From jargon to clarity

On patient access to electronic health records and patient-friendly terminology

van Mens, H.J.T.

Publication date
2023

Document Version
Final published version

From jargon to clarity

On patient access to electronic health records and patient-friendly terminology

Hugo J.T. van Mens

From jargon to clarity

On patient access to electronic health records and patient-friendly terminology

Hugo J.T. van Mens

Colophon

Doctoral thesis

From jargon to clarity: On patient access to electronic health records and patient-friendly terminology

Cover and layout design by Timo Wolf Kamp | Persoonlijk Proefschrift

Printed by Ridderprint, the Netherlands

This thesis was prepared at ChipSoft B.V. and Amsterdam UMC. The author carried out the research and development for this thesis at ChipSoft B.V., Amsterdam, the Netherlands, within the Department of Research and Development, Care Portal. Additionally, the research was carried out at Amsterdam UMC, University of Amsterdam, at the Department of Medical Informatics, Meibergdreef 9, Amsterdam, the Netherlands, within the Amsterdam Public Health research institute, research programs Quality of Care, Methodology and Digital Health, Amsterdam, the Netherlands.

The research for/publication of this doctoral thesis received financial assistance from ChipSoft B.V., Amsterdam UMC and UMC Utrecht, which is gratefully acknowledged.

ISBN 978-94-6483-108-5

NUR 970

From jargon to clarity

On patient access to electronic health records and patient-friendly terminology

ACADEMISCH PROEFSCHRIFT

ter verkrijging van de graad van doctor

aan de Universiteit van Amsterdam

op gezag van de Rector Magnificus

prof. dr. ir. P.P.C.C. Verbeek

ten overstaan van een door het College voor Promoties ingestelde commissie,

in het openbaar te verdedigen in de Agnietenkapel

op dinsdag 13 juni 2023, te 15.00 uur

door Hugo Johan Theodoor van Mens

geboren te Naarden

Promotiecommissie

Promotores:	prof. dr. N.F. de Keizer	AMC-UvA
	dr. ir. R. Cornet	AMC-UvA
Overige leden:	prof. dr. M.W.M. Jaspers	AMC-UvA
	prof. dr. E.M.A. Smets	AMC-UvA
	prof. dr. J.C.M. van Weert	Universiteit van Amsterdam
	dr. C.W.M. Dedding	Vumc
	prof. dr. A.P.J. van den Bosch	Universiteit Utrecht
	dr. F. Hielkema-Raadsveld	Nictiz

Faculteit der Geneeskunde

Contents


Anamnesis
He has been ve ll and suffered from dyspnea and
lethargia. Hist roenteritis due to SARS-CoV-2.
Physical
Atrial simulatio quency 73/minute. QRS and QT/QTc
normal. liac decompensation. Complaints
arose dmill t .

1

General introduction

1.1 What cows, abduction, denial and cabbage have in common

Have you ever seen nurses walking around with a cow when you visited a hospital? Did they really make a scan of your cat and abduct your arm? Why did your doctors write down that you deny drinking and tell you they will give you a cabbage? These are all examples of medical jargon used in clinical practice. Patients frequently find it difficult to understand medical terminology.[1-4] Medical language can be confusing, awkward or even offending.[5,6] Outside veterinary medicine, cows are common in clinical practice as well, as COW is an abbreviation for a Computer On Wheels.[7,8] A CAT scan refers to a scan made by Computerized Axial Tomography.[9] Aliens were not involved in your arm abduction, but it is a medical term for moving away a body part from another, e.g. moving away your arm from your trunk.[10] Cabbage, in the example above, does not refer to a vegetable crop. Clinicians pronounce the abbreviation CABG as cabbage, which refers to Coronary Artery Bypass Graft.[7] Clinicians use formal, objective language that may alienate patients. Positive has a different connotation in plain language than in the context of test results. Clinicians might write that a patient denied something, which does not mean they do not trust what the patient was saying, but that the patient confirmed something was not the case (e.g. that the patient stated he or she does not drink alcohol).[5,6] Even among clinicians confusion is not uncommon. Medical specialties each have their peculiar terms. Paget's disease may refer to a carcinoma (Paget's disease of the skin) or a bone disorder (Paget's disease of bone or osteitis deformans).[9] The same abbreviations may mean different things, even within the same medical specialty. In cardiology for instance, where MI may refer to myocardial infarction or mitral insufficiency.[9] Confusion on the meaning of communicated terms between clinicians and patients, and among clinicians, might lead to adverse events and affect patient safety.[11-13]

1.2 History of medical record keeping

Medical terms have a long history and the practice of medical record-keeping dates back to ancient civilizations. Oven-baked clay records have been preserved from Mesopotamia (2114 – 2004 BC).[14] Medical case histories recorded on papyrus were found in Egypt (1600 – 1700 BC). However, these ancