Data Elements, Alignment and Terminology Assessment

The analysis in pursuit of the common core set of data elements included reviewing outputs from the IMI EHR4CR project, the IMI EHR2EDC project, PHIRI and IPS documents (ISO, HL7, IHE and eHN). These sources of data elements and categories were then compared with HL7 FHIR and CDISC SDTM and CDASH standards with respect to how to 'model' the data at different levels. The elements identified for the core set were those that were deemed to be important for healthcare, research and public health. There was an initial assessment of a potential method to incorporate relevant terminology. The resulting core set of data elements will be reviewed by more stakeholders to build consensus around the elements as well as their priority (e.g. mandatory, recommended, optional). The elements will also be validated through use and testing in Health Information Domains and the xShare business use cases. The following section provides an overview of each standard or data element source, with applicable considerations and what was gleaned from each. Links to the standards referred to in this section are provided in chapter 7.

2.2.1 HL7 IPS

The HL7 FHIR IPS Implementation Guide (IG) was reviewed for specific data elements that had codes assigned. Discrete data elements could easily be leveraged for research use cases such as populating IPS repositories; directly populating research forms, such as case report forms; or for evaluation as eligibility criteria or feasibility of protocol. Where there were no discrete data elements or thousands of terms and codes in the valueset, a couple of examples were mapped to see how this would work with the focus of looking for what can easily come over and whether there were gaps that would need to be harmonised.

The initial focus of the work was on the Required IPS Categories (see Figure 2). However, since there were discrete data elements in many recommended or optional categories, all content was reviewed with a view of "what can be gleaned for easy transformation to support research from HL7 FHIR to an electronic data capture system (and the associated case report forms)". In the kick off meeting, the xShare teams determined that Pregnancy status should be a mandatory field for research. Additionally, the pregnancy related, and women's health section of the IG has several discretely defined terms to inform public health and accentuate clinical research for women. The Vital Signs category was also determined to be very important for research (not "Optional" as currently specified in the HL7 IPS IG). See 2.2.7 and 2.2.8 for analyses of other IPS documents. The analysis of the HL7 IPS IG resulted in 21 data elements defined at various levels of granularity.

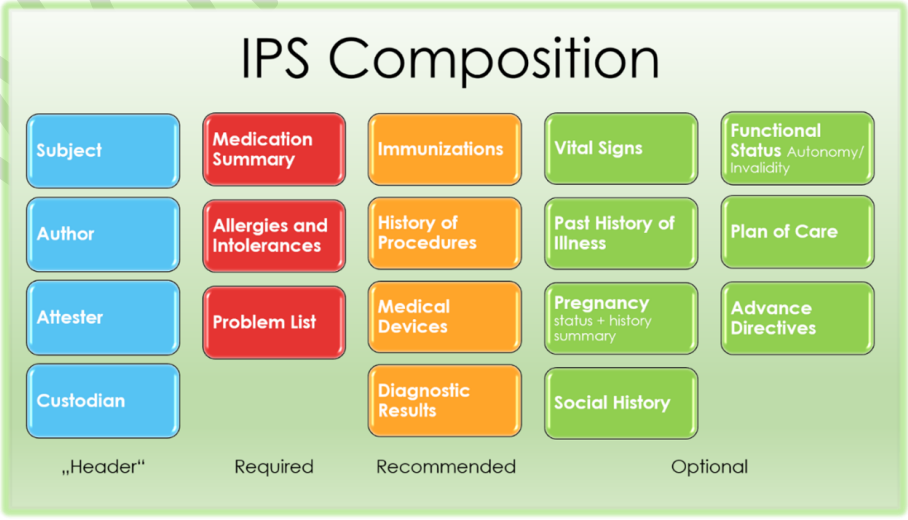

Figure 2: The HL7 IPS Composition from the HL7 IPS FHIR IG

2.2.2 HL7 FHIR IPA and US Core

The International Patient Access Implementation Guide provides additional HL7 FHIR profiles with data elements and valuesets. The US Core Implementation Guide provides a set of HL7 FHIR profiles in support of the US Core Data for Interoperability (USCDI). These IGs were reviewed for discrete data elements to further inform the work on the set of core data elements. Since HL7 FHIR has defined a discrete set of data elements for the Vital Signs category, these are the same for all three HL7 FHIR IGs (IPS, IPA, and US Core). The three IGs were reviewed for similarities and differences to further inform this work. The profiles consistently use the resources with some variations in the leveraging of the four Medication Resources and vary through the use of extensions, constraints, and valuesets.

2.2.3 CDISC, CDASH and SDTM

The CDISC standards are used globally for many kinds of medical, academic and clinical research. These standards were developed over the past 25 plus years by subject matter experts consisting of data managers, statisticians, clinical researchers, protocol authors, subject matter experts and others to inform and streamline clinical research from protocol through data collection, tabulation, analysis and reporting. The CDISC standards for tabulation (SDTM) and analysis (ADaM) are required by certain regulators, including regulators, USA FDA and Japanese PMDA. The CDISC foundational standards identify those elements that support and are common across all research studies (i.e. core data elements). There are also CDISC standards, which augment the foundational standards, that support over fifty different therapeutic areas and paediatrics. Controlled terminology underpins the CDISC standards and is maintained and supported by the NIH NCI EVS.

CDISC has supported global clinical research for decades with end-to-end harmonized standards, gradually working with partners to add specific therapeutic area standards. CDISC collaborations for standards development include (but are not limited to) the Coalition For Accelerating Standards and Therapies (CFAST), FDA, PMDA, WHO, C-Path, TransCelerate; the National Institute of Allergy and Infectious Disease (NIAID) for HIV studies, pharmacovigilance, and meta-analysis; the National Institute of Diabetes and Digestive and Kidney Diseases (NIDDK) as a member of C-Path Polycystic Kidney Disease TA consortium; the National Institute of Neurological Disorders and Stroke (NINDS) for CDEs used in Parkinson's and Traumatic Brain Injury research; National Institute of Mental Health (NIMH) contributions for Schizophrenia and in future CDEs alignment for Post Traumatic Stress; Oxford and IDDO for vaccine standards; and the National Institute of Child Health and Human Development (NICHD) for paediatric terminologies, along with the conect4children (c4c) IMI development of a Paediatric User Guide. CDISC also participated in IMI/IHI EHR4CR and EHR2EDC.

CDISC has also been interested in linking its research standards with healthcare for quite some time. CDISC started the Clinical Research Special Interest Group within HL7. In collaboration with FDA and others, this became RCRIM and is now the BR&R workgroup. CDISC participated in the HL7 Board when HL7 FHIR was adopted and transitioned from working with HL7 V3 and HL7 CDA to begin using HL7 FHIR. CDISC has published a mapping IG between HL7 FHIR and CDISC SDTM; this will be updated and augmented as part of the CDMH P3 project.

In addition, through the Healthcare Link Initiative and CDISC Europe Foundation, CDISC has developed IHE Interoperability Specifications to support population of research case report forms direction with EHR data and also to support activities identified within structured protocols (which has now transitioned to the ICH M11 project, TransCelerate USDM and Vulcan UDP initiatives.

For the xShare WP5 task 5.2 for this working paper (5.1), the initial focus from a research perspective was on CDASH, which was developed to facilitate the flow of data from individual patient (case report form data collection) into the multi-patient Study Data Tabulation Model (SDTM), which is the aggregated tabular view with which regulators and statisticians wish to work. The initial publication of CDASH was indeed a core set of foundational data elements that are common across all clinical research studies. The initial version of CDASH, published in 2006, had 16 domains and just over 100 elements. These are represented as questions on case report forms. (Standard eCRFs are available for users to initiate clinical research studies, which can reduce study start-up time by ~ 70%).

Using this foundational CDASH standard, the CDASH Domains were compared to IPS Categories (Figure 3). The next step was to discover how the questions on the case report forms surrounding the data items (discrete data elements) would truly inform this xShare working paper. There is still an outstanding step in terms of how to count data elements, specifically whether to count the metadata (which provide the context) as discrete elements or only the 'parent' data element. For example, medication is an element; however, dose and dose form and route of administration provide context yet may not be considered elements. (See Additional Considerations.)

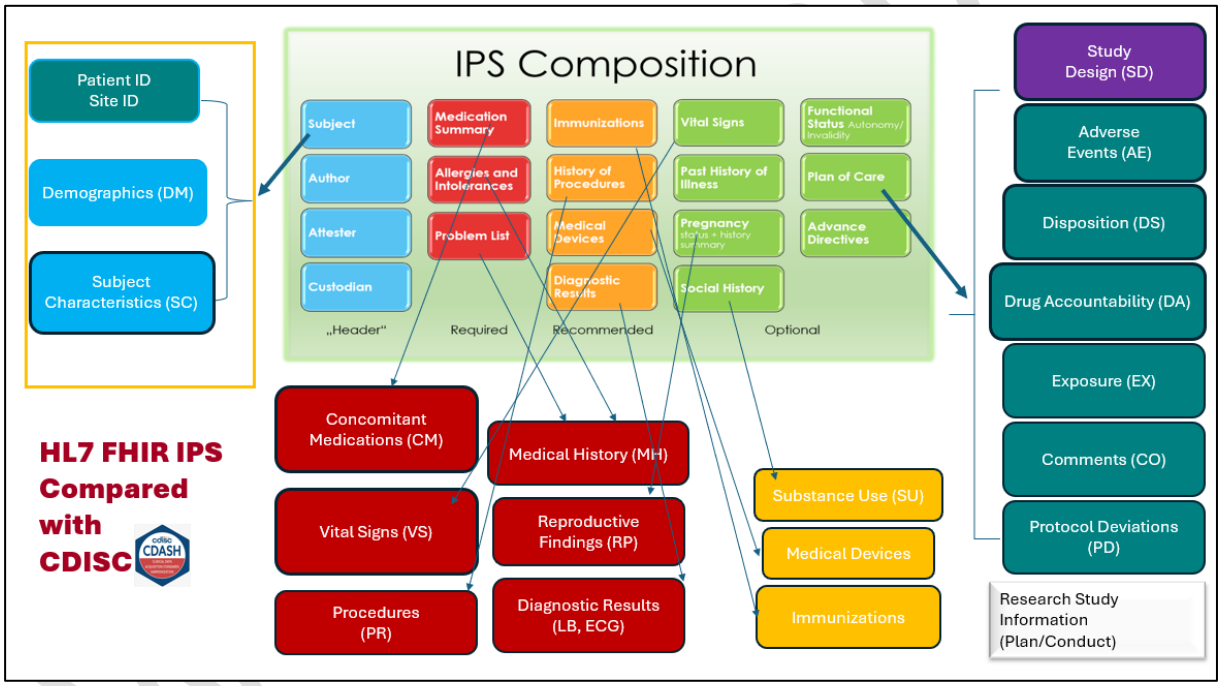

Figure 3: Diagram of Core CDASH Domains Aligned with HL7 IPS Categories

After the analysis of IPS, the CDISC case report forms were reviewed, and the domains identified to further inform the work. Any data elements identified here were added to the mapping with the idea of connecting the standards to leverage both for research (clinical and public health) through the semantics. There is more information about connecting the various standards in the Additional Considerations section.

2.2.4 IMI/IHI EHR4CR and EHR2EDC Data Elements

Contributions of lists of data elements were received from prior IMI/IHI projects namely the EHR4CR and EHR2EDC. These inventories of data elements were compiled by those projects by examining multiple cross-sponsor clinical trial protocols, with a particular focus on eligibility criteria, seeking to

elicit the commonest occurring elements that were therapeutic area agnostic. The 81 data elements contributed from the EHR4CR project included: 5 demographic/subject category; 13 problem list/diagnosis/medical history category; 7 medication summary category, 3 procedure category, 7 vital sign category, 41 laboratory findings of the diagnostic results category, and finally, 5 scores/classification clinical tools (not a category in the IPS). The 55 data elements contributed from the EHR2EDC project included: 5 demographic/subject category; 11 vital sign category, 6 medication summary category, and 37 laboratory findings of the diagnostic results category. These were added to the data element list to inform the core set of data elements. Reviewing these data elements pointed back to the questions from the CDISC CDASH case report forms. The findings appear in the appendix, and as supplement to be shared in the xShare website.

2.2.5 PHIRI Public Health Data Elements

The Public Health/Population Health and cross-border threats work package (WP4) for xShare aims to identify public health and population use cases. WP4 contributed 200 data elements from the Population Health Information Infrastructure Project (PHIRI)[21]. A number of these overlap with data elements already identified (largely laboratory and microbiology elements), and a number were related to healthcare operations and were thus determined not applicable for this working paper for xShare. After review, 35 data elements were applicable to clinical research and public health (subject category, diagnostic results category for general laboratory tests and microbiology tests) and 45 specific tests (data items) were for microbiology pathogens. These were added to the data element list to inform the core set of data elements for xShare that should support healthcare, research and public health.

2.2.6 ISO IPS (27269:2021)

The ISO/TC 215 Health Informatics - International Patient Summary (IPS) document purpose is to provide the patient summary to facilitate and support clinical decision-making at the point of care. It provides a general structure and guideline for the IPS pointing to CEN for the technical specification and SNOMED CT GPS for the terminology, defining the core data set for “cross organization care”. ISO IPS was analysed for similarities and differences between the HL7 IPS. They were assessed with respect to whether they were indeed the same, and whether they stayed true to the vision of ensuring a standardized International Patient Summary. The essential structure of each contains very similar concepts. Figure 4 shows the data blocks visualized in the IPS ISO. The figures for HL7 and ISO IPS are very similar even down to the colour coding for the required, recommended and optional fields. Data elements related to healthcare encounters were gleaned and deemed to add value to the core set. The ISO IPS (Figure 2) has 19 boxes with Cross organization, Healthcare Provider, Address book, and Patient’s Story while HL7 IPS has 18 boxes that are semantically similar, for the most part, with the exception of Author, Attester, and Custodian. These terms are related to the documentation of the International Patient Summary and fall under the same box as the Healthcare provider.

[21] https://www.phiri.eu/

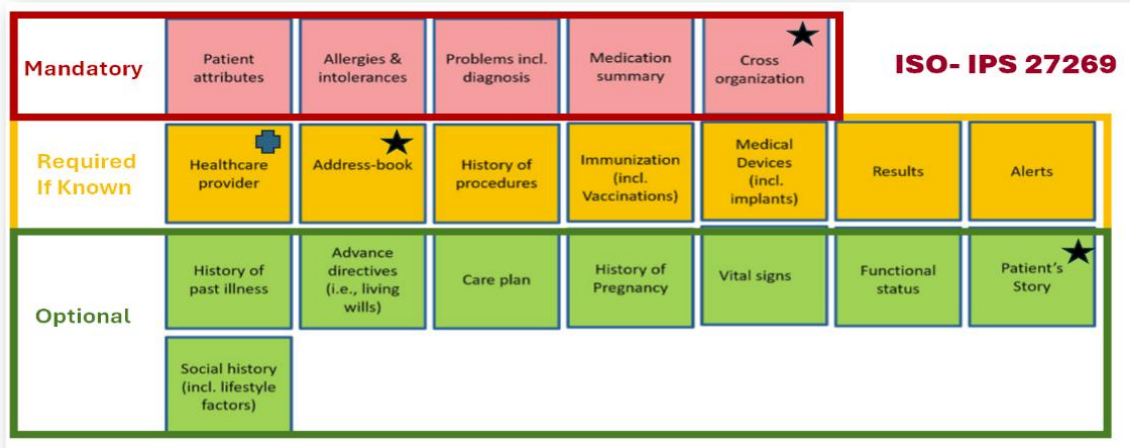

Figure 4: ISO-IPS Datablocks/Categories (indicates the ones that differ from HL7 IPS.)*

The ISO IPS provides content on healthcare encounters not found in the HL7 IPS. The three boxes with stars indicate the categories that are in the ISO IPS and not in the HL7 IPS. The box with the cross indicates where there is also additional information not in both IPS documents.

2.2.7 IHE IPS

Integrating Healthcare Enterprise (IHE) is an initiative by healthcare professionals and industry to improve the way computer systems in healthcare share information (IHE https://wiki.ihe.net/index.php/Main_Page). IHE IPS profile uses the HL7 IPS IG providing additional descriptive information to support multiple international use cases supporting planned and unplanned care across borders. It contains a full mapping to ISO/EN 17269 data elements to HL7 FHIR and CDA documents. Comparing and contrasting the IHE IPS to the ISO and HL7 IPS identified that it aligns with the standard as well. The IHE IPS includes data elements for healthcare encounters, including admission and discharge information.

2.2.8 IPS Document Analysis

HL7 IPS, ISO IPS and IHE IPS provide comprehensive, predefined set of data elements in a structured document format while the HL7 FHIR IPS represents the data using resources with defined attributes providing a mechanism to exchange. The data elements represented across ISO IPS, IHE IPS, and HL7 FHIR IPS are similar and include Patient Demographics, Patient Identifiers, Allergies and Intolerances, Medications, Observations, Procedures, Immunizations, and History of Past Problems. Notably the healthcare encounter is missing from the HL7 FHIR IPS. While the three standards share common data elements, the naming conventions of the data elements vary.

2.3 Related Projects Informing the Identification of the Core Set

2.3.1 eHealth Network Guideline on the electronic exchange of health data under Cross-Border Directive 2011/24/EU (European IPS)

The eHealth Network connects national authorities responsible for eHealth in designated member states working towards facilitating greater interoperability for sharing health data and empowering citizens to access and share their health data. The eHealth Network Guideline[22] was reviewed to further glean or support the core data element set. Data element concepts for laboratory test,

[22] https://health.ec.europa.eu/publications/ehn-guideline-patient-summary_en, https://art-decor.org/ad/#/ehng-/project/overview

results (including units), specimen and health encounter further support the core data elements gleaned from the IMI and PH contributions. There is a notable recommendation for controlled lists with valueset catalogues with agreed valuesets made easily available to implementers (p.16). Note there is an entire section on standards for units referencing Nomenclature for Properties and Units (NPU) and Unified Code for Units of Measure (UCUM) (p.20).

2.3.2 UNICOM Project and ISO/IDMP

The EU project Up-scaling the global univocal identification of medicines in context of Digital Single Market strategy (UNICOM) provides a view of big data pharmaco-epidemiology networks with the purpose of improving patient safety focusing on the Implementation of the ISO suite of Identification of Medicinal Products (IDMP) standards. This work informs the Medication Summary data elements in the core data element set and has the potential to connect clinical research to Identification of Medicinal Products (IDMP) product information. Aligning the core medicinal data elements such as substance/product, dose form, ingredient, and strength will support data exchange. Their work highlights differences in the different terminologies related to dose forms, therefore, aligning on a data element concept with an aligned codelist or valueset would enable precision in the descriptions, definitions, and characteristics to enable robust data aggregation. UNICOM has promoted conformance to the ISO IDMP standard and Gravitate-Health[23] in collaboration with the Vulcan accelerator developed an HL7 FHIR Implementation Guide for the representation of medical product information (ePI), including but going beyond what is required in an EHR or IPS; this work also advocated for the use of EDQM for dose forms. It has piloted and advocates for the generation of PhPIDs by the WHO as a global identifier that could be included in the patient summary to enable global recognition of a patient's medication list.

2.4 Technical and Semantic Considerations in the Core Data Element Set

Taking into account the many inputs to propose a core data element set for IPS across health and care, public/population health and clinical research following a path to inform a possible research extension to the IPS (the IPS+R), these design considerations have sought to balance on the one side being comprehensive enough that the data elements can be understood in their clinical context and present sufficiently meaningful information to inform future clinical decision makers caring for the patient, but on the other side to avoid making the resulting IPS+R over complicated or burdensome to populate, or introducing expectations on data availability that is unlikely to have values recorded in the majority of electronic health record systems in use today.

The National Information Institute of Standards and Technology (NIST) define a data element as "a basic unit that has a unique meaning and subcategories (data items) of distinct value[24]. Examples include gender, race, and geographic location. The data elements may be grouped into related categories (or domains) and assigned detailed *terminologies*/codelists/valuesets. Additional consideration appears in the literature[25].

The data elements are within a hierarchy that starts with groups of elements and continues to more detail through *terminologies*/codelists/valuesets. Each level clusters the data elements that would normally be captured from a patient at a single moment, for example during a clinical encounter, and authored in an electronic health record by one person during that encounter. An example of this would be a prescribed drug, for which the name of the medicinal product, the dose, the route and intended duration of the treatment would normally be specified at the same time by the prescriber.

[23]

[24] NIST. Computer Securiy Resource Center. Accessed 24 May 2024. Available from https://csrc.nist.gov/glossary/term/data_element

[25] https://ebooks.iospress.nl/publication/50513

Similarly, a systolic and diastolic blood pressure, and optionally the assessment method, would normally be captured as a single entry in an EHR. In contrast, height and weight might not be captured at the same time: they might be entered by different authors on different occasions when the patient is examined and are therefore not clustered under a containing data item.

The definitions used within this xShare working paper are described in this section.

2.4.1 Hierarchy of a Core Data Element Set

It is important to structure a core data element set such that it is readily comprehensible and that the layers of detail are clear. The proposed set is structured into categories or domains of elements that are then further described with metadata and defined as data element concepts with each data element associated with relevant variables and terminologies/vocabularies, the latter being essential to semantic interoperability. Figure 5 shows the relationship between these items.

Explicitly, Figure 5 shows examples of how the category of Medication has a data element of Medication Name with the various metadata pieces and the associated controlled terminology for Medication Route of Administration. DEC refers to a Data Element Concept per ISO 11179 (3.1.3, 3.1.4 and Figure 6).

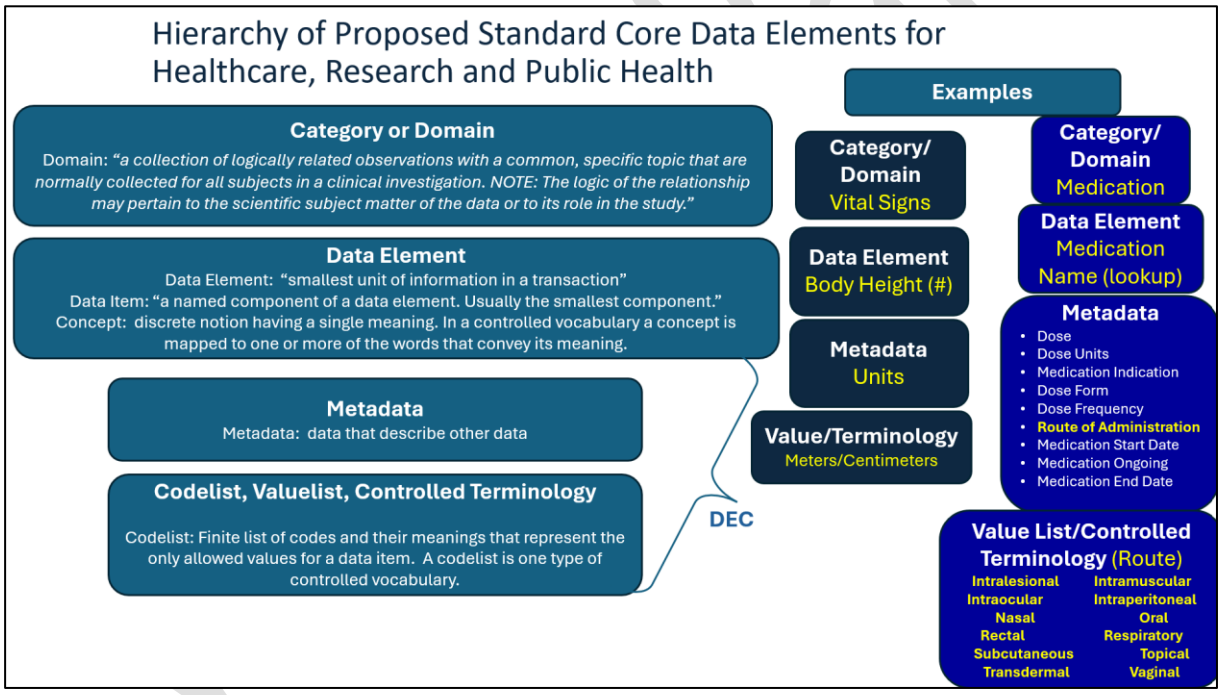

Figure 5: Hierarchy of Proposed Standard Core Data Elements for Healthcare, Research and Public Health

2.4.2 Categories or Domain

Per the CDISC Glossary, a domain is "a collection of logically related observations with a common, specific topic that are normally collected for all subjects in a clinical investigation. NOTE: The logic of the relationship may pertain to the scientific subject matter of the data or to its role in the study."

For the purpose of this core data element set, a Domain is considered to be essentially equivalent to a Category in the IPS and is not specific to a 'clinical investigation'. Thus, a Category or Domain is 'a collection of logically related observations with a common, specific topic'.

2.4.3 Data Elements and Variables or Metadata

Data elements are defined as the “smallest unit of information in a transaction” while a data item (which is the term used in the CDISC transport XML standard is “a named component of a data element. Usually the smallest component.” A concept is a discrete notion, having a single meaning. In a controlled vocabulary a concept is mapped to one or more of the words that convey its meaning. In the xShare hierarchy, a data element may have metadata or variables describing that element, such as dose for medication. There is still discussion as to whether each important and relevant piece of metadata should be counted as an element. That said, it is important to consider that each data element may have a valid valueset/codelist or set of controlled terminology that enable the meaning of the data entered in the data element field can be transported to another location/user and retain its meaning, i.e. semantic interoperability.

2.4.4 ISO/IEC 11179 MDR Structure of a Data Element

ISO (the International Organization for Standardization) and IEC (the International Electrotechnical Commission) consist of national bodies that are members of ISO or IEC to develop a worldwide standardized system. The ISO/IEC 11179 standard promotes a standard description of data, common understanding of data across organizational elements and between organizations, re-use, harmonization, management of the components and re-use. The ISO 11179 standard states that “descriptive data are known as metadata.” This provides a framework for any kind of data and any purpose and is essential to understanding the meaning of an object through the object descriptions (in this case health data elements and surrounding metadata) independent of an application.

2.4.5 Data elements and data element concepts

ISO 11179 differentiates data elements, data element concepts, definition and metadata. As the identification of the core set of data elements ensued, we realized and noted the variation of the levels of granularity of health data within the IPS categories further defined in the IPS FHIR IG. It is important to understand what a data element is to grasp the direction of this work. ISO 11179 has four corners noted for the structure of a data element: data element concept, conceptual domain, data element, and value domain (see Figure 6). First, compare and contrast the data element and data element concept: ISO 11179 defines a data element as “unit of data for which the definition, identification, representation and permissible values are represented by means of set of attributes”; and a Data Element Concept as “concept that can be represented in the form of a data element, described independently of any particular representation”. Further descriptive definitions aid in comprehending this: definition is “representation of a concept by a descriptive statement which serves to differentiate it from related concepts”; metadata is “data that defines and describes other data”; and Value domain is a “set of permissible values”. A metadata item is “instance of a metadata object”, while the metadata object is “object type defined by the metamodel” and a name is “designation of an object by linguistic expression”.

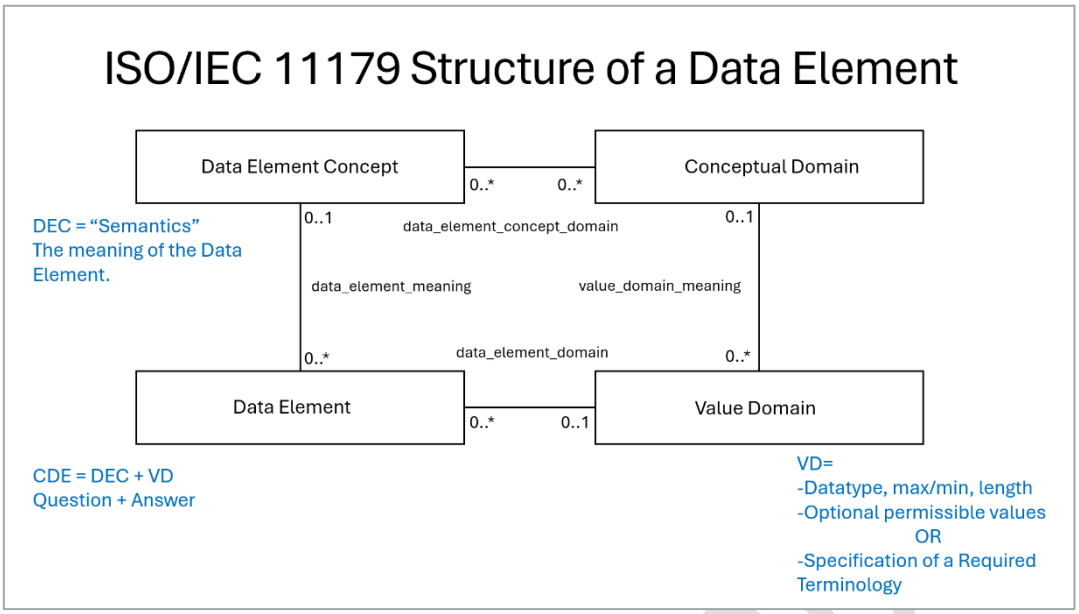

Figure 6: The ISO-IEC Structure of a Data Element (ISO 11179, MDR)

The application of the ISO 11179 is shown in Figure 7. Conceptual layers are data elements, while the implementation layer varies according to the specific standard. The figure shows how this would be represented in a CDISC SDTM table; a format easy for humans to review and digest versus machine readable format such as JSON or xml.

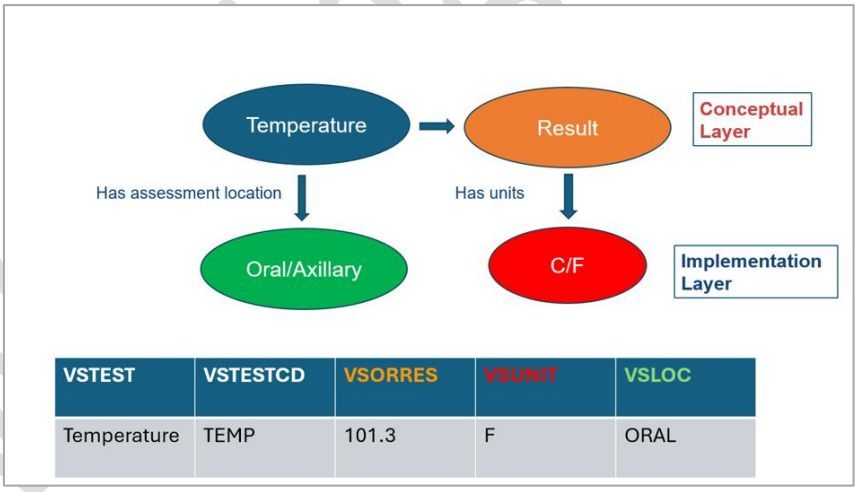

Figure 7: The ISO-IEC Structure of a Data Element Applied to Standards

The data element concept in Figure 7 further demonstrates the ISO 11179 structure of a data element. The Conceptual Layer has a Conceptual Domain of Vital Signs, the data element concept is Temperature thus the Question and Answer is: "What is the Body Temperature?". The Implementation Layer for Temperature "Has assessment location" of "Oral/Axillary" and the Result "Has units" of "C/F"; the Value Domain is the permissible values for the results or in other words, the answer (seen in the table VSORRES, the Vital Signs in Original Result Units). The location and unit are metadata for the test or data element concept of temperature further describing the data element concept of Temperature. The topic variable is the vital sign test of temperature.

To bring this full circle Figure 8 takes this one step further back to the topic of medication further driving home the point that each piece of the data element concept metadata provides questions with answers that have insights on the topic variable which in this case is medication.

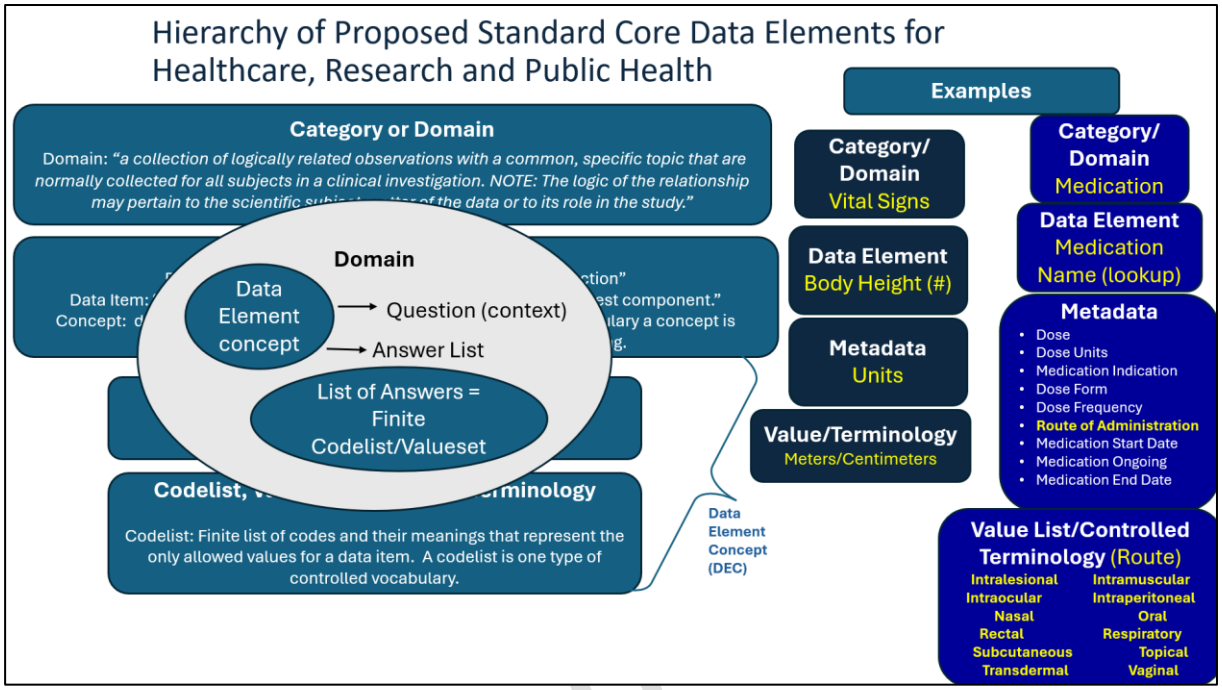

Figure 8: Core Data Element Structure with Question and Answer (Data Element Concept)

2.4.6 Applying the ISO 11179 Principles to the ISO Core Set

During the data element gathering process, the contributions were viewed in terms of the metadata items provided. For example, there are lists of microbiology tests used in public health and laboratory tests used across health care and clinical research. "If fasting" is a concept that was provided within the general list. This is metadata about the data item and will impact laboratory tests such as glucose or cholesterol measurements. The Public Health data element contributions provided a list of descriptive data elements that related to health care such as details about the specific test that was done by whom and where (operational health care data fields) and a number of specimen and test related elements (Specimen ID, Specimen Material, Microorganism, Test Method, Test Date Time, Reference ranges-upper and lower limits, etc.). Following the guidelines for common data element structure and definition will ideally support semantic interoperability across models. An xShare core data element set for health care, clinical research, and public health globally will help sharing of patient summary across borders, baseline clinical research screening, and provide signals of epidemics informing treatment and safety information.

2.5 Data Concept Identification Process

The data element concept identification is critical to the establishment of a common core of harmonized data elements for HL7 FHIR IGs that will be part of X-Bundles for xShare. Furthermore, understanding what data elements will inform the EHRxF IPS+R provides the underpinning for exchange of research data. It will enable the development of the research extensions to IPS+R and provide the basis for testing, demonstrating and validating data availability, quality, and utility of

IPS+R in the real world with real world data sets. This will lay the foundation for carrying out the business use case leveraging real world EHR.

First, gathering of data elements commenced by reviewing the IPS FHIR IG (this means there was a coding system and code assigned to it that can easily be leveraged for research), then the second set was identified through review of the contributed data elements from IMI and Public Health along with the CDASH IG case report forms providing additional supporting information. The core set was consolidated and confirmed by reviewing the other initiative documents.

Once the gathering of available discrete data elements occurred, an analysis of these data items (discrete values) and the data variables (data elements that cover the concept that essentially asks a question and has a codelist or answer list) was performed. These were reviewed based on source of the data element in terms of what group recommended/requested them, did all the groups recommend them and is there enough information surrounding the data element concept (variable) that the observation collected (electronically) has meaning. For instance, if a medication dose is known (a qualifier variable) and the medication (topic) it is related to is not, the information will be unusable. In research there are certain qualifier variables that are relevant to the objective of the study. An example is the location for collecting the Vital Sign of Temperature. If the location is Rectal the value is very different than a location of Axillary. If the research study Inclusion criteria for Sepsis had a criterion of “Temperature greater than 40 degrees Celsius”, the temperature measurement location would make a difference.

In the review of data element concepts (variables not data items), we noted that there are several data elements (variables) in the CDISC SDTM that are used to standardize the original data either to the standardized terminology or the standardized units thus changing the results. A decision was made not to include these as this would be part of the data processing or data management once the electronic health record original result/condition/procedure/treatment data is obtained from the sites (e.g., PRDECOD/Procedure Dictionary-Derived Term).

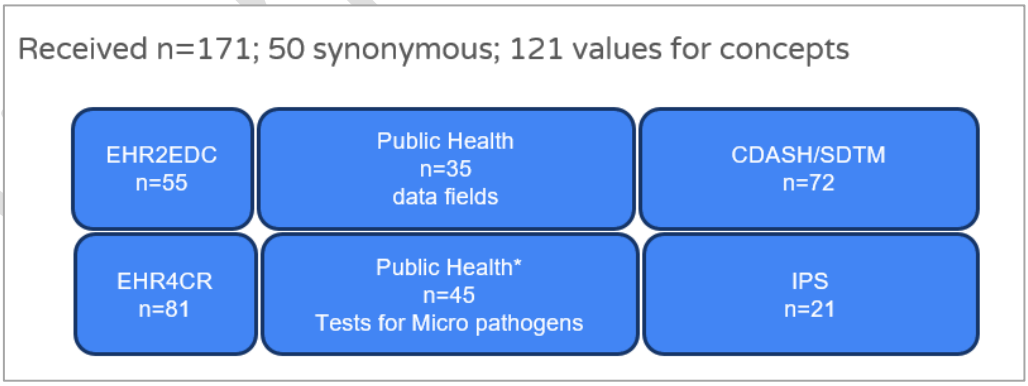

Figure 9: Data Element Collection Breakdown

Ultimately, an analysis of the data elements gleaned from the various sources included a compilation of the three IPS documents (ISO, HL7 FHIR, IHE), EHR4CR, EHR2EDC, PHIRI, and the CDASH IG. A total of 171 data elements were received (health operations elements were excluded), with 50 synonyms resulting in 121 data elements that varied in granularity. These were analysed further to determine

what IPS categories these supported and what CDISC domains were covered. A breakdown is shown in Figure 9 and Figure 10.

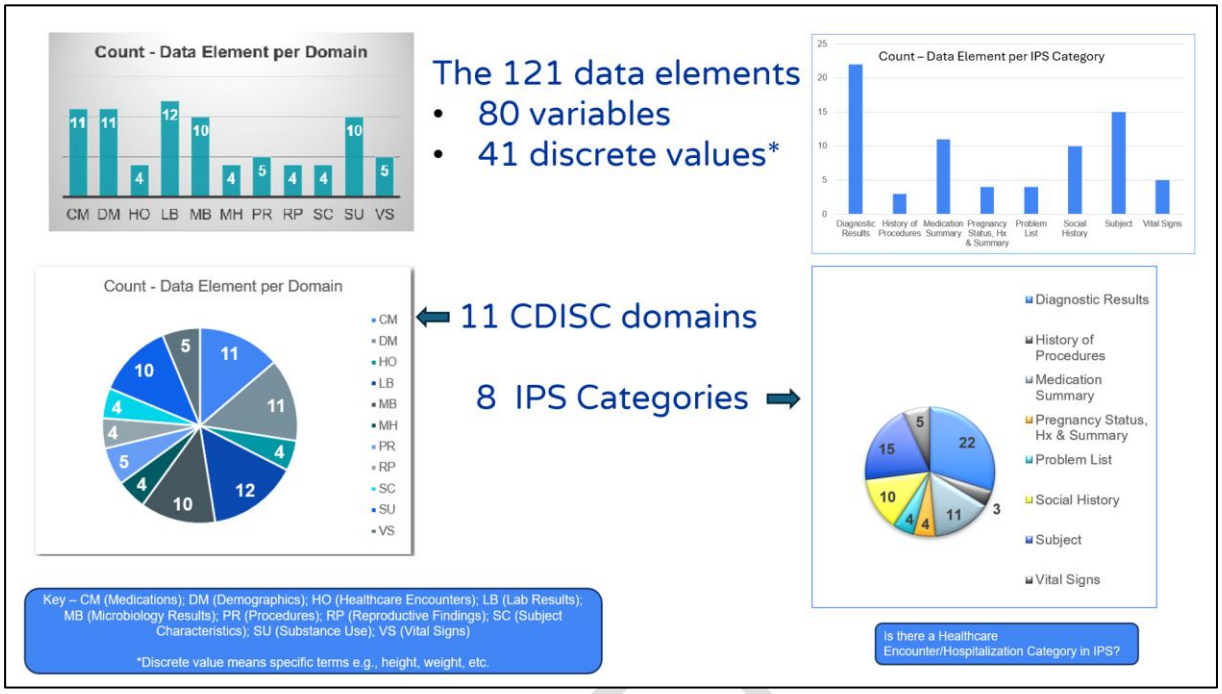

Figure 10: Data Element Set by IPS Category and CDISC Domain

2.6 Codelists and Valuesets

In the HL7 IPS IG terminology artifacts are defined and were reviewed. A number of data elements have specific coding systems and codes assigned for the data item which makes interoperability achievable. These can be mapped from one terminology and model to another terminology and model. Other data elements have extensive valuesets associated with the terminology bindings.

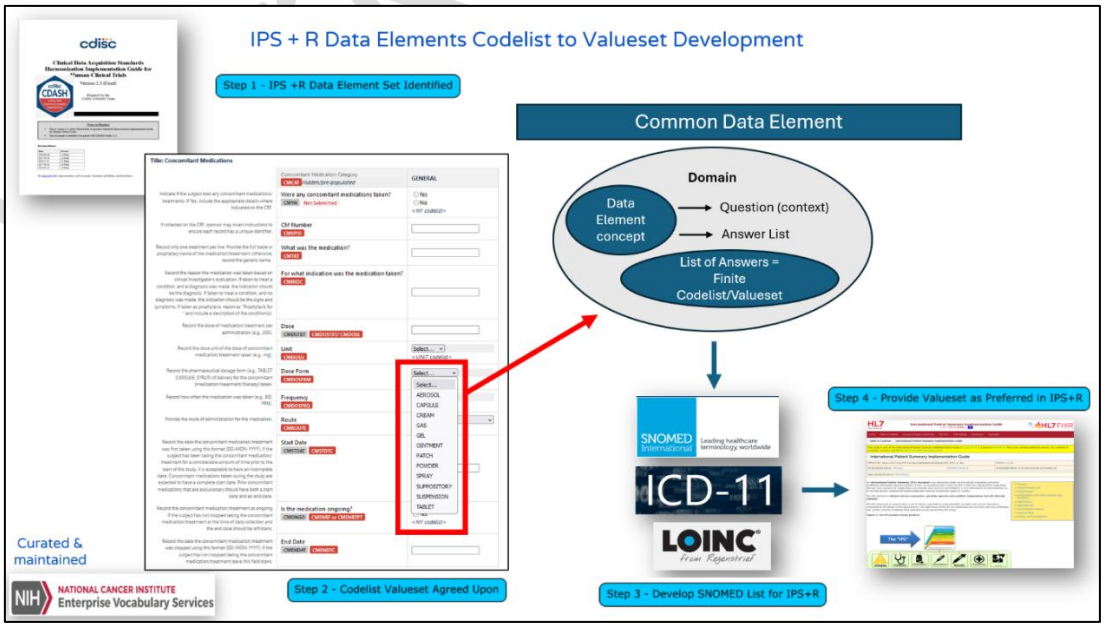

Figure 11: Proposed IPS+R Data Element Valueset Development from Codelists

2.6.1 Additional Considerations from Initial Review

After receiving the IHI EHR4CR, the EHR2EDC, and the Public Health data elements and analysing across these contributions of data elements – those gleaned from the IPS, any gleaned from CDASH, then IHI and Public Health data elements – it was realized/recognized that we are looking for are the questions surrounding the topics in these domains: be they a problem which is a condition, and which is a medical history item in the CDISC realm. It is interesting to consider the term "Problem". In the midpoint discussion, time was spent analysing and dissecting this. Comments were made that this may be more than a "Condition" and can include things like "Homelessness" or "Problems that are of importance of a patient's care". When analysing the HL7 FHIR IPS IG further after this conversation, it is represented using the Condition Resource and a SNOMED CT code. Is it a condition or a problem? In electronic healthcare records, the Problem List is the list of Conditions which are generally medical conditions.

This could impact the mapping of the Problem List down the road because "Homelessness" would be a "Subject Characteristic" which is something about a person that is not collected in other domains (nice of CDISC to have a literal junk drawer!). If one dives in deeper, it is a place for information about a person that may change over time such as Marital Status, Homelessness, National Origin (rarely changes but it can- A person may have been born in Iran, but moved to France and became a French citizen 30 years ago and now consider themselves French. Perhaps, they live in Italy half the year, they still consider their nation of origin France). Then if one considers all the various terminologies SNOMED CT, LOINC, ICD, RxNORM, WHODRUG, MedDRA, NDC and then CPT which is different in every country it adds further complexity. What data element concept does one start with and how is it used? This is a problem.

Thus, if one considers the research world and the role of the CDISC global standards used for submissions to regulatory bodies. It seems one can start there for the core data element concepts and point all the terminologies to the data element concepts (developed over 20 years through a global standards development process using consensus) the possibility of connecting research and healthcare through semantics becomes a possibility. Even when the problems are so vague and difficult, when everyone points to a specific discretely defined set of data elements, we can go from point A to point B seamless at each level. This is where the SNOMED CT terminology hierarchy can play a role in analysing and inter-connecting diverse interpretations of this data element (and potentially others). SNOMED CT terminology has layers upon layers that can aid machine learning and sending the right data to the right place. Connecting the SNOMED CT terminology to the others starting with the CDISC variables will yield an even more powerful world in clinical research.

Now consider connecting HL7 FHIR resources (starting with either an electronic health record or a personal health record or app of a person's choice for health data) to the CDISC variables enabling the data to go from point A (health) to point B (research), this will enable data aggregation across pharma studies, observational studies and now we can connect the distributed learning observational data from OMOP models via OHDSI in addition to other common data models.

2.6.2 Comparison of HL7 and CDISC Terminologies

Part of the assessment included a review of the HL7 FHIR IPS and CDISC terminologies comparing the HL7 FHIR IPS standard valuesets to the CDISC codelists (Figure 11). The two are semantically similar in many cases; however, the HL7 V2 and V3 lists are the same as the CDISC codelists. See Figure 12 and Figure 13 for details. The M, F, U were used with V2 and V3 and align exactly with the CDISC

terminology (with the exception of 'Intersex', which was added in 2023). Note that the terminology is updated and maintained biannually. Unfortunately, over time the standards terminologies started to diverge and that can be seen here, specifically looking at the FHIR codes.

Expansion based on AdministrativeGender v5.0.0 (CodeSystem)

Code	System	Display	Definition	v2 map for AdministrativeGender	v3 map for AdministrativeGender
male	http://hl7.org/fhir/administrative-gender	Male	Male.	~M	~M
female	http://hl7.org/fhir/administrative-gender	Female	Female.	~F	~F
other	http://hl7.org/fhir/administrative-gender	Other	Other.	>A *(Source concept 'other' is broader than target concept 'Ambiguous' because target concept does not include 'Other')* >O *(Source concept 'other' is broader than target concept 'Other' because target concept does not include 'Ambiguous')*	<UN
unknown	http://hl7.org/fhir/administrative-gender	Unknown	Unknown.	~U	~UNK

Figure 12: HL7 IPS Valuesets for Administrative Gender for FHIR, V2, and V3

Code	Codelist Code	Codelist Extensible (Yes/No)	Codelist Name	CDISC Submission Value	CDISC Synonym(s)	CDISC Definition	NCI Preferred Term
C66731		No	Sex	SEX	Sex	The assemblage of physical properties or qualities by which male is distinguished from female; the physical difference between male and female; the distinguishing peculiarity of male or female. (NCI)	CDISC SDTM Sex of Individual Terminology
C16576	C66731		Sex	F	Female	A person who belongs to the sex that normally produces ova. The term is used to indicate biological sex distinctions, or cultural gender role distinctions, or both. (NCI)	Female
C45908	C66731		Sex	INTERSEX		A person (one of unisexual specimens) who is born with genitalia and/or secondary sexual characteristics of indeterminate sex, or which combine features of both sexes. (NCI)	Intersex
C20197	C66731		Sex	M	Male	A person who belongs to the sex that normally produces sperm. The term is used to indicate biological sex distinctions, cultural gender role distinctions, or both. (NCI)	Male
C17998	C66731		Sex	U	U; UNK; Unknown	Not known, not observed, not recorded, or refused. (NCI)	Unknown

Figure 13: CDISC Codelist for Sex

To promote interoperability across countries for patient care and clinical research a smaller term list/set will be helpful in ensuring the same terms are used to enable exchange. An example of shorter codelist or valueset is the Dose Form list for the data collection standard (CDASH) shown in Figure 14. Contrast this to, Figure 15, the data tabulation (SDTM) codelist that has 190 terms and the IPS valueset that has 993 terms. The IPS list originates from the European Directorate for the Quality of Medicines and Healthcare Dose Form Codes (EDQM) and supports the product medicine quality use case thus requiring such granular specificity (see the section on UNICOM). See Figure 16 for a partial view of this[26]. The purpose of the CDASHIG Terminology lists is to provide a smaller more manageable list for case report forms. The SDTMIG codelist for Pharmaceutical Dose Form has 189 data element items for the Data Element Concept of Dose Form.

[26] For the full list of terms for dose form and copyright details see http://hl7.org/fhir/uv/ips/STU1.1/ValueSet-medicine-doseform.html

Code	Codelist Code	Codelist Extensible (Yes/No)	Codelist Name	CDISC Submission Value	CDISC Synonym(s)	CDISC Definition	NCI Preferred Term
C78418		Yes	Concomitant Medication Dose Form	CMDOSFRM	Concomitant Medication Dose Form	A terminology subset of the CDISC SDTM Pharmaceutical Dosage Form codelist created for CDASH Concomitant Medication Dose Form codelist. (NCI)	CDISC CDASH Concomitant Medication Dose Form Terminology
C42887	C78418		Concomitant Medication Dose Form	AEROSOL	aer	A product that is packaged under pressure and contains therapeutically active ingredients that are released upon activation of an appropriate valve system; it is intended for topical application to the skin as well as local application into the nose (nasal aerosols), mouth (lingual aerosols), or lungs (inhalation aerosols).	Aerosol Dosage Form
C25158	C78418		Concomitant Medication Dose Form	CAPSULE	cap	A solid pharmaceutical dosage form that contains medicinal agent within either a hard or soft soluble container or shell, usually used for the oral administration of medicine. The shells are made of a suitable form of gelatin or other substance. (NCI)	Capsule Dosage Form
C28944	C78418		Concomitant Medication Dose Form	CREAM		A semisolid emulsion of either the oil-in-water or the water-in-oil type, ordinarily intended for topical use. (NCI)	Cream Dosage Form
C42933	C78418		Concomitant Medication Dose Form	GAS		Any elastic aeriform fluid in which the molecules are separated from one another and have free paths. (NCI)	Gas Dosage Form
C42934	C78418		Concomitant Medication Dose Form	GEL		A semisolid (1) dosage form that contains a gelling agent to provide stiffness to a solution or a colloidal dispersion (2). A gel may contain suspended particles. Note 1: A semisolid is not pourable; it does not flow or conform to its container at room temperature. It does not flow at low shear stress and generally exhibits plastic flow behavior. Note 2: A colloidal dispersion is a system in which particles of colloidal dimension (i.e., typically between 1 nm and 1 micrometer) are distributed uniformly throughout a liquid.	Gel Dosage Form
C42966	C78418		Concomitant Medication Dose Form	OINTMENT	oint	A suspension or emulsion, semisolid (1) dosage form, usually containing less than 20 percent water and volatiles (2) and greater than 50 percent hydrocarbons, waxes, or polyols as the vehicle. This dosage form is generally for external application to the skin or mucous membranes. Note 1: A semisolid is not pourable; it does not flow or conform to its container at room temperature. It does not flow at low shear stress and generally exhibits plastic flow behavior. Note 2: Percent water and volatiles are measured by a loss on drying test in which the sample is heated at 105 degrees C until constant weight is achieved.	Ointment Dosage Form
C42968	C78418		Concomitant Medication Dose Form	PATCH		A drug delivery system that often contains an adhesive backing that is usually applied to an external site on the body. Its ingredients either passively diffuse from, or are actively transported from, some portion of the patch. Depending upon the patch, the ingredients are either delivered to the outer surface of the body or into the body. A patch is sometimes synonymous with the terms Extended Release Film and System.	Patch Dosage Form
C42972	C78418		Concomitant Medication Dose Form	POWDER		An intimate mixture of dry, finely divided drugs and/or chemicals that may be intended for internal or external use. (NCI)	Powder Dosage Form
C42989	C78418		Concomitant Medication Dose Form	SPRAY		A liquid minutely divided as by a jet of air or steam. (NCI)	Spray Dosage Form
C42993	C78418		Concomitant Medication Dose Form	SUPPOSITORY	supp	A solid body of various weights and shapes, adapted for introduction into the rectal, vaginal, or urethral orifice of the human body; they usually melt, soften, or dissolve at body temperature.	Suppository Dosage Form
C42994	C78418		Concomitant Medication Dose Form	SUSPENSION	Ready to Use Suspension; susp	A liquid dosage form that contains solid particles dispersed in a liquid vehicle. Note: A liquid is pourable; it flows and conforms to its container at room temperature. It displays Newtonian or pseudoplastic flow behavior.	Suspension Dosage Form
C42998	C78418		Concomitant Medication Dose Form	TABLET	tab	A solid dosage form containing medicinal substances with or without suitable diluents. (NCI)	Tablet Dosage Form

Figure 14: CDISC CDASH Codelist for Dose Form

Code	Codelist Code	Codelist Extensible (Yes/No)	Codelist Name	CDISC Submission Value	CDISC Synonym(s)	CDISC Definition	NCI Preferred Term
C66726		Yes	Pharmaceutical Dosage Form	FRM	Pharmaceutical Dosage Form	The form of the completed pharmaceutical product, e.g. tablet, capsule, injection, elixir, suppository. Dosage form can have a significant effect on the onset, duration and intensity of the pharmacological action of a drug. A pharmaceutical dosage form controls the rate at which the drug is released into the biological fluids. This release rate affects its intrinsic absorption pattern and therefore, the bioavailability of the drug.	CDISC SDTM Pharmaceutical Dosage Form Terminology
C42887	C66726		Pharmaceutical Dosage Form	AEROSOL	aer	A product that is packaged under pressure and contains therapeutically active ingredients that are released upon activation of an appropriate valve system; it is intended for topical application to the skin as well as local application into the nose (nasal aerosols), mouth (lingual aerosols), or lungs (inhalation aerosols).	Aerosol Dosage Form
C42888	C66726		Pharmaceutical Dosage Form	AEROSOL, FOAM		A dosage form containing one or more active ingredients, surfactants, aqueous or non-aqueous liquids, and the propellants; if the propellant is in the internal (discontinuous) phase (i.e., of the oil-in-water type), a stable foam is discharged, and if the propellant is in the external (continuous) phase (i.e., of the water-in-oil type), a spray or a quick-breaking foam is discharged.	Aerosol Foam Dosage Form
C42960	C66726		Pharmaceutical Dosage Form	AEROSOL, METERED		A pressurized dosage form consisting of metered dose valves which allow for the delivery of a uniform quantity of spray upon each activation. (NCI)	Metered Aerosol Dosage Form
C42971	C66726		Pharmaceutical Dosage Form	AEROSOL, POWDER		A product that is packaged under pressure and contains therapeutically active ingredients, in the form of a powder, that are released upon activation of an appropriate valve system. (NCI)	Powder Aerosol Dosage Form
C42889	C66726		Pharmaceutical Dosage Form	AEROSOL, SPRAY		An aerosol product which utilizes a compressed gas as the propellant to provide the force necessary to expel the product as a wet spray; it is applicable to solutions of medicinal agents in aqueous solvents. (NCI)	Aerosol Spray Dosage Form
C42892	C66726		Pharmaceutical Dosage Form	BAR, CHEWABLE		A solid dosage form usually in the form of a rectangle that is meant to be chewed. (NCI)	Chewable Bar Dosage Form
C42890	C66726		Pharmaceutical Dosage Form	BEAD		A solid dosage form in the shape of a small ball. (NCI)	Bead Dosage Form
C43451	C66726		Pharmaceutical Dosage Form	BEAD, IMPLANT, EXTENDED RELEASE		A small sterile solid mass consisting of a highly purified drug intended for implantation in the body which would allow at least a reduction in dosing frequency as compared to that drug presented as a conventional dosage form. (NCI)	Extended Release Bead Implant Dosage Form
C42891	C66726		Pharmaceutical Dosage Form	BLOCK		Solid dosage form, usually in the shape of a square or rectangle. (NCI)	Block Dosage Form
C97197	C66726		Pharmaceutical Dosage Form	CAPLET		A solid dosage form in which a tablet has been compacted into capsule shape.	Caplet Dosage Form
C25158	C66726		Pharmaceutical Dosage Form	CAPSULE	cap	A solid pharmaceutical dosage form that contains medicinal agent within either a hard or soft soluble container or shell, usually used for the oral administration of medicine. The shells are made of a suitable form of gelatin or other substance. (NCI)	Capsule Dosage Form
C42896	C66726		Pharmaceutical Dosage Form	CAPSULE, COATED PELLETS		A solid dosage form in which the drug is enclosed within either a hard or soft	Coated Pellet in Capsule Dosage Form

Figure 15: CDISC SDTM Codelist for Pharmaceutical Dosage Form

14.62.1.2 Expansion

This value set contains 993 concepts

Expansion based on http://standardterms.edqm.eu version 5 March 2019

Code	System	Display
10100500	http://standardterms.edqm.eu	Concentrate for oral suspension
10101000	http://standardterms.edqm.eu	Oral drops, solution
10102000	http://standardterms.edqm.eu	Oral drops, suspension
10103000	http://standardterms.edqm.eu	Oral drops, emulsion
10104000	http://standardterms.edqm.eu	Oral liquid
10105000	http://standardterms.edqm.eu	Oral solution
10106000	http://standardterms.edqm.eu	Oral suspension
10107000	http://standardterms.edqm.eu	Oral emulsion
10108000	http://standardterms.edqm.eu	Oral gel
10109000	http://standardterms.edqm.eu	Oral paste
10110000	http://standardterms.edqm.eu	Powder for oral solution
10111000	http://standardterms.edqm.eu	Powder for oral suspension
10112000	http://standardterms.edqm.eu	Granules for oral solution
10113000	http://standardterms.edqm.eu	Granules for oral suspension
10114000	http://standardterms.edqm.eu	Powder and solvent for oral solution
10115000	http://standardterms.edqm.eu	Powder and solvent for oral suspension
10115500	http://standardterms.edqm.eu	Powder and solvent for syrup
10116000	http://standardterms.edqm.eu	Lyophilisate for suspension
10117000	http://standardterms.edqm.eu	Syrup
10118000	http://standardterms.edqm.eu	Powder for syrup
10119000	http://standardterms.edqm.eu	Granules for syrup
10120000	http://standardterms.edqm.eu	Soluble tablet
10121000	http://standardterms.edqm.eu	Dispersible tablet
10121500	http://standardterms.edqm.eu	Dispersible tablets for dose dispenser
10122000	http://standardterms.edqm.eu	Herbal tea
10201000	http://standardterms.edqm.eu	Oral powder
10202000	http://standardterms.edqm.eu	Instant herbal tea
10203000	http://standardterms.edqm.eu	Effervescent powder
10204000	http://standardterms.edqm.eu	Granules
10205000	http://standardterms.edqm.eu	Effervescent granules
10206000	http://standardterms.edqm.eu	Gastro-resistant granules
10207000	http://standardterms.edqm.eu	Prolonged-release granules

Figure 16: HL7 IPS Valueset for Medicine Doseform

This subset approach is a consideration for Codelists and Valuesets for xShare Core Set of Data Element Concepts for IPS+R. The shorter list approach would also benefit the provider documentation burden; the lists that pop up for providers to choose from are vast and overwhelming.

Part of the vision for the IPS+R is streamlining the data from the health records to the research records. When going from real world data to CDISC domains and codelists, the coding system code and label meaning will drive where the data goes into SDTM. For example, laboratory tests are Observations in HL7 FHIR with the specific terminology defining the test that occurred. The CDISC domains have multiple domains for laboratory data since it is a heterogeneous mix of data types. For example, a routine laboratory test would go into the Laboratory Findings domain (LB), specimen results for detection, quantification, and other characterizations of microorganisms go in the Microbiology domain (MB), and histopathology findings and microscopic findings go in the Microscopic Findings (MI) domain. A process to develop the IPS+R Data Element Codelists could be developed as shown previously in Figure 11. Start with the CDISC Codelist (where applicable), since

CDISC is the global standard for randomised controlled trials, then translate the codelist to a valueset in SNOMED CT or the preferred terminology for healthcare/EHRs.

3 Gap Analysis: IPS and Core Data Element List and Definitions

3.1 Core Elements Set with Definitions

Table 1 provides a list of the proposed xShare Core Data Elements for the IPS+R and their definition. **The light blue rows highlight the main data element concept (Category/Domain)** for Medication, Procedure, Substance Use, Healthcare Encounter, Medical History (equivalent to Problem, Diagnosis, or Condition), Reproductive Findings (Pregnancy related), Vital Signs, Subject/Demographic, Body System Diagnostic Test Results and Laboratory Results (both general laboratory and microbiology). After initial review, the Adverse Event data elements were added for research purposes. CDISC definitions were used to stay consistent. The rows without colour indicate the 'attributes or metadata' elements associated with the main data element concept (Category/Domain). The complete table of the core element data set is included in appendix I and compared to IPS and USCDI 4.0 (see Appendix).

Table 1: xShare Core Data Element Set

xShare Core Data Element	Definition
Reported Name of Drug, Med, or Product	Verbatim medication name or treatments (include only treatments with data collection characteristics similar to medications).
CM Dose per Administration	The dose of medication/treatment (e.g., --TRT) given at one time, represented as a numeric value.
Concomitant Meds Dose Description	The dose of medication/treatment taken per administration.
CM Dose Units	The unit associated with the concomitant medication/treatment/therapy taken (e.g., mg in "2 mg 3 times per day").
CM Indication	The condition, disease, symptom, or disorder that the concomitant (non-study) medication/treatment/therapy was used to address or investigate (e.g., why the medication/treatment/therapy was taken or administered).
CM Dose Form	The pharmaceutical dosage form in which the CMTRT is physically presented.
CM Dosing Frequency per Interval	The number of doses given/administered/taken during a specific interval.
CM Route of Administration	The route of administration of the concomitant medication/treatment/therapy.
Concomitant Meds Start Date	The start date is when the concomitant medication/treatment/therapy was first taken, represented in an unambiguous date format (e.g., DD-MON-YYYY).
Ongoing Concomitant Meds	Indication the concomitant medication/treatment/therapy is ongoing when no end date is provided.
Concomitant Meds End Date	The date that the subject ended/stopped taking the concomitant medication/treatment/therapy, represented in an unambiguous date format (e.g., DD-MON-YYYY).
Reported Name of Procedure	The verbatim surgical, therapeutic, or diagnostic procedure's name.
Procedure Start Date	The date or start date of when the procedure started or was performed, represented in an unambiguous date format (e.g., DD-MON-YYYY).
Procedure Indication	The condition, disease, symptom, or disorder that the procedure was used to address or investigate (e.g., why the therapy was taken or administered, why the procedure was performed).
Ongoing Procedure	Indication the procedure is ongoing when no end date is provided.
Procedure End Date	The end date of the procedure, represented in an unambiguous date format (e.g., DD-MON-YYYY).

xShare Core Data Element	Definition
Reported Name of Substance	The type of substance (e.g., TOBACCO, ALCOHOL, CAFFEINE or CIGARETTES, CIGARS, COFFEE).
Never Current Former Usage	Indication the prespecified substance was used.
Substance Dose Description	The amount of substance used (e.g., 1-2 packs, 8 oz).
Substance Dose per Administration	The dose of substance (e.g., --TRT) taken at one time, represented as a numeric value.
Substance Dose Units	The unit associated with the substance taken (e.g., pack in "1 pack per day").
Substance Use Frequency per Interval	The number/amount of the of substance consumed per a specific interval.
Substance Use Start Date	The date substance use started, represented in an unambiguous date format (e.g., DD-MON-YYYY).
Substance Use End Date	The date substance use ended, represented in an unambiguous date format (e.g., DD-MON-YYYY).
Substance Use Collected Duration	Collected duration of the substance use.
Substance Use Collected Duration Unit	Unit of the collected duration of the substance use. Used only if duration was collected on the CRF.
Reported Term for Healthcare Encounter	The reported or prespecified name of the healthcare encounter.
Healthcare Encounter Start Date	The start date the healthcare encounter, represented in an unambiguous date format (e.g., DD-MON-YYYY).
Healthcare Encounter End Date	The end date of the healthcare encounter (e.g., date of discharge), represented in an unambiguous date format (e.g., DD-MON-YYYY).
Reason for the Healthcare Encounter	Denotes the reason for the healthcare encounter.
Medical History Collection Date	The date on which the medical history was collected, represented in an unambiguous date format (e.g., DD-MON-YYYY).
Reported Term for the Medical History	The reported or prespecified name of the medical condition or event.
Medical History Event Start Date	The start date of medical history event or condition, represented in an unambiguous date format (e.g., DD-MON-YYYY).
Medical History Event End Date	The end date of medical history event or condition, represented in an unambiguous date format (e.g., DD-MON-YYYY).
Reproductive System Findings Test Name	Descriptive name for reproductive system finding.
RP Result or Finding in Original Units	Result of the finding defined in reproductive system finding, as originally received or collected.
RP Original Units	The unit of the result as originally received or collected.
Reproductive System Finding Date	The date on which the reproductive system result or finding was collected, represented in an unambiguous date format (e.g., DD-MON-YYYY).
Subject Characteristic	Descriptive name of the subject characteristic of interest.
Subject Characteristic Collection Date	The date of collection represented in an unambiguous date format (e.g., DD-MON-YYYY).
SC Result or Finding in Original Units	Result of the subject characteristic as originally received or collected.

xShare Core Data Element	Definition
SC Result or Finding in Original Units	Result of the subject characteristics as originally received or collected.
Vital Signs Test Name	Descriptive name of the test or examination used to obtain the measurement or finding.
Vital Signs Date	The date of the vital signs measurement, represented in an unambiguous date format (e.g., DD-MON-YYYY).
Vital Signs Time	The time of measurement, represented in an unambiguous time format (e.g., hh:mm:ss).
VS Result or Finding in Original Units	Result of the vital signs measurement as originally received or collected.
VS Original Units	The unit of the result as originally received or collected.
Body System Diagnostic Test Name	Descriptive name of the test or examination used to obtain the measurement or finding.
Body System Diagnostic Test Result in Original Units	Result of the measurement or finding as originally received or collected.
Body System Diagnostic Test Original Unit	The unit of the result as originally received or collected.
Body System Diagnostic Test Anatomical Location	Anatomical location used for the measurement.
Specimen Collection Date	The date of specimen collection, represented in an unambiguous date format (e.g., DD-MON-YYYY).
Specimen Collection Time	The time of specimen collection, represented in an unambiguous time format (e.g., hh:mm:ss).
Lab Specimen Type	The type of sample material taken from a biological entity.
Lab Fasting Status	An indication that the subject has abstained from food/water for the specified amount of time.
Lab Test or Examination Name	Descriptive name of the lab test or examination used to obtain the measurement or finding. Any test normally performed by a clinical laboratory is considered a lab test.
Lab Result or Finding in Original Units	Result of the measurement or finding as originally received or collected.
Lab Original Units	The unit of the result as originally received or collected.
Lab Specimen ID	An internal or external identifier (e.g., specimen identifier).
Lab Method of Test or Examination	Method of the test or examination.
Lab Ref Range Lower Limit in Orig Unit	The lower end of normal range or reference range for continuous results stored in LBORRES.
Lab Ref Range Upper Limit in Orig Unit	The upper end of normal range or reference range for continuous results stored in LBORRES.
Lab Reference Range Indicator	An indication or description of how the value compares to the normal range or reference range.
MB Specimen Collection Date	The date of specimen collection, represented in an unambiguous date format (e.g., DD-MON-YYYY).
MB Specimen Collection Time	The time of specimen collection, represented in an unambiguous time format (e.g., hh:mm:ss).
Microbiology Test or Finding Name	Descriptive name of the microbiology test or examination used to obtain the measurement or finding.

xShare Core Data Element	Definition
MB Result or Finding in Original Units	Result of the measurement or finding as originally received or collected.
MB Original Units	The unit of the result as originally received or collected.
MB Specimen Type	The type of specimen used for a measurement.
MB Method of Test or Examination	Method of the test or examination.
MB Reference ID	An internal or external identifier such as specimen identifier.
MB Specimen Collection Location	A description of the anatomical location of the subject relevant to the collection of specimen.
Microbiology Examination Detail	Detail of the microbiology examination used to obtain the measurement or finding.
Reported Term for the Adverse Event	The reported or prespecified name of the adverse event.
Adverse Event Start Date	The start date of the adverse event, represented in an unambiguous date format (e.g., DD-MON-YYYY).
Start Time of Adverse Event	The start time of the adverse event, represented in an unambiguous time format (e.g., hh:mm:ss).
Ongoing Adverse Event	Indication that an adverse event is ongoing when no end date is provided.
Adverse Event End Date	The date when the adverse event resolved/ended, represented in an unambiguous date format (e.g., DD-MON-YYYY).
End Time of Adverse Event	The time when the adverse event ended/resolved, represented in an unambiguous time format (e.g., hh:mm:ss).
AE Severity/Intensity	The severity or intensity of the event.
AE Standard Toxicity Grade	The grade of the severity of the event using a standard "toxicity" scale (e.g., NCI CTCAE).
AE Serious Event	An indication of whether the adverse event is determined to be "serious," based on what is defined in the regulations/protocol.
AE Results in Death	An indication the serious adverse event resulted in death.
AE Is Life Threatening	An indication the serious adverse event was life threatening.
AE Requires or Prolongs Hospitalization	An indication the serious adverse event resulted in an initial or prolonged hospitalization.
AE Persistent or Significant Disability/Incapacity	An indication the serious adverse event was associated with a persistent or significant disability or incapacity.
AE Needs Intervention to Prevent Impairment	An indication an adverse event required medical or surgical intervention to preclude permanent impairment of a body function, or prevent permanent damage to a body structure, due to the use of a medical product.
AE Other Medically Important Serious Event	An indication additional categories for seriousness apply.
AE Involves Cancer	An indication the serious event was associated with the development of cancer.
AE Causality	An indication the study treatment had a causal effect on the adverse event, as determined by the clinician/investigator.
Actions Taken with Device	A description of the action taken, with respect to a device used in a study (which may or may not be the device under study), as a result of the event.

xShare Core Data Element	Definition
Any Other Actions Taken	An indication whether any other actions were taken in response to the adverse event that were unrelated to study treatment dose changes or other non-study treatments given because of this adverse event.
Outcome of Adverse Event	A description of the outcome of an event.
Research Study Identifier	A unique identifier for a study.
Research Subject Identifier for the Study	A unique subject identifier within a site and a study.
Unique Subject Identifier for the Study	Identifier used to uniquely identify a subject across all studies for all applications or submissions involving the product. This must be a unique value, and could be a compound identifier formed by concatenating STUDYID-SITEID-SUBJID.
Age	The age of the subject, expressed in AGEU.
Age Units	Units of time routinely used to express the age of a person.
Demographics Collection Date	The date of collection, represented in an unambiguous date format (e.g., DD-MON-YYYY).
Sex	Sex of the subject, as determined by the investigator.
Birth Date	A subject's date of birth (with or without the time of birth). The complete Date of Birth is made from the temporal components of Birth Year, Birth Month, Birth Day, and Birth Time.
Date/Time of Death	Date/time of death for any subject who died, in ISO 8601 format. Should represent the date/time that is captured in the clinical-trial database.
Subject Death Flag	Indicates the subject died. Should be "Y" or null. Should be populated even when the death date is unknown

3.2 Proposed Core Data Elements for IPS + R Indicating Hierarchy and Current 'Gaps'

Using the prior definitions another way to review the proposed Core Data Elements for xShare IPS + R is to view them in categories (or research domains) that indicate the data elements and the essential metadata/variables for each data item. These are the recommended categories for this initial working paper. See the table in Section 4.3 for a gap analysis and prospective HL7 FHIR IPS additions to address these gaps.

3.2.1 Category for Patient Information or Subject Characteristics Domain

The IPS Patient Information Category is basically equivalent with the CDISC Subject Characteristics and Demographics Domains. The key elements are Subject Characteristic Item, Result and Collection Date. Each of these elements is then associated with the requested item list in the lower part of the diagram: Age, Birth Data, Sex/Gender and Death (Figure 17, Table 2). Associated with each of these items is the collection date and some have additional essential metadata, such as age units. It should be noted that Gender in the HL7 FHIR IPS IG and Sex in CDISC are not precisely specified. It is assumed here to be an administrative interpretation. It may be important in future versions of the IPS+R to distinguish data elements for sex reported at birth, gender identity, anatomical sex and sexual orientation. The Research Subject Identifier is a number provided when a person enters a

research study. It is unique and allows the clinician to match or re-identify the research subject with a patient since research subject PHI is confidential to the study sponsor. Required for IPS and Research (Gap: Research Subject Identifier for Research).

Table 2 Patient Information Category in IPS Corresponds to CDISC Domains for Demographics and Subject Characteristics

Data Element	DRAFT CDASHIG Definition	Question Text	Prompt
Subject Characteristic Collection Date	The date of collection represented in an unambiguous date format (e.g., DD-MON-YYYY).	What was the date the subject characteristics were collected?	Date
Subject Characteristic	Descriptive name of the subject characteristic of interest.	What is the subject characteristics name?	[Subject Characteristic Test Name]
Subject Characteristic Finding Value	Result of the subject characteristic as originally received or collected.	What is the subject characteristic?	(Result)
Subject Characteristic Finding Value Units	Result of the subject characteristics as originally received or collected.	What is the subject's [SCTEST]?	[SCTEST] Result
Study Identifier	A unique identifier for a study.	What is the study identifier?	[Protocol/Study]
Study Site Identifier	A unique identifier for a site within a study.	What is the site identifier?	Site (Identifier)
Research Subject Identifier for the study	A unique subject identifier within a site and a study.	What [is/was] the (study) [subject/participant] identifier?	[Subject/Participant] (Identifier)
Research Unique Subject Identifier	Identifier used to uniquely identify a subject across all studies for all applications or submissions involving the product. This must be a unique value, and could be a compound identifier formed by concatenating STUDYID-SITEID-SUBJID.	What [is/was] the (StudyID)-(SiteID) (study) [subject/participant] identifier?	[STUDYID-SITEID-SUBJID]
Research Study Identifier	A unique identifier for a study.	What is the study identifier?	[Protocol/Study]
Age	The age of the subject, expressed in AGEU.	What is the subject's age?	Age
Age Units	Units of time routinely used to express the age of a person.	What is the age unit used?	Age Unit
Demographics Collection Date	The date of collection, represented in an unambiguous date format (e.g., DD-MON-YYYY).	What is the date of collection?	Collection Date
Sex (Administrative/clinical use?)	Sex of the subject, as determined by the investigator.	What is the sex of the subject?	Sex
Birth Date	A subject's date of birth (with or without the time of birth). The complete Date of Birth is made from the temporal components of Birth Year, Birth Month, Birth Day, and Birth Time..	What is the subject's date of birth?	Birth Date
Deceased Date	Date/time of death for any subject who died, in ISO 8601 format. Should represent the date/time that is captured in the clinical-trial database.	What was the subject's date/time of death?	Death Date
Deceased Flag	Indicates the subject died. Should be "Y" or null. Should be populated even when the death date is unknown	Was the subject dead?	Subject Death Flag
STUDYID	What is the study identifier?	[Protocol/Study]	

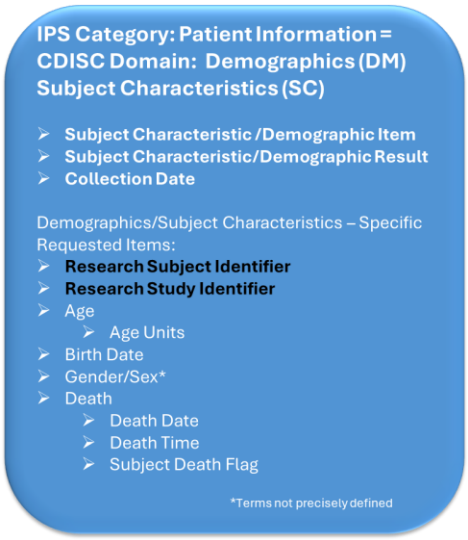

Figure 17: Patient Information Category in IPS Corresponds to CDISC Domains for Demographics and Subject Characteristics

3.2.2 Category of Problem List or Domain for Medical History

The IPS Category of Problem List is basically the equivalent of the CDISC Domain for Medical History. The Elements would consist of terms from a look-up/valuelist related to condition or problem (Figure 18, Table 3). The metadata would be the Event Start Date and End Date (or if there is no End Date, is it likely to be Ongoing for an extended period). Examples could include Alcohol Abuse or many others. It is required for Required for IPS and Research.

Table 3: Problem list Category in IPS Corresponds to CDISC Domains for Medical History

Data Element	DRAFT CDASHIG Definition	Question Text	Prompt
Medical History Collection Date	The date on which the medical history was collected, represented in an unambiguous date format (e.g., DD-MON-YYYY).	What was the date the medical history was collected?	Collection Date
Medical History Reported Term	The reported or prespecified name of the medical condition or event.	What is the medical condition or event term?	Medical History Term
Medical History Event Start Date	The start date of medical history event or condition, represented in an unambiguous date format (e.g., DD-MON-YYYY).	What [is/was] the [medical event or condition/category of the event] start date?	Start Date
Medical History Event End Date	The end date of medical history event or condition, represented in an unambiguous date format (e.g., DD-MON-YYYY).	What[is/was] the[medical event or condition/category of the event] end date?	End Date

Figure 18: Problem List Category in IPS (current or past) Corresponds to the CDISC Domain for Medical History

3.2.3 Category or Domain for Medications

Medications that a patient is currently taking are listed with the Name (look-up) along with Dose and Indication (Figure 19, Table 4). The Dose Element has a number of important variables or metadata items. The medication Start Date is important. The presence of an End Date is also important as its absence would normally imply that this medicine is an ongoing treatment.

Table 4: The CDISC Elements for the Domain of Concomitant Medications (IPS Category of Medication Summary)

Data Element	DRAFT CDASHIG Definition	Question Text	Prompt
Medication	Verbatim medication name or treatments (include only treatments with data collection characteristics similar to medications).	What was the (concomitant) [medication/treatment/therapy] (name/term)?	(Concomitant) [Medication/Treatment/Therapy]
Medication Dose	The dose of medication/treatment (e.g., --TRT) given at one time, represented as a numeric value.	What was the individual dose (of the concomitant (medication/treatment/therapy] per administration)?	[Dose/Amount] (per administration)
Medication Dose Text (for range doses)	The dose of medication/treatment taken per administration.	What was the individual dose of the (concomitant) [medication/treatment/therapy]?	Dose
Medication Dose Unit	The unit associated with the concomitant medication/treatment/therapy taken (e.g., mg in "2 mg 3 times per day").	What is the unit (for the dose of concomitant [medication/treatment/therapy])?	(Dose) Unit
Medication Indication	The condition, disease, symptom, or disorder that the concomitant (non-study) medication/treatment/therapy was used to address or investigate (e.g., why the medication/treatment/therapy was taken or administered).	For what indication was the (concomitant) [medication/treatment/therapy] taken?	Indication
Medication Dose Form	The pharmaceutical dosage form in which the CMTRT is physically presented.	What was the dose form of the (concomitant)]medication/treatment/therapy]?	Dose Form
Medication Dose Frequency	The number of doses given/administered/taken during a specific interval.	What was the frequency of the (concomitant) [medication/treatment/therapy]?	Frequency
Medication Route of Administration	The route of administration of the concomitant medication/treatment/therapy.	What was the route of administration of the (concomitant) [medication/treatment/therapy]?	Route
Medication Start Date	The start date is when the concomitant medication/treatment/therapy was first taken, represented in an unambiguous date format (e.g., DD-MON-YYYY).	What was the (concomitant) [medication/treatment/therapy/dose] start date?	Start Date
Medication Ongoing	Indication the concomitant medication/treatment/therapy is ongoing when no end date is provided.	Was the (concomitant) [medication/treatment/therapy] ongoing (as of [the study-specific time point or period])?	Ongoing (as of [the study-specific time point or period])
Medication End Date	The date that the subject ended/stopped taking the concomitant medication/treatment/therapy, represented in an unambiguous date format (e.g., DD-MON-YYYY).	What was the (concomitant) [medication/treatment/therapy/dose] end date?	End Date

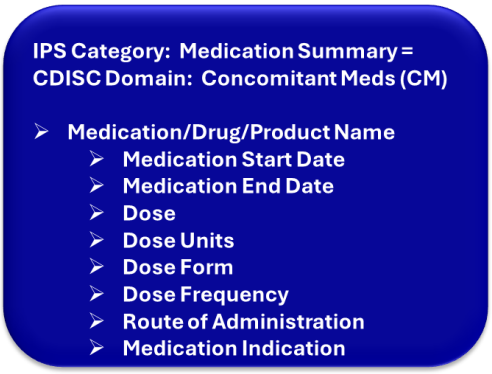

Figure 19: Medication Summary Category in the IPS Corresponds to the CDISC Domain for Concomitant Medications

3.2.4 Category or Domain for Procedures

The Procedures domain or History of Procedures Category consists of a name of a procedure (from a look up list or value list), along with the indication for the procedure and Start Date or End Date (or if no End Date, is likely to be Ongoing) (see Figure 20, Table 5). Examples of procedures could be a stress test, a colonoscopy or others.

Table 5: The Data Elements for the CDISC Domain of Procedure (IPS Category of History of Procedures)

Data Element	DRAFT CDASHIG Definition	Question Text	Prompt
Procedure	The verbatim surgical, therapeutic, or diagnostic procedure's name.	What was the procedure name?	[Procedure Name]; (Specify) Other
Procedure Start Date	The date or start date of when the procedure started or was performed, represented in an unambiguous date format (e.g., DD-MON-YYYY).	What was the procedure (start) date ?	(Start) Date
Procedure Indication	The condition, disease, symptom, or disorder that the procedure was used to address or investigate (e.g., why the therapy was taken or administered, why the procedure was performed).	For what indication was the [PRTRT] performed?	Indication
Procedure Ongoing	Indication the procedure is ongoing when no end date is provided.	Was the procedure ongoing (as of the [study-specific timepoint or period])?	Ongoing (as of the [study-specific timepoint or period])
Procedure End Date	The end date of the procedure, represented in an unambiguous date format (e.g., DD-MON-YYYY).	What was the procedure (end) date?	(End) Date

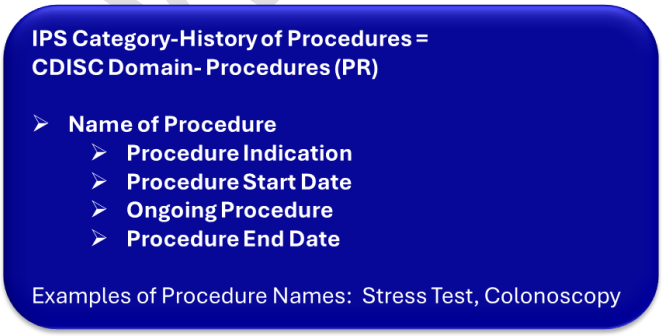

Figure 20: IPS Category of Procedures Corresponds to the CDISC Domain of Procedures (PR).

3.2.5 Domain of Reproductive Status

This domain is important for research since it is frequent for research protocols to exclude women who are pregnant or trying to conceive (see Figure 21, Table 6). There is typically a question in case report forms that hides this data field if the patient is male or post-menopausal. Other possible data elements to include here are menopause status and age at menarche.

Table 6: The Elements of the CDISC Domain for Reproductive Findings (IPS Category of Pregnancy Status)

Data Element	DRAFT CDASHIG Definition	Question Text	Prompt
Reproductive Finding Name	Descriptive name for reproductive system finding.	What is the reproductive finding name?	[Reproductive System Findings Test Name]
Reproductive Finding Result Value	Result of the finding defined in reproductive system finding, as originally received or collected.	What was the result for the reproductive system question?	(Result)
Reproductive Finding Result Value Units	The unit of the result as originally received or collected.	What was the unit of the result?	Unit
Reproductive Finding Date	The date on which the reproductive system result or finding was collected, represented in an unambiguous date format (e.g., DD-MON-YYYY).	What was the date the reproductive system question was collected?	Collection Date

Figure 21: IPS Category for Pregnancy Status Corresponds to the CDISC Domain for Reproductive Findings

3.2.6 Vital Signs Category or Domain

The Vital Signs category or domain is equivalent in both the IPS and for research (Table 7, Figure 22). The metadata that are important for a given vital sign item (test) are date, time and units. There are a number of vital sign tests that are desired; these are listed and may include additional information, such as whether blood pressure is diastolic or systolic and, depending on the use, the position of the patient when it is measured. Vital signs are considered optional for the IPS but they are quite important in many research studies. The vital signs tests of Blood Pressure (D/S), Heart Rate, Respiratory Rate, Oxygen Saturation, Body Height and Weight, and Body Surface Area are required in clinical research.

Table 7: CDISC Elements for the Domain of Vital Signs (IPS Category Vital Signs)

Data Element	DRAFT CDASHIG Definition	Question Text	Prompt
Vital Signs Test Name	Descriptive name of the test or examination used to obtain the measurement or finding.	What is the vital sign test name?	[Vital Signs Test Name]
Vital Signs Test Date	The date of the vital signs measurement, represented in an unambiguous date format (e.g., DD-MON-YYYY).	What was the date of the measurement(s)?	[VSTEST] Date
Vital Signs Test Time	The time of measurement, represented in an unambiguous time format (e.g., hh:mm:ss).	What was the time of the measurement(s)?	[VSTEST] Time
Vital Signs Result Value	Result of the vital signs measurement as originally received or collected.	What was the result of the [VSTEST] measurement?	[VSTEST] (Result)
Vital Signs Result Value Units	The unit of the result as originally received or collected.	What was the unit of the [VSTEST] measurement?	[VSTEST] Unit

Figure 22 The IPS Category of Vital Signs Corresponds to the CDISC Domain for Vital Signs

3.2.7 Diagnostic Results Category or Diagnostic Findings and Lab and Microbiology Results

The IPS Diagnostics Results Category include all diagnostic test results which include body system tests, laboratory tests, and microbiology tests (Figure 23). In CDISC these Findings Class domains follow a similar structure. The diagnostic tests/diagnostic findings such as radiology reports would be represented in the domain related to the topic of the body system finding (e.g., chest x-ray would go in Respiratory System Findings (RE) domain or echocardiography diagnostic results would go in Cardiovascular System Findings (CV).

The CDISC specimen domains have different domains related to the nature of the test. The Test Name would be from a list of lab or microbiology tests. The necessary metadata include units and collection date and time. It may also be necessary, depending on the test, to include location of the test or method. For more details, for laboratory tests see Table 8, and for microbiology tests see Table 9. The priority list of laboratory tests that have been considered of high importance for research and are shown in Figure 24 and for microbiology tests in Figure 25.

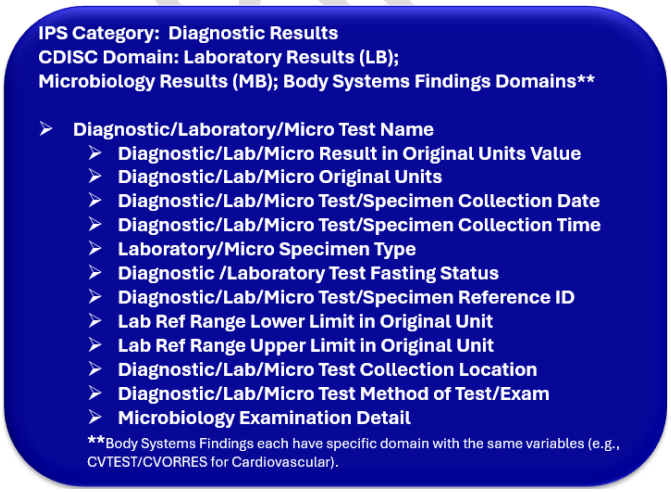

Figure 23: IPS Category for Diagnostic Results Corresponds to the CDISC Domains for Laboratory and Microbiology

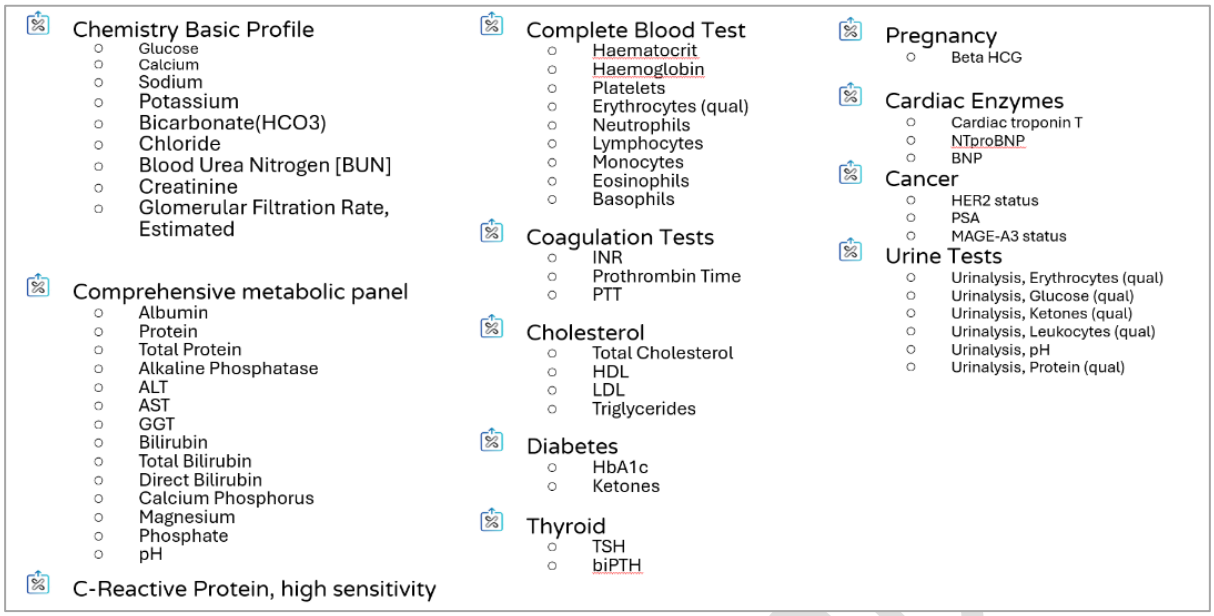

Figure 24: Priority List of Laboratory Tests for Clinical Research identified for the xSHARE Core Element Set.

1. ADENOVIRUS (RESPIRATORY)
2. BARTONELLA
3. BORDETELLA PERTUSSIS
4. BORRELIA BURGDORFERI
5. CAMPYLOBACTER
6. CHLAMYDIA PSITTACI
7. CHLAMYDIA TRACHOMATIS
8. CRYPTOCOCCUS
9. CRYPTOSPORIDIUM
10. CYCLOSPORA
11. ENTAMOEBA HISTOLYTICA
12. ENTEROVIRUS IN CSF
13. ESCHERIA COLI (VTEC - EHEC - STEC)
14. GIARDIA
15. HAEMOPHILUS INFLUENZAE (invasive disease)
16. HANTAVIRUS
17. HEPATITIS A VIRUS (HAV)
18. HEPATITIS B VIRUS (HBV)
19. HEPATITIS C VIRUS
20. HEPATITIS E VIRUS
21. HERPES SIMPLEX 3 VIRUS - VARICELLA ZOSTER VIRUS (Encephalitis)
22. HUMAN RESPIRATORY SYNCTIAL VIRUS (RSV)
23. HERPES SIMPLEX VIRUS 1-2 in CSF
24. HUMAN RESPIRATORY SYNCTIAL VIRUS (RSV)
25. INFLUENZA (A+B)
26. LEGIONELLA PNEUMOPHILA
27. LISTERIA
28. METAPNEUMOVIRUS
29. MMSE
30. MONKEYPOX
31. MORBILLIVIRUS (MEASLES/MAZELEN/ROUGEOLE)
32. MUMPS VIRUS (BOF/OREILLONS)
33. MYCOPLASMA PNEUMONIAE
34. NEISSERIA GONORRHOEAE
35. NEISSERIA MENINGITIDIS (INVASIVE DISEASE)
36. NOROVIRUS
37. PARAINFLUENZA
38. PLASMODIUM
39. ROTAVIRUS
40. RUBIVIRUS (RUBELLA/RUBéOLE)
41. SALMONELLA
42. SARS-CoV-2
43. SHIGELLA
44. STREPTOCOCCUS PNEUMONIAE
45. STREPTOCOCCUS PYOGENES
46. TREPONEMA PALLIDUM (ACTIVE)
47. YERSINIA ENTEROCOLITICA

Figure 25:Priority List of Microbiology Tests identified for the xSHARE Core Element Set

Table 8: CDISC Elements for the Laboratory Domain (IPS Category for Diagnostic Results)

Data Element	DRAFT CDASHIG Definition	Question Text	Prompt
Laboratory Specimen Collection Date	The date of specimen collection, represented in an unambiguous date format (e.g., DD-MON-YYYY).	What was the (start) date of the lab specimen collection?	Specimen Collection (Start) Date
Laboratory Specimen Collection Time	The time of specimen collection, represented in an unambiguous time format (e.g., hh:mm:ss).	What was the (start) time of the lab specimen collection?	Specimen Collection (Start) Time
Laboratory Specimen Type	The type of sample material taken from a biological entity.	What is the specimen material type?	Specimen Type
Laboratory Fasting Status	An indication that the subject has abstained from food/water for the specified amount of time.	Was the subject fasting?	Fasting
Laboratory Test Name	Descriptive name of the lab test or examination used to obtain the measurement or finding. Any test normally performed by a clinical laboratory is considered a lab test.	What was the lab test name?	[Laboratory Test Name]
Laboratory Result Value	Result of the measurement or finding as originally received or collected.	What was the result of the lab test?	(Result)
Laboratory Result Value Units	The unit of the result as originally received or collected.	What was the unit of the lab result?	Unit
Laboratory Specimen ID	An internal or external identifier (e.g., specimen identifier).	What was the (laboratory test) [reference identifier/accession number]?	(Laboratory) [Reference identifier/Accession Number]
Laboratory Method of Test	Method of the test or examination.	What was the method used for the lab test or examination?	Method of Test or Examination
Lab Ref Range Lower Limit in Orig Unit	The lower end of normal range or reference range for continuous results stored in LBORRES.	What was the lower limit of the reference range for this lab test?	Normal Range Lower Limit
Lab Ref Range Upper Limit in Orig Unit	The upper end of normal range or reference range for continuous results stored in LBORRES.	What was the high limit of the reference range for this lab test?	Normal Range Upper Limit
Laboratory Reference Indicator	An indication or description of how the value compares to the normal range or reference range.	How [did/do] the reported values compare within the [reference/normal/expected] range?	Comparison to [Reference/Expected/Normal] Range

Table 9: CDISC Elements for the Domain of Microbiology (IPS Category for Diagnostic Results)

Data Element	DRAFT CDASHIG Definition	Question Text	Prompt
Microbiology Specimen Collection Date	The date of specimen collection, represented in an unambiguous date format (e.g., DD-MON-YYYY).	What was the (start) date of the (microbiology) specimen collection?	Specimen Collection (Start) Date
Microbiology Specimen Collection Time	The time of specimen collection, represented in an unambiguous time format (e.g., hh:mm:ss).	What was the (start) time of the (microbiology) specimen collection?	Specimen Collection (Start) Time
Microbiology Test	Descriptive name of the microbiology test or examination used to obtain the measurement or finding.	What was the microbiology examination test name?	[Microbiology Test Name]
Microbiology Test Result Value	Result of the measurement or finding as originally received or collected.	What was the result of the examination?	(Result)
Microbiology Test Result Value Units	The unit of the result as originally received or collected.	What was the unit of the result?	Unit
Microbiology Specimen Type	The type of specimen used for a measurement.	What is the specimen material type?	Specimen Type
Microbiology Test Method	Method of the test or examination.	What was the method used for the test or examination?	Method
Microbiology Reference ID	An internal or external identifier such as specimen identifier.	What was the (microbiology test) [reference identifier/accession number]?	(Microbiology Test) [Reference Identifier/Accession Number]
Microbiology Specimen Collection Location	A description of the anatomical location of the subject relevant to the collection of specimen.	What was the anatomical location where the specimen was collected?	Anatomical Location
Microbiology Examination Test Detail	Detail of the microbiology examination used to obtain the measurement or finding.	What was the microbiology examination detail?	[Examination Name Detail]

3.2.8 Social History Category or Substance Use

The Social History IPS Category or Substance Use Domain shown in Figure 26, Table 10)are primarily devoted to collecting information about use of tobacco or alcohol (generally, not abusive). However, the elements pertain to any type of substance one might be interested in tracking. Metadata items are listed such as dose, dose units, frequency and Start Date and End Date unless it is ongoing.

Table 10: CDISC Elements for the Domain of Substance Use (IPS Category Social History)

Data Element	DRAFT CDASHIG Definition	Question Text	Prompt
Substance Use Reported Name	The type of substance (e.g., TOBACCO, ALCOHOL, CAFFEINE or CIGARETTES, CIGARS, COFFEE).	What [is/was] the [name/type] of (the) substance used?	[Type of Substance]
Substance Usage	Indication the prespecified substance was used.	Has the subject ever [used/consumed] [SUTRT/SUCAT]?	([Substance]) Usage
Substance Dose Description	The amount of substance used (e.g., 1-2 packs, 8 oz).	What is/was the amount of [SUTRT] used/consumed?	Amount
Substance Dose (nontext)	The dose of substance (e.g., --TRT) taken at one time, represented as a numeric value.	What was the individual dose (of the concomitant [medication/treatment/therapy] per administration)?	[Dose/Amount] (per administration)
Substance Dose Units	The unit associated with the substance taken (e.g., pack in "1 pack per day").	What is the unit (for the dose of the substance])?	(Dose) Unit
Substance Use Frequency	The number/amount of the of substance consumed per a specific interval.	What [is/was] the frequency of [SUTRT] [use/consumption]?	Frequency
Substance Use Start Date	The date substance use started, represented in an unambiguous date format (e.g., DD-MON-YYYY.	What was the start date of [SUTRT/SUCAT] use/consumption?	Start Date
Substance Use End Date	The date substance use ended, represented in an unambiguous date format (e.g., DD-MON-YYYY).	What was the end date of [SUTRT/SUCAT] use/consumption?	End Date
Substance Use Duration	Collected duration of the substance use.	What was the duration of [SUTRT/SUCAT] use/consumption?	Duration
Substance Use Duration Unit	Unit of the collected duration of the substance use. Used only if duration was collected on the CRF.	What was the unit of duration of [SUTRT/SUCAT] use/consumption?	(Duration) Unit

Figure 26: IPS Social History Category Corresponds with the CDISC Domain for Substance Use

3.2.9 Admission or Discharge Information or Healthcare Encounters Domain

The information on Admission to the Hospital, Discharge from a Hospital or Visits to a Hospital can be quite important in research (Table 11, Figure 27). These elements are not included in the ISO or HL7 FHIR IPS; however, they are noted in the IPS document from IHE. This is a Gap: as we need Healthcare Encounters for Research.

Table 11: CDISC Elements for the Domain of Health Encounter

Data Element	DRAFT CDASHIG Definition	Question Text	Prompt
Healthcare Encounter	The reported or prespecified name of the healthcare encounter.	What was the healthcare encounter?; If [HODECOD], specify	[Healthcare Encounter]; [Specify]
Healthcare Encounter Start Date	The start date the healthcare encounter, represented in an unambiguous date format (e.g., DD-MON-YYYY).	What was the [healthcare encounter/HOTERM] [start/admission] date?	([HOTERM])[Start/Admission] Date
Healthcare Encounter End Date	The end date of the healthcare encounter (e.g., date of discharge), represented in an unambiguous date format (e.g., DD-MON-YYYY).	What was the [healthcare encounter/HOTERM] [end/discharge] date?	([HOTERM])[End/Discharge] Date
Reason for Healthcare Encounter	Denotes the reason for the healthcare encounter.	What was the reason for the [healthcare encounter/HOTERM]?	Reason for the Healthcare Encounter

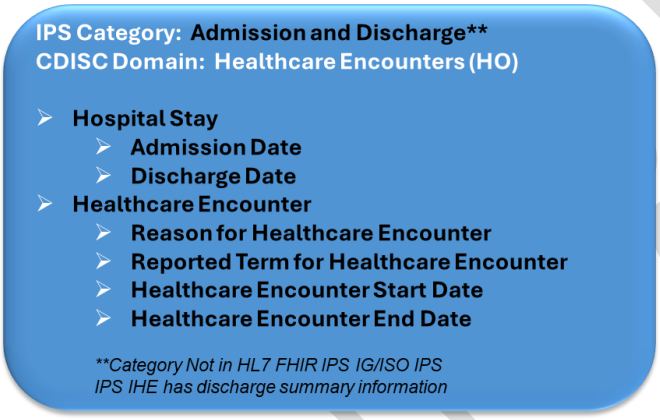

Figure 27: IPS Category of Admission and Discharge Corresponds with the CDISC Domain for Health Encounter (NOTE: The European IPS does not include health encounter information, in contrast to IHE IPS and ISO IPS.)

3.2.10 Adverse Events Domain

The information on Adverse Events (AE) is important in research (Table 12, Figure 28). The ICH defines an Adverse Event as "Any untoward medical occurrence in a patient or clinical investigation subject administered a pharmaceutical product and which does not necessarily have to have a causal relationship with this treatment". Metadata items are listed such as severity, serious, and outcomes are standard assessments collected for each new condition or event. This is a Gap: Adverse Event for Research.

Table 12: CDISC Elements for the Domain of Adverse Events

Data Element	DRAFT CDASHIG Definition	Question Text	Prompt
Adverse Event (AE)	The reported or prespecified name of the adverse event.	What is the adverse event term?	Adverse Event
AE Start Date	The start date of the adverse event, represented in an unambiguous date format (e.g., DD-MON-YYYY).	What is the adverse event start date?	Start Date
Adverse Event Start Time	The start time of the adverse event, represented in an unambiguous time format (e.g., hh:mm:ss).	What is the adverse event start time?	Start Time
Ongoing Adverse Event	Indication that an adverse event is ongoing when no end date is provided.	Is the adverse event ongoing (as of [the study-specific time point or period])?	Ongoing (as of [the study-specific time point or period])
AE End Date	The date when the adverse event resolved/ended, represented in an unambiguous date format (e.g., DD-MON-YYYY).	What was the adverse event end date?	End Date
End Time of AE	The time when the adverse event ended/resolved, represented in an unambiguous time format (e.g., hh:mm:ss).	What was the adverse event end time?	End Time
AE Severity	The severity or intensity of the event.	What is the severity of the adverse event?	Severity
AE Toxicity Grade	The grade of the severity of the event using a standard "toxicity" scale (e.g., NCI CTCAE).	What is the [NCI CTCAE/Name of scale (toxicity) grade] of the adverse event?	[NCI CTCAE/ Name of the scale] (Toxicity) Grade
AE Serious Event	An indication of whether the adverse event is determined to be "serious," based on what is defined in the regulations/protocol.	Was the adverse event serious?	Serious
AE Results in Death	An indication the serious adverse event resulted in death.	Did the adverse event result in death?	Death
AE is Life Threatening	An indication the serious adverse event was life threatening.	Was the adverse event life threatening?	Life Threatening
AE Requires or Prolongs Hospitalization	An indication the serious adverse event resulted in an initial or prolonged hospitalization.	Did the adverse event result in initial or prolonged hospitalization for the subject?	Hospitalization (initial or prolonged)
AE Persist or Signify Disability/Incapacity	An indication the serious adverse event was associated with a persistent or significant disability or incapacity.	Did the adverse event result in disability or permanent damage?	Disability or Permanent Damage
AE Congenital Anomaly or Birth Defect	An indication the serious adverse event was associated with a congenital anomaly or birth defect.	Was the adverse event associated with a congenital anomaly or birth defect?	Congenital Anomaly or Birth Defect
AE Needs Intervention to Prevent Impairment	An indication an adverse event required medical or surgical intervention to preclude permanent impairment of a body function, or prevent permanent damage to a body structure, due to the use of a medical product.	Did the adverse event require intervention to prevent permanent impairment or damage resulting from the use of a medical product?	Needs Intervention to Prevent Impairment
AE Other Medically Important Serious Event	An indication additional categories for seriousness apply.	Was the adverse event a medically important event not covered by other serious criteria?	Other Serious (Important Medical Events)
AE Involves Cancer	An indication the serious event was associated with the development of cancer.	Was the adverse event associated with the development of cancer?	Cancer
AE Causality	An indication the study treatment had a causal effect on the adverse event, as determined by the clinician/investigator.	Was this adverse event related to study treatment?	Relationship to Study Treatment
Actions Taken with Device	A description of the action taken, with respect to a device used in a study (which may or may not be the device under study), as a result of the event.	What action was taken with a device used in the study?	Action Taken with Device
Any Other Actions Taken	An indication whether any other actions were taken in response to the adverse event that were unrelated to study treatment dose changes or other non-study treatments given because of this adverse event.	Were any other actions taken in response to this adverse event?	Any Other Action(s) Taken
Outcome of AE	A description of the outcome of an event.	What is the outcome of this adverse event?	Outcome

Figure 28: IPS Category Proposed Data Element for Adverse Events for research Corresponds to the CDISC Domain for Adverse Events

3.3 Gap analysis

Reviewing the HL7 IPS IG against the xShare Core Data Element Set, the data elements researchers would find useful to add would support the surrounding concepts for observations in the categories listed in the following table (Table 13).

The elements identified through this gap analysis would be 'ideal' additions for research although further discussion should take place to determine the optimal way these could be supported in EHRs. For example, all of the adverse event elements may not need to be 'native required elements' should there be agreement to leverage the IHE RFD Interoperability specification with a serious adverse event form that presents in the user's screen in the event of a serious adverse event occurs.

In addition, discussions should occur in terms of the status of these elements in the IPS. For example, certain elements required for research are in the IPS under 'optional'. This is an initial proposed core data element set for IPS+R, which will need iterations and consensus-building to become a required standard for data collection. Such data elements are already required for research purposes globally.

Table 13: Gaps identified between HL7 FHIR IPS and xShare Core Element Set and proposed resolution

IPS Category	xShare Core Data Element	FHIR IPS	Prospect FHIR IPS Addition
Subject	Research Subject Identifier Research Study Identifier	Patient identifier is there – would require a connection between ResearchSubject/ ResearchStudy and Patient	Add rudimentary ResearchStudy, add ResearchSubject to point to IPS Patient
	Deceased Deceased Date/Time	Deceased dateTime (not Must Support) Deceased Boolean (not Must Support)	Add Must Support Flag
	Age		Add Age Observation
Medication Summary	Medication Route	Medication Route (not Must Support)	Add Must Support Flag
	Medication Ongoing	Medication with no end date indication is ongoing	No additions needed
Health Encounter	Healthcare Encounter Start Date		Add Encounter that fits the need for the required administrative items of an actual encounter
	Healthcare Encounter End Date		
	Healthcare Encounter Reason		
Adverse Event	Reported Term for the Adverse Event Adverse Event Start Date/Time Adverse Event End Date/Time Ongoing Adverse Event		Add AdverseEvent to reflect most of the needed items
	Severity/Intensity		Part of AdverseEvent
	Standard Toxicity Grade		Add Toxicity Observation
	Serious Event • Results in Death • Is Life Threatening • Requires or Prolongs Hospitalization • Persist or Significant Disability/Incapacity • Needs Intervention to Prevent Impairment • Other Medically Important Serious Event Involves Cancer		Part of these “serious” items are Coded Concepts of the outcome
	Causality		This is part of the SuspectedEntity→ Causality chain

4 Next steps

This working paper has presented a core set of elements that builds upon the IPS Categories and data elements for healthcare to also support clinical research and public health.

Once the initial core elements have been agreed by xShare partners and external experts and there have been appropriate reviews to build consensus, there will need to be an effort to include requirements for controlled terminology, codelists and/or valuelists to make it possible to exchange the key data while preserving the meaning (semantic interoperability) and ensuring quality. This will be undertaken within the time frame of xShare, as far as possible. The finally agreed inventory of data elements needed for research or public health that are not already required for the IPS will be proposed as IPS extensions.

A WP5 *European EHRxF in Clinical Research: Core Set, and IPS+R* working paper in the coming months will be a more detailed specification for these elements, including terminology, valuesets or codelists, as appropriate. The data element specifications will include the necessary metadata. There will be no new data modelling: the goal is that existing data models will adhere to and adopt the core data elements and associated metadata and terminology. It is anticipated that this core data element set will introduce approximately 10% new core data elements, in addition to the existing IPS elements. However, the addition of standard controlled terminology, codelists or valuesets (as appropriate for each element) will be necessary for all elements. In addition, many of the 'optional' elements in IPS should be required to meet the needs of the research and public health communities.

The coming use cases (D5.2) and other use cases from WP4 for public health will put these core data elements through testing that will no doubt identify necessary changes/modifications. Hence, this working paper will be considered an initial draft, and iterations will be anticipated to create the xShare standard core dataset during the course of this three-year project. This core set of data elements will also be tested in WP2 through the xBundle implementations for business use cases for healthcare and public health.

The European IPS (from the European Patient Summary Guideline V3.2 (June 2023), shown in Appendix II, is currently HL7 CDA. Mapping this to HL7 FHIR, while FHIR paths are associated with the xSHARE core data element set, would also be a relevant next step for this project.

This working paper (D5.1) has analysed data elements from a number of sources to identify and to some degree align the healthcare/IPS data elements of greatest research value and proposes the data elements that could usefully be added to the existing IPS in order to maximise its value for public/population health and clinical research in addition to its primary value for healthcare. Following stakeholder review, CDISC and HL7 FHIR implementation guides will be produced. This will set the groundwork for IPS+R and proposals for additional health information domains or priority data categories in the EHDS. Supplementary resources are available on the *xShare portal*.

5 Conclusion

This proposed core data element set, which is designed to support research and public health in addition to healthcare, should represent a global standard. Clinical research and outbreaks are frequently international; also, research sponsors wish to have their data accepted by regulatory authorities around the world. It would also be extremely valuable within healthcare so that there is a semantic standard to which EHRs and other mobile devices that collect healthcare and real-world data can aspire to meet for meaningful data exchange for the benefit of the patient. Ideally, this core element set will become a 'target' to which EHRs and other mobile devices collecting 'real world data' can map and/or harmonize the data they store.

Eventually, if xShare is successful, an EEHRxF and IPS with standard data elements and terminologies to support semantic interoperability will represent at least an EU standard, if not a global standard. This will no doubt improve data quality as vendors align with these requirements for at least a core set of data. At some point, it is hoped that, as with clinical research data standards, the core xShare data element set will set a global standard and will eventually be extended to support paediatrics as well as numerous therapeutic areas.

6 References to standards cited in this document

FDA. Sentinel Common Data Model. Accessed 24 May 2024 at https://www.sentinelinitiative.org/methods-data-tools/sentinel-common-data-model.

HL7. International Patient Summary Implementation Guide v1.1.0 – STU1 Update 1. Published 22 Nov 2022. Accessed on 26 April 2024 at http://hl7.org/fhir/uv/ips/.

HL7. International Patient Access Implementation Guide v1.0.0 – STU1. Published 26-March-2023. Accessed on 26 April 2024 at http://hl7.org/fhir/uv/ipa/

CDISC. Clinical Data Interchange Standards Consortium – CDASH, SDTM and ADAM. Accessed on 12 May 2024 at http://cdisc.org/standards.

Joint Initiative Council for Global Informatics Standardization. IPS Terminology Browser. Published 8 Dec 2022. Accessed on 26 April 2024 at https://international-patient-summary.net/ips-terminology-browser/

IHE. IHE IPS Profile. Published 3 Nov 2021. Accessed on 26 Apr 2024 at https://international-patient-summary.net/ips-terminology-browser/

eHealth Network. Guideline on the electronic exchange of health data under Cross-Border Directive 2011/24/EU. Published Nov 2023. Accessed on 26 Apr 2024 at https://health.ec.europa.eu/ehealth-digital-health-and-care/eu-cooperation/ehealth-network_en.

ISO. NEN-EN-ISO 27269:2022 International Standard: Health informatics-International patient summary. Published Apr 2021.

ISO. ISO/IEC 11179-1:2023 Information technology Metadata registries (MDR). Published Jan 2023. Accessed on 26 April 2024. at https://www.iso.org/standard/78914.html

NIH. National Cancer Institute Biomedical Research Integrated Domain Group (BRIDG) Model. 23 April 2023 at https://bridgmodel.nci.nih.gov/.

OHDSI. Observational Health Data Sciences and Informatics. Standardized Data: The OMOP Common Data Model. 2024. Accessed 24 May 2024 at https://www.ohdsi.org/data-standardization.

PCORnet. PCORnet Common Data Model. Accessed 24 May 2024 at https://pcornet.org/data/.

7 Appendix I: xShare Core Data Element Set v1.0

IPS Category	Domain	Data Element	USCDI v4.0	CDASHIG Variable	CDASHIG Variable Label	DRAFT CDASHIG Definition	Question Text	Prompt
Subject	DM	Study Identifier		STUDYID	Study Identifier	A unique identifier for a study.	What is the study identifier?	[Protocol/Study]
Subject	DM	Study Site Identifier		SITEID	Study Site Identifier	A unique identifier for a site within a study.	What is the site identifier?	Site (Identifier)
Subject	DM	Research Subject Identifier for the study	Member Identifier (Patient Identifier- would link to the research subject in blinded fashion at site level only known to investigator and team)	SUBJID	Subject Identifier for the Study	A unique subject identifier within a site and a study.	What [is/was] the (study) [subject/participant] identifier?	[Subject/Participant] (Identifier)
Subject	DM	Research Unique Subject Identifier		USUBJID	Unique Subject Identifier for the Study	Identifier used to uniquely identify a subject across all studies for all applications or submissions involving the product. This must be a unique value, and could be a compound identifier formed by concatenating STUDYID-SITEID-SUBJID.	What [is/was] the (StudyID)-(SiteID) (study) [subject/participant] identifier?	[STUDYID-SITEID-SUBJID]
Subject	DM	Research Study Identifier		STUDYID	Study Identifier	A unique identifier for a study.	What is the study identifier?	[Protocol/Study]
Subject	DM	Age		AGE	Age	The age of the subject, expressed in AGEU.	What is the subject's age?	Age
Subject	DM	Age Units		AGEU	Age Units	Units of time routinely used to express the age of a person.	What is the age unit used?	Age Unit
Subject	DM	Demographics Collection Date		DMDAT	Demographics Collection Date	The date of collection, represented in an unambiguous date format (e.g., DD-MON-YYYY).	What is the date of collection?	Collection Date
Subject	DM	Sex (Administrative/clinical use?)	Sex	SEX	Sex	Sex of the subject, as determined by the investigator.	What is the sex of the subject?	Sex

IPS Category	Domain	Data Element	USCDI v4.0	CDASHIG Variable	CDASHIG Variable Label	DRAFT CDASHIG Definition	Question Text	Prompt
Subject	DM	Birth Date	Date of Birth	BRTHDAT	Birth Date	A subject's date of birth (with or without the time of birth). The complete Date of Birth is made from the temporal components of Birth Year, Birth Month, Birth Day, and Birth Time..	What is the subject's date of birth?	Birth Date
Subject	DM	Deceased Date	Date of Death	DTHDTC	Date/Time of Death	Date/time of death for any subject who died, in ISO 8601 format. Should represent the date/time that is captured in the clinical-trial database.	What was the subject's date/time of death?	Death Date
Subject	DM	Deceased Flag		DTHFL	Subject Death Flag	Indicates the subject died. Should be "Y" or null. Should be populated even when the death date is unknown	Was the subject dead?	Subject Death Flag
Subject	SC	Subject Characteristic Collection Date		SCDAT	Subject Characteristic Collection Date	The date of collection represented in an unambiguous date format (e.g., DD-MON-YYYY).	What was the date the subject characteristics were collected?	Date
Subject	SC	Subject Characteristic	Data included for individual patient information; Physical Activity; Sexual Orientation; Gender Identity; Preferred Language; Occupation; Occupation Industry	SCTEST	Subject Characteristic	Descriptive name of the subject characteristic of interest.	What is the subject characteristics name?	[Subject Characteristic Test Name]
Subject	SC	Subject Characteristic Finding Value		SCORRES	SC Result or Finding in Original Units	Result of the subject characteristic as originally received or collected.	What is the subject characteristic?	(Result)
Subject	SC	Subject Characteristic Finding Value Units		[SCTESTCD]_ SCORRES	SC Result or Finding in Original Units	Result of the subject characteristics as originally received or collected.	What is the subject's [SCTEST]?	[SCTEST] Result
Problem List	MH	Medical History Collection Date		MHDAT	Medical History Collection Date	The date on which the medical history was collected, represented in an unambiguous date format (e.g., DD-MON-YYYY).	What was the date the medical history was collected?	Collection Date
Problem List	MH	Medical History Reported Term	Problem (Condition, diagnosis, or reason for	MHTERM	Reported Term for the Medical History	The reported or prespecified name of the medical condition or event.	What is the medical condition or event term?	Medical History Term

IPS Category	Domain	Data Element	USCDI v4.0	CDASHIG Variable	CDASHIG Variable Label	DRAFT CDASHIG Definition	Question Text	Prompt
			seeking medical attention)					
Problem List	MH	Medical History Event Start Date	Date of Diagnosis	MHSTDAT	Medical History Event Start Date	The start date of medical history event or condition, represented in an unambiguous date format (e.g., DD-MON-YYYY).	What [is/was] the [medical event or condition/category of the event] start date?	Start Date
Problem List	MH	Medical History Event End Date	Date of Resolution	MHENDAT	Medical History Event End Date	The end date of medical history event or condition, represented in an unambiguous date format (e.g., DD-MON-YYYY).	What[is/was] the[medical event or condition/category of the event] end date?	End Date
Medication Summary	CM	Medication	Medications	CMTRT	Reported Name of Drug, Med, or Therapy	Verbatim medication name or treatments (include only treatments with data collection characteristics similar to medications).	What was the (concomitant) [medication/treatment/therapy] (name/term)?	(Concomitant) [Medication/Treatment/Therapy]
Medication Summary	CM	Medication Dose	Dose	CMDOSE	CM Dose per Administration	The dose of medication/treatment (e.g., --TRT) given at one time, represented as a numeric value.	What was the individual dose (of the concomitant [medication/treatment/therapy] per administration)?	[Dose/Amount] (per administration)
Medication Summary	CM	Medication Dose Text (for range doses)		CMDSTXT	Concomitant Meds Dose Description	The dose of medication/treatment taken per administration.	What was the individual dose of the (concomitant) [medication/treatment/therapy]?	Dose
Medication Summary	CM	Medication Dose Unit	Dose Unit of Measure	CMDOSU	CM Dose Units	The unit associated with the concomitant medication/treatment/therapy taken (e.g., mg in "2 mg 3 times per day").	What is the unit (for the dose of concomitant [medication/treatment/therapy])?	(Dose) Unit
Medication Summary	CM	Medication Indication	Indication	CMINDC	CM Indication	The condition, disease, symptom, or disorder that the concomitant (non-study) medication/treatment/therapy was used to address or investigate (e.g., why the medication/treatment/therapy was taken or administered).	For what indication was the (concomitant) [medication/treatment/therapy] taken?	Indication

IPS Category	Domain	Data Element	USCDI v4.0	CDASHIG Variable	CDASHIG Variable Label	DRAFT CDASHIG Definition	Question Text	Prompt
Medication Summary	CM	Medication Dose Form		CMDOSFRM	CM Dose Form	The pharmaceutical dosage form in which the CMTRT is physically presented.	What was the dose form of the (concomitant) [medication/treatment/therapy]?	Dose Form
Medication Summary	CM	Medication Dose Frequency		CMDOSFRQ	CM Dosing Frequency per Interval	The number of doses given/administered/taken during a specific interval.	What was the frequency of the (concomitant) [medication/treatment/therapy]?	Frequency
Medication Summary	CM	Medication Route of Administration		CMROUTE	CM Route of Administration	The route of administration of the concomitant medication/treatment/therapy.	What was the route of administration of the (concomitant) [medication/treatment/therapy]?	Route
Medication Summary	CM	Medication Start Date		CMSTDAT	Concomitant Meds Start Date	The start date is when the concomitant medication/treatment/therapy was first taken, represented in an unambiguous date format (e.g., DD-MON-YYYY).	What was the (concomitant) [medication/treatment/therapy/dose] start date?	Start Date
Medication Summary	CM	Medication Ongoing		CMONGO	Ongoing Concomitant Meds	Indication the concomitant medication/treatment/therapy is ongoing when no end date is provided.	Was the (concomitant) [medication/treatment/therapy] ongoing (as of [the study-specific time point or period])?	Ongoing (as of [the study-specific time point or period])
Medication Summary	CM	Medication End Date		CMENDAT	Concomitant Meds End Date	The date that the subject ended/stopped taking the concomitant medication/treatment/therapy, represented in an unambiguous date format (e.g., DD-MON-YYYY).	What was the (concomitant) [medication/treatment/therapy/dose] end date?	End Date
History of Procedures	PR	Procedure	Procedures	PRTRT	Reported Name of Procedure	The verbatim surgical, therapeutic, or diagnostic procedure's name.	What was the procedure name?	[Procedure Name]; (Specify) Other
	PR	Procedure Start Date	Performance Time	PRSTDAT	Procedure Start Date	The date or start date of when the procedure started or was performed, represented in an unambiguous date format (e.g., DD-MON-YYYY).	What was the procedure (start) date ?	(Start) Date

IPS Category	Domain	Data Element	USCDI v4.0	CDASHIG Variable	CDASHIG Variable Label	DRAFT CDASHIG Definition	Question Text	Prompt
History of Procedures	PR	Procedure Indication		PRINDC	Procedure Indication	The condition, disease, symptom, or disorder that the procedure was used to address or investigate (e.g., why the therapy was taken or administered, why the procedure was performed).	For what indication was the [PRTRT] performed?	Indication
	PR	Procedure Ongoing		PRONGO	Ongoing Procedure	Indication the procedure is ongoing when no end date is provided.	Was the procedure ongoing (as of the [study-specific timepoint or period])?	Ongoing (as of the [study-specific timepoint or period])
History of Procedures	PR	Procedure End Date		PRENDAT	Procedure End Date	The end date of the procedure, represented in an unambiguous date format (e.g., DD-MON-YYYY).	What was the procedure (end) date?	(End) Date
Pregnancy Status, Hx & Summary	RP	Reproductive Finding Name	Pregnancy Status	RPTEST	Reproductive System Findings Test Name	Descriptive name for reproductive system finding.	What is the reproductive finding name?	[Reproductive System Findings Test Name]
Pregnancy Status, Hx & Summary	RP	Reproductive Finding Result Value		RPORRES	RP Result or Finding in Original Units	Result of the finding defined in reproductive system finding, as originally received or collected.	What was the result for the reproductive system question?	(Result)
Pregnancy Status, Hx & Summary	RP	Reproductive Finding Result Value Units		RPORRESU	RP Original Units	The unit of the result as originally received or collected.	What was the unit of the result?	Unit
Pregnancy Status, Hx & Summary	RP	Reproductive Finding Date		RPDAT	Reproductive System Finding Date	The date on which the reproductive system result or finding was collected, represented in an unambiguous date format (e.g., DD-MON-YYYY).	What was the date the reproductive system question was collected?	Collection Date
Vital Signs	VS	Vital Signs Test Name	Vital Signs: Systolic Blood Pressure; Diastolic Blood Pressure; Average Blood Pressure; Heart Rate; Respiratory Rate; Body Temperature; Body Height; Body Weight; Pulse Oximetry; Inhaled Oxygen Concentration	VSTEST	Vital Signs Test Name	Descriptive name of the test or examination used to obtain the measurement or finding.	What is the vital sign test name?	[Vital Signs Test Name]

IPS Category	Domain	Data Element	USCDI v4.0	CDASHIG Variable	CDASHIG Variable Label	DRAFT CDASHIG Definition	Question Text	Prompt
Vital Signs	VS	Vital Signs Test Date		[VSTESTCD]_VSDAT	Vital Signs Date	The date of the vital signs measurement, represented in an unambiguous date format (e.g., DD-MON-YYYY).	What was the date of the measurement(s)?	[VSTEST] Date
Vital Signs	VS	Vital Signs Test Time		[VSTESTCD]_VSTIM	Vital Signs Time	The time of measurement, represented in an unambiguous time format (e.g., hh:mm:ss).	What was the time of the measurement(s)?	[VSTEST] Time
Vital Signs	VS	Vital Signs Result Value		[VSTESTCD]_VSORRES	VS Result or Finding in Original Units	Result of the vital signs measurement as originally received or collected.	What was the result of the [VSTEST] measurement?	[VSTEST] (Result)
Vital Signs	VS	Vital Signs Result Value Units		[VSTESTCD]_VSORRESU	VS Original Units	The unit of the result as originally received or collected.	What was the unit of the [VSTEST] measurement?	[VSTEST] Unit
Diagnostic Results	MULTIPLE	Body System Diagnostic Test Date		--DAT	[Body System Diagnostic] Test Assessment Date	The date the [Body System Diagnostic] measurement was performed, represented in an unambiguous date format (e.g., DD-MON-YYYY).	What was the date the [Body System Diagnostic] measurement was taken?	Date
Diagnostic Results	MULTIPLE	Body System Diagnostic Test Name	Tests	--TEST	[Body System Diagnostic] Test Name	Descriptive name of the test or examination used to obtain the measurement or finding.	What is the [Body System Diagnostic] test name?	[[Body System Diagnostic] Test Name]
Diagnostic Results	MULTIPLE	Body System Diagnostic Test Result Value	Values/Results	--ORRES	[Body System Diagnostic] Test Result in Original Units	Result of the measurement or finding as originally received or collected.	What was the result of the measurement?	[[Body System Diagnostic] --TEST] Result
Diagnostic Results	MULTIPLE	Body System Diagnostic Test Result Value Unit	Result Unit of Measure	--ORRESU	[Body System Diagnostic] Test Original Unit	The unit of the result as originally received or collected.	What was the unit of the result?	Unit
Diagnostic Results	MULTIPLE	Body System Diagnostic Test Anatomical Location		--LOC	[Body System Diagnostic] Anatomical Location	Location used for the measurement.	What was the anatomical location where the measurement was taken?	Anatomical Location
Diagnostic Results	LB	Laboratory Specimen Collection Date		LBDAT	Specimen Collection Date	The date of specimen collection, represented in an unambiguous date format (e.g., DD-MON-YYYY).	What was the (start) date of the lab specimen collection?	Specimen Collection (Start) Date
Diagnostic Results	LB	Laboratory Specimen Collection Time		LBTIM	Specimen Collection Time	The time of specimen collection, represented in an unambiguous time format (e.g., hh:mm:ss).	What was the (start) time of the lab specimen collection?	Specimen Collection (Start) Time

IPS Category	Domain	Data Element	USCDI v4.0	CDASHIG Variable	CDASHIG Variable Label	DRAFT CDASHIG Definition	Question Text	Prompt
Diagnostic Results	LB	Laboratory Specimen Type	Specimen Type	LBSPEC	LB Specimen Type	The type of sample material taken from a biological entity.	What is the specimen material type?	Specimen Type
Diagnostic Results	LB	Laboratory Fasting Status		LBFAST	Lab Fasting Status	An indication that the subject has abstained from food/water for the specified amount of time.	Was the subject fasting?	Fasting
Diagnostic Results	LB	Laboratory Test Name	Laboratory Tests	LBTEST	Lab Test or Examination Name	Descriptive name of the lab test or examination used to obtain the measurement or finding. Any test normally performed by a clinical laboratory is considered a lab test.	What was the lab test name?	[Laboratory Test Name]
Diagnostic Results	LB	Laboratory Result Value	Values/Results	LBORRES	Lab Result or Finding in Original Units	Result of the measurement or finding as originally received or collected.	What was the result of the lab test?	(Result)
Diagnostic Results	LB	Laboratory Result Value Units	Result Unit of Measure	LBORRESU	Lab Original Units	The unit of the result as originally received or collected.	What was the unit of the lab result?	Unit
Diagnostic Results	LB	Laboratory Specimen ID	Specimen Identifier	LBREFID	Lab Specimen ID	An internal or external identifier (e.g., specimen identifier).	What was the (laboratory test) [reference identifier/accession number]?	(Laboratory) [Reference identifier/Accession Number]
Diagnostic Results	LB	Laboratory Method of Test		LBMETHOD	Lab Method of Test or Examination	Method of the test or examination.	What was the method used for the lab test or examination?	Method of Test or Examination
Diagnostic Results	LB	Lab Ref Range Lower Limit in Orig Unit	Result Reference Range	LBORNRLO	Lab Ref Range Lower Limit in Orig Unit	The lower end of normal range or reference range for continuous results stored in LBORRES.	What was the lower limit of the reference range for this lab test?	Normal Range Lower Limit
Diagnostic Results	LB	Lab Ref Range Upper Limit in Orig Unit	Result Reference Range	LBORNRHI	Lab Ref Range Upper Limit in Orig Unit	The upper end of normal range or reference range for continuous results stored in LBORRES.	What was the high limit of the reference range for this lab test?	Normal Range Upper Limit
Diagnostic Results	LB	Laboratory Reference Indicator	Result Interpretation	LBNRIND	Lab Reference Range Indicator	An indication or description of how the value compares to the normal range or reference range.	How [did/do] the reported values compare within the [reference/normal/expected] range?	Comparison to [Reference/Expected/Normal] Range

IPS Category	Domain	Data Element	USCDI v4.0	CDASHIG Variable	CDASHIG Variable Label	DRAFT CDASHIG Definition	Question Text	Prompt
Diagnostic Results	MB	Microbiology Specimen Collection Date		MBDAT	MB Specimen Collection Date	The date of specimen collection, represented in an unambiguous date format (e.g., DD-MON-YYYY).	What was the (start) date of the (microbiology) specimen collection?	Specimen Collection (Start) Date
Diagnostic Results	MB	Microbiology Specimen Collection Time		MBTIM	MB Specimen Collection Time	The time of specimen collection, represented in an unambiguous time format (e.g., hh:mm:ss).	What was the (start) time of the (microbiology) specimen collection?	Specimen Collection (Start) Time
Diagnostic Results	MB	Microbiology Test	Laboratory Tests	MBTEST	Microbiology Test or Finding Name	Descriptive name of the microbiology test or examination used to obtain the measurement or finding.	What was the microbiology examination test name?	[Microbiology Test Name]
Diagnostic Results	MB	Microbiology Test Result Value	Values/Results	MBORRES	MB Result or Finding in Original Units	Result of the measurement or finding as originally received or collected.	What was the result of the examination?	(Result)
Diagnostic Results	MB	Microbiology Test Result Value Units	Result Unit of Measure	MBORRESU	MB Original Units	The unit of the result as originally received or collected.	What was the unit of the result?	Unit
Diagnostic Results	MB	Microbiology Specimen Type	Specimen Type	MBSPEC	MB Specimen Type	The type of specimen used for a measurement.	What is the specimen material type?	Specimen Type
Diagnostic Results	MB	Microbiology Test Method		MBMETHOD	MB Method of Test or Examination	Method of the test or examination.	What was the method used for the test or examination?	Method
Diagnostic Results	MB	Microbiology Reference ID	Specimen Identifier	MBREFID	MB Reference ID	An internal or external identifier such as specimen identifier.	What was the (microbiology test) [reference identifier/accession number]?	(Microbiology Test) [Reference Identifier/Accession Number]
Diagnostic Results	MB	Microbiology Specimen Collection Location	Specimen Source Site	MBLOC	MB Specimen Collection Location	A description of the anatomical location of the subject relevant to the collection of specimen.	What was the anatomical location where the specimen was collected?	Anatomical Location
Diagnostic Results	MB	Microbiology Examination Test Detail		MBTSTDTL	Microbiology Examination Detail	Detail of the microbiology examination used to obtain the measurement or finding.	What was the microbiology examination detail?	[Examination Name Detail]
Social History	SU	Substance Use Reported Name	Alcohol Use; Substance Use; Smoking Status	SUTRT	Reported Name of Substance	The type of substance (e.g., TOBACCO, ALCOHOL, CAFFEINE or CIGARETTES, CIGARS, COFFEE).	What [is/was] the [name/type] of (the) substance used?	[Type of Substance]

IPS Category	Domain	Data Element	USCDI v4.0	CDASHIG Variable	CDASHIG Variable Label	DRAFT CDASHIG Definition	Question Text	Prompt
Social History	SU	Substance Usage	Alcohol Use; Substance Use; Smoking Status	SUNCF	Never Current Former Usage	Indication the prespecified substance was used.	Has the subject ever [used/consumed] [SUTRT/SUCAT]?	([Substance]) Usage
Social History	SU	Substance Dose Description		SUDSTXT	Substance Dose Description	The amount of substance used (e.g., 1-2 packs, 8 oz).	What is/was the amount of [SUTRT] used/consumed?	Amount
Social History	SU	Substance Dose (nontext)		SUDOSE	Substance Dose per Administration	The dose of substance (e.g., --TRT) taken at one time, represented as a numeric value.	What was the individual dose (of the concomitant [medication/treatment/therapy] per administration)?	[Dose/Amount] (per administration)
Social History	SU	Substance Dose Units		SUDOSU	Substance Dose Units	The unit associated with the substance taken (e.g., pack in "1 pack per day").	What is the unit (for the dose of the substance])?	(Dose) Unit
Social History	SU	Substance Use Frequency		SUDOSFRQ	Substance Use Frequency per Interval	The number/amount of the of substance consumed per a specific interval.	What [is/was] the frequency of [SUTRT] [use/consumption]?	Frequency
Social History	SU	Substance Use Start Date		SUSTDAT	Substance Use Start Date	The date substance use started, represented in an unambiguous date format (e.g., DD-MON-YYYY.	What was the start date of [SUTRT/SUCAT] use/consumption?	Start Date
Social History	SU	Substance Use End Date		SUENDAT	Substance Use End Date	The date substance use ended, represented in an unambiguous date format (e.g., DD-MON-YYYY).	What was the end date of [SUTRT/SUCAT] use/consumption?	End Date
Social History	SU	Substance Use Duration		SUCDUR	Substance Use Collected Duration	Collected duration of the substance use.	What was the duration of [SUTRT/SUCAT] use/consumption?	Duration
Social History	SU	Substance Use Duration Unit		SUCDURU	Substance Use Collected Duration Unit	Unit of the collected duration of the substance use. Used only if duration was collected on the CRF.	What was the unit of duration of [SUTRT/SUCAT] use/consumption?	(Duration) Unit
	HO	Healthcare Encounter	Encounter Information	HOTERM	Reported Term for Healthcare Encounter	The reported or prespecified name of the healthcare encounter.	What was the healthcare encounter?; If [HODECOD], specify	[Healthcare Encounter]; [Specify]
	HO	Healthcare Encounter Start Date	Encounter Time	HOSTDAT	Healthcare Encounter Start Date	The start date the healthcare encounter, represented in an unambiguous date format (e.g., DD-MON-YYYY).	What was the [healthcare encounter/HOTERM]	([HOTERM])[Start /Admission] Date

IPS Category	Domain	Data Element	USCDI v4.0	CDASHIG Variable	CDASHIG Variable Label	DRAFT CDASHIG Definition	Question Text	Prompt
							[start/admission] date?	
	HO	Healthcare Encounter End Date		HOENDAT	Healthcare Encounter End Date	The end date of the healthcare encounter (e.g., date of discharge), represented in an unambiguous date format (e.g., DD-MON-YYYY).	What was the [healthcare encounter/HOTERM] [end/discharge] date?	([HOTERM])[End/ Discharge] Date
	HO	Reason for Healthcare Encounter		HOREAS	Reason for the Healthcare Encounter	Denotes the reason for the healthcare encounter.	What was the reason for the [healthcare encounter/HOTERM] ?	Reason for the Healthcare Encounter
Allergies and Intolerances	AE	Adverse Event	Allergies and Intolerances	AETERM	Reported Term for the Adverse Event	The reported or prespecified name of the adverse event.	What is the adverse event term?	Adverse Event
Allergies and Intolerances	AE	Adverse Event Start Date		AESTDAT	Adverse Event Start Date	The start date of the adverse event, represented in an unambiguous date format (e.g., DD-MON-YYYY).	What is the adverse event start date?	Start Date
Allergies and Intolerances	AE	Adverse Event Start Time		AESTTIM	Start Time of Adverse Event	The start time of the adverse event, represented in an unambiguous time format (e.g., hh:mm:ss).	What is the adverse event start time?	Start Time
Allergies and Intolerances	AE	Ongoing Adverse Event		AEONGO	Ongoing Adverse Event	Indication that an adverse event is ongoing when no end date is provided.	Is the adverse event ongoing (as of [the study-specific time point or period])?	Ongoing (as of [the study-specific time point or period])
Allergies and Intolerances	AE	Adverse Event End Date		AEENDAT	Adverse Event End Date	The date when the adverse event resolved/ended, represented in an unambiguous date format (e.g., DD-MON-YYYY).	What was the adverse event end date?	End Date
Allergies and Intolerances	AE	End Time of Adverse Event		AEENTIM	End Time of Adverse Event	The time when the adverse event ended/resolved, represented in an unambiguous time format (e.g., hh:mm:ss).	What was the adverse event end time?	End Time
Allergies and Intolerances	AE	Adverse Event Severity		AESEV	AE Severity/Intensity	The severity or intensity of the event.	What is the severity of the adverse event?	Severity

IPS Category	Domain	Data Element	USCDI v4.0	CDASHIG Variable	CDASHIG Variable Label	DRAFT CDASHIG Definition	Question Text	Prompt
Allergies and Intolerances	AE	Adverse Event Toxicity Grade		AETOXGR	AE Standard Toxicity Grade	The grade of the severity of the event using a standard "toxicity" scale (e.g., NCI CTCAE).	What is the [NCI CTCAE/Name of scale (toxicity) grade] of the adverse event?	[NCI CTCAE/ Name of the scale] (Toxicity) Grade
Allergies and Intolerances	AE	Adverse Event Serious Event		AESER	AE Serious Event	An indication of whether the adverse event is determined to be "serious," based on what is defined in the regulations/protocol.	Was the adverse event serious?	Serious
Allergies and Intolerances	AE	Adverse Event Results in Death		AESDTH	Results in Death	An indication the serious adverse event resulted in death.	Did the adverse event result in death?	Death
Allergies and Intolerances	AE	Adverse Event is Life Threatening		AESLIFE	Is Life Threatening	An indication the serious adverse event was life threatening.	Was the adverse event life threatening?	Life Threatening
Allergies and Intolerances	AE	AE Requires or Prolongs Hospitalization		AESHOSP	Requires or Prolongs Hospitalization	An indication the serious adverse event resulted in an initial or prolonged hospitalization.	Did the adverse event result in initial or prolonged hospitalization for the subject?	Hospitalization (initial or prolonged)
Allergies and Intolerances	AE	AE Persist or Signif Disability/Incapacity		AESDISAB	Persist or Signif Disability/Incapacity	An indication the serious adverse event was associated with a persistent or significant disability or incapacity.	Did the adverse event result in disability or permanent damage?	Disability or Permanent Damage
Allergies and Intolerances	AE	AE Congenital Anomaly or Birth Defect		AESCONG	Congenital Anomaly or Birth Defect	An indication the serious adverse event was associated with a congenital anomaly or birth defect.	Was the adverse event associated with a congenital anomaly or birth defect?	Congenital Anomaly or Birth Defect
Allergies and Intolerances	AE	AE Needs Intervention to Prevent Impairment		AESINTV	Needs Intervention to Prevent Impairment	An indication an adverse event required medical or surgical intervention to preclude permanent impairment of a body function, or prevent permanent damage to a body structure, due to the use of a medical product.	Did the adverse event require intervention to prevent permanent impairment or damage resulting from the use of a medical product?	Needs Intervention to Prevent Impairment
Allergies and Intolerances	AE	AE Other Medically Important Serious Event		AESMIE	Other Medically Important Serious Event	An indication additional categories for seriousness apply.	Was the adverse event a medically important event not covered by other serious criteria?	Other Serious (Important Medical Events)

IPS Category	Domain	Data Element	USCDI v4.0	CDASHIG Variable	CDASHIG Variable Label	DRAFT CDASHIG Definition	Question Text	Prompt
Allergies and Intolerances	AE	AE Involves Cancer		AESCAN	Involves Cancer	An indication the serious event was associated with the development of cancer.	Was the adverse event associated with the development of cancer?	Cancer
Allergies and Intolerances	AE	AE Causality		AEREL	AE Causality	An indication the study treatment had a causal effect on the adverse event, as determined by the clinician/investigator.	Was this adverse event related to study treatment?	Relationship to Study Treatment
Allergies and Intolerances	AE	Actions Taken with Device		AEACNDEV	Actions Taken with Device	A description of the action taken, with respect to a device used in a study (which may or may not be the device under study), as a result of the event.	What action was taken with a device used in the study?	Action Taken with Device
Allergies and Intolerances	AE	Any Other Actions Taken		AEACNOYN	Any Other Actions Taken	An indication whether any other actions were taken in response to the adverse event that were unrelated to study treatment dose changes or other non-study treatments given because of this adverse event.	Were any other actions taken in response to this adverse event?	Any Other Action(s) Taken
Allergies and Intolerances	AE	Outcome of Adverse Event		AEOUT	Outcome of Adverse Event	A description of the outcome of an event.	What is the outcome of this adverse event?	Outcome

8 Appendix II: European Patient Summary Guideline V3.2 (June 2023)

A.1 Patient Summary Header Data Elements (v3.2, July 2023)			
ID	**Data Element**	**Description**	**Preferred Code Systems * ****
A.1.1	**Identification of the patient/subject**		
A1.1.1.1	National healthcare patient ID	Country ID, unique to the patient in that country. Example: ID for Portuguese patient	
A.1.1.2	Family name/surname	The given name/first name of the patient (also known as forename or first name). This field can contain more than one element.	
A.1.1.3	Given name	The family name/surname/last name of the patient. This field can contain more than one element or multiple fields could be present	
A.1.1.4	Date of birth	The date of birth of the patient [ISO TS 22220]. As age of the patient might be important for correct interpretation of the test result values, complete date of birth should be provided.	Complete date, without time, following the ISO 8601
A.1.1.5	Gender	This field must contain a recognised valid value for "administrative gender".If different, "physiological gender" should be communicated elsewhere.	HL7 Administrative Gender
A.1.1.6	Country of affiliation	Name of country of affiliation	ISO 3166
A.1.2	**Contact information**		
A.1.2.1	**Patient address**		
A.1.2.1.1	Street	Example: Rua dos Campeões	
A.1.2.1.2	House number	Example: 246	
A.1.2.1.3	City	Example: Porto	
A.1.2.1.4	Post code	Example: 4500	
A.1.2.1.5	State or province	Example: Vila Nova de Gaia	
A.1.2.1.6	Country	Example: PT	ISO 3166
A.1.2.1.7	Telephone no.	Example: +351 20 7025 6161	
A.1.2.1.8	Email	Example: jens@bigsmile.eu	
A.1.2.2	**Preferred HP to contact**	This section can be repeated and linked to any specific information in the Patient Summary, for example a link between a rare disease problem (A.2.3.1) and the rare disease specialist responsible for the care of the individual patient (this section).	
A.1.2.2.1	Name of the HP	Name of the Health Professional that has been treating or taking responsibility for the patient.	
A.1.2.2.2	Role of the HP	Health professional role	
A.1.2.2.3	HP Organisation	Health professional organisation	

A.1 Patient Summary Header Data Elements (v3.2, July 2023)			
ID	**Data Element**	**Description**	**Preferred Code Systems * ****
A.1.2.2.4	Telephone no.	Example: +45 20 7025 6161	
A.1.2.2.5	Email	Email of the HP/legal organisation	
A.1.2.2.6	Network affiliation	The HP organization is affiliated with a European network, for example an ERN.	
A.1.2.2.7	Related with	Identify the entry or entries of this Patient Summary for which the health professional is the preferred contact.	
A.1.2.3	**Contact person/ legal guardian**		
A.1.2.3.1	Role of that person	Role of the contact person: legal guardian, next of kin, other person to contact	HL7 RoleClass
A.1.2.3.2	Relationship level	Relationship type with the patient (e.g. father, wife, daughter)	HL7 RoleCode
A.1.2.3.3	Given name	The first name of the contact person/guardian (example: Peter). This field can contain more than one element.	
A.1.2.3.4	Family name/surname	This field can contain more than one element. Example: Español Smith	
A.1.2.3.5	Telephone no.	Example: +45 20 7025 6161	
A.1.2.3.6	Email	Email of the contact person/legal guardian	
A.1.3	**Insurance information**		
1.3.1	Insurance number	Example: QQ 12 34 56 A	
A.1.4	**Document data**		
A.1.4.1	Date created	Date on which PS was generated	ISO 8601
A.1.4.2	Date of last update	Date on which PS was updated (date of most recent version)	ISO 8601
A.1.4.3	Nature of the PS	Defines the context in which it was generated. Distinguishes between three methodological approaches for generating the PS: direct human intervention by an HP, automatically generated approach and mixed approach	
A.1.5	**Author and Organisation**		
A.1.5.1	Author organisation	At least one Author and Organisation shall be listed. In the event that there is no Author, at least one Organisation shall be listed. This Author should be able to provide further information on the provenance of the data present in the patient summary. Different authors contributing to individual sections and/or entries could be provided at the relevant level.	
A.1.5.2	Legal authenticator	Legal entity (Health Professional or Health Care Provider) who authenticated the Patient Summary	
A.1.6	**Additional information / Knowledge resources**		

A.1 Patient Summary Header Data Elements (v3.2, July 2023)			
ID	**Data Element**	**Description**	**Preferred Code Systems * ****
A.1.6.1	External reference	A reference leading to Clinical Practice Guidelines (CPG), emergency and anesthesia guidelines or other clinical relevant guidelines. This should be used only for providing specific guidance which may not be readily available to the health professional caring for the patient.	
A.1.6.2	Related with	Identify the entry or entries of this Patient Summary for which this additional information relates with, for example a link between a rare disease problem (section A.2.3.1) and guidelines for that particular rare disease (this section).	

A.1 Patient Summary Body Data Elements (v3.2, July 2023)			
ID	**Data Element**	**Description**	**Preferred Code Systems * ****
A.2.1	**Alerts**		
A.2.1.1	Allergy		
A.2.1.1 .1	Allergy description	Textual description of the allergy or intolerance	
A.2.1.1 .3	Type of propensity	This element describes whether this condition refers to an allergy, non-allergy intolerance, or unknown class of intolerance (not known to be allergy or intolerance)	SNOMED CT GPS
A.2.1.1 .4	Allergy manifestation	Description of the clinical manifestation of the allergic reaction. Example: anaphylactic shock, angioedema (the clinical manifestation also gives information about the severity of the observed reaction)	SNOMED CT GPS
A.2.1.1 .4	Severity	Severity of the clinical manifestation of the allergic reaction.	SNOMED CT GPS
A.2.1.1 .5	Criticality	Potential risk for future life-threatening adverse reactions when exposed to a substance known to cause an adverse reaction.	SNOMED CT GPS
A.2.1.1 .6	Onset date	Date of the observation of the reaction	ISO 8601
A.2.1.1 .7	End Date	Date of resolution of the allergy (e.g. when the clinician deemed there is no longer any need to track the underlying condition)	ISO 8601
A.2.1.1 .8	Status	Current status of the allergy or intolerance, for example, whether it is active, in remission, resolved, etc.	SNOMED CT GPS
A.2.1.1 .9	Certainty	Assertion about the certainty associated with a propensity, or potential risk, of a reaction to the identified substance. Diagnostic and/or clinical evidence of condition.	SNOMED CT GPS

A.1 Patient Summary Body Data Elements (v3.2, July 2023)			
ID	**Data Element**	**Description**	**Preferred Code Systems * ****
A.2.1.1 .10	Agent or Allergen	A specific allergen or other agent/substance (drug, food, chemical agent, etc.) to which the patient has an adverse reaction propensity.	SNOMED CT GPS(for non-drug allergy) ATC* (for drug allergy)(IDMP, when available)
A.2.1.2	Medical alert information (other alerts not included in allergies)		
A.2.1.2 .1	Healthcare alert description	Description of medical alerts in textual format: any clinical information that is imperative to know so that the life or health of the patient does not come under threat. - Example 1: intolerance to aspirin due to gastrointestinal bleeding. Example 2: intolerance to captopril because of cough (the patient is not allergic but can't tolerate it because of persistent cough) Example 3: the patient has a rare disease that requires special treatment Example 4: Airway Alert / Difficult Intubation Example 5: Diagnoses such as malignant hyperthermia, porphyria, and bleeding disorders; special treatments like anticoagulants or immunosuppressants; implanted devices. Example 6: transplanted organs illustrate other information that has to be taken into account in a healthcare contact. Example 7: participation in a clinical trial that has to be taken into account in a healthcare contact.	
A.2.2	**Medical history**		
A.2.2.1	**Vaccination/ prophylaxis information**		
A.2.2.1 .1	Disease or agent targeted	Disease or agent that the vaccination provides protection against	ICD-10* SNOMED CT GPS
A.2.2.1 .2	Vaccine/prophyl axis	Generic description of the vaccine/prophylaxis or its component(s)	SNOMED CT GPS, ATC*, (IDMP, when available)
A.2.2.1 .3	Vaccine medicinal product name	Brand name of the vaccine medicinal product.z	
A.2.2.1 .3.1	Identifier of the vaccine medicinal product	Identifier for the vaccine medicinal product. It could be MPID according to ISO 11615, EMA PMS ID and/or a national identifier.	EMA PMS
A.2.2.1 .4	Marketing Autorisation Holder	Marketing Authorisation Holder	EMA's Organisations System data (SPOR)

A.1 Patient Summary Body Data Elements (v3.2, July 2023)			
ID	**Data Element**	**Description**	**Preferred Code Systems * ** **
A.2.2.1.5	Number in a series of vaccinations/doses	Order in the vaccination course	
A.2.2.1.6	Batch/lot number	A distinctive combination of numbers and/or letters which specifically identifies a batch	
A.2.2.1.7	Date of vaccination	The date when the vaccination was administered	ISO 8601
A.2.2.1.8	Administering centre	Name/code of administering centre or a health authority responsible for the vaccination event	
A.2.2.1.9	Health Professional identification	Name or health professional code responsible for administering the vaccine or prophylaxis	
A.2.2.1.10	Country of vaccination	The country in which the individual has been vaccinated	ISO 3166
A.2.2.1.11	Next vaccination date	The date when the vaccination is planned to be given/repeated (e.g. next dose)	ISO 8601
A.2.2.2	**Resolved, closed or inactive problems**		
A.2.2.2.1	Problem description	Problems or diagnoses that the patient suffered in the past, and which have been resolved, closed or declared as inactive (not included in "current problems or diagnosis"). Example: hepatic cyst (the patient has been treated with a hepatic cystectomy that solved the problem and the problem is therefore closed).	ICD-10* or SNOMED CT GPS Orphacode if rare disease is diagnosed
A.2.2.2.2	Onset date	Date of problem onset	ISO 8601
A.2.2.2.3	End date	Problem resolution date	ISO 8601
A.2.2.2.4	Resolution circumstances	Describes the reason for which the status of the problem changed from current to inactive (e.g. surgical procedure, medical treatment, etc.). This field includes "free text" if the resolution circumstances are not already included in other fields such as surgical procedure, medical device, etc., e.g. hepatic cystectomy (this will be the resolution circumstances for the problem "hepatic cyst" and will be included in surgical procedures).	
A.2.2.3	**Medical history**		
A.2.2.3.1	Medical history	This section may provide both synthetic anamnesis (e.g. description of phases of the pathology as a chronological summary of clustered clinical information) and anecdotal evidence that clinicians can collect from the patient, and can read in narrative. S2,ChII,Art7(c)	
A.2.3	**Medical problems**		

A.1 Patient Summary Body Data Elements (v3.2, July 2023)			
ID	**Data Element**	**Description**	**Preferred Code Systems * ****
A.2.3.1	**Current problems**		
A.2.3.1 .1	Problem / diagnosis description	Health conditions affecting the health of the patient and are important to be known for a health professional during a health encounter.	ICD-10* SNOMED CT GPS Orphacode if rare disease is diagnosed
A.2.3.1 .2	Onset date	Date of problem onset	ISO 8601
A.2.3.1 .3	Diagnosis assertion status	Assertion about the certainty associated with a diagnosis. Diagnostic and/or clinical evidence of condition.	HL7
A.2.3.2	**Medical devices and implants**		
A.2.3.2 .1	Device and implant description	Describes the patient's implanted and external medical devices and equipment upon which their health status depends. Includes devices such as cardiac pacemakers, implantable fibrillator, prosthesis, ferromagnetic bone implants, etc. of which the HP needs to be aware.	SNOMED CT GPS* EMDN
A.2.3.2 .2	Device ID	Normalised identifier of the device instance such as UDI according to REGULATION (EU) 2017/745	
A.2.3.2 .3	Implant date	Date when procedure was performed	ISO 8601
A.2.3.2 .4	End date	Date when the device was explanted from the patient or the external device was no longer in use; likewise when the device is planned to be explanted	ISO 8601
A.2.3.3	**Procedures**		
A.2.3.3 .1	Procedure description	Describes the type of procedure	SNOMED CT GPS*
A.2.3.3 .2	Body site	Procedure target body site	SNOMED CT GPS*
A.2.3.3 .3	Procedure date	Date when procedure was performed	ISO 8601
A.2.3.4	**Functional status**		
A.2.3.4 .1	Description	Need for the patient to be continuously assessed by third parties; functional status may influence decisions about how to plan and administer treatments	
A.2.3.4 .2	Onset Date	Onset date of a condition	ISO 8601
A.2.3.4 .3	Functional assessment description	Description of the functional assessment	ICF

A.1 Patient Summary Body Data Elements (v3.2, July 2023)			
ID	**Data Element**	**Description**	**Preferred Code Systems * ** **
A.2.3.4.4	Functional assessment date	Date of the functional assessment	ISO 8601
A.2.3.4.5	Functional assessment result	Functional assessment result value	ICF
A.2.4	**Medication summary**		
A.2.4.1	Current and relevant past medicines	Relevant prescribed medicines whose period of time indicated for the treatment has not yet expired whether it has been dispensed or not, or medicines that influence current health status or are relevant to a clinical decision	
A.2.4.1.1	Medication reason	This is the reason why the medication is being prescribed or used. It provides a link to the Past or current health conditions or problems that the patient has had or has.	ICD-10* or SNOMED CT GPS Orphacode if rare disease is diagnosed
A.2.4.1.2	Intended use	Indication intended use as: prevention or treatment Example: prophylaxis, treatment, diagnostic, anaesthesia, care of equipment,	
A.2.4.1.3	Brand name	Brand name if biological medicinal product or when justified by the health professional (ref. Commission Directive 2012/52/EU)	
A.2.4.1.4	Active ingredient lists	Substance that alone or in combination with one or more other ingredients produces the intended activity of a medicinal product. Example: "paracetamol"	ATC* (IDMP identifier, when available)
A.2.4.1.5	Strength	The content of the active ingredient expressed quantifiably per dosage unit, per unit of volume or per unit of weight, according to the pharmaceutical dose form. Example: 500 mg per tablet	UCUM, EDQM Standard Terms
A.2.4.1.6	Pharmaceutical dose form	The form in which a pharmaceutical product is presented in the medicinal product package (e.g. tablet, syrup)	EDQM Standard Terms
A.2.4.1.7	Dosage Regimen	Number of units per intake and frequency of intake over a specified duration of time. Example: 1 tablet every 24h, for 10 days	
A.2.4.1.8	Route of administration	Path by which the pharmaceutical product is taken into or makes contact with the body.	EDQM Standard Terms
A.2.4.1.9	Date of onset of treatment	Date when patient needs to start taking the medicine prescribed	ISO 8601
A.2.5	**Social history**		
A.2.5.1	Social history observations related to health	Health related lifestyle factors or lifestyle observations and social determinants of health. Example: cigarette smoker, alcohol consumption	SNOMED CT GPS
A.2.5.2	Reference date range	Example: from 1974 to 2004	

A.1 Patient Summary Body Data Elements (v3.2, July 2023)			
ID	**Data Element**	**Description**	**Preferred Code Systems * ****
A.2.6	**Pregnancy history**		
A.2.6.1	**Current pregnancy status**		
A.2.6.1.1	Date of observation	Date on which the observation of the pregnancy state was made	ISO 8601
A.2.6.1.2	Status	Provides the woman's current state at the date the observation was made: e.g. pregnant, not pregnant, unknown	SNOMED CT GPS
A.2.6.1.3	Expected date of delivery	Date on which the woman is due to give birth.	ISO 8601
A.2.6.2	**History of previous pregnancies**		
A.2.6.2.1	Previous pregnancies status	Information on the woman's previous pregnancies: Yes, previous pregnancies; No, previous pregnancies; Unknown	SNOMED CT GPS
A.2.6.2.2	Previous pregnancies description		
A.2.6.2.2.1	Outcome date	Date referred to the previous pregnancies outcome	ISO 8601
A.2.6.2.2.2	Outcome	Outcome of the previous pregnancies	SNOMED CT GPS
A.2.6.2.2.3	Number of children	Number of children/fetus in this specific pregnancy	
A.2.7	**Patient provided data**		
A.2.7.1	**Travel history**		
A.2.7.1.1	Country	Country(s) visited	ISO 3166
A.2.7.1.2	Period	Date of entry and departure	ISO 8601
A.2.7.2	**Advance Directive**		
A.2.7.2.1	Documentation	Existence of documentation on living will	SNOMED CT GPS
A.2.8	**Results**		

A.1 Patient Summary Body Data Elements (v3.2, July 2023)			
ID	**Data Element**	**Description**	**Preferred Code Systems * ****
A.2.8.1	**Result observations**	(A list of observation results pertaining to the subject of care's health condition and which might have impact on future treatments)	
A.2.8.1 .1	Date	Date and time of the observation	ISO 8601
A.2.8.1 .2	Observation type	Observation results types that may be measurements, laboratory results, anatomic pathology results, radiology results or other imaging or clinical results. Examples: Diagnostic results (Blood group, Laboratory Observations, Imaging results etc.) Physical findings (Vital signs observations)	HL7 Observation CategoryCodes
A.2.8.1 .3	Result description	Narrative representation of the observation result and findings.	
A.2.8.1 .4	Observation details	Observation details including code that identifies observation, specification of the observed body structure or specimen, date and time of the specimen collection.	LOINC SNOMED CT GPS NPU
A.2.8.1 .5	Observation results	Result of the observation including numeric and coded results of the measurement, details about how the tests were done to get the result values, information about referential ranges and result interpretation. Content of the observation result will vary according to the type of the observation.	SNOMED CT GPS (for ordinal or nominal scale results) UCUM (for units)
A.2.8.1 .6	Performer	Identifies the originator/author and provides provenance information about the source of the results data that may have not originated with the source of the whole PS document.	
A.2.8.1 .7	Reporter	With certain observation results, e.g. there may also be an interpreter or a person responsible for validation.	
A.2.9	**Plan of Care**	**(Therapeutic recommendations that do not include pharmacologic treatments, such as diet, physical exercise, planned surgeries)**	
A.2.9.1	Plan of care	Narrative containing the plan including proposals, goals, and order requests for monitoring, tracking, or improving the condition of the patient. In the future it is expected that this section could be provided in a structured and coded format.	

* In a foreseeable future, the suggested preferred vocabularies might be superseded or complemented, as mentioned in Guidelines Article 11(2).
** The Preferred code system(s) has been selected based on adequacy to convey the information using the methodology of the Subgroup on Semantics. When more alternative international code systems are available, all are listed when it is assumed to be unlikely that agreement can be reached short term. Mapping between code systems could be proposed for specific use cases.

9 Appendix III: United States Common Data Core (USCDI v4)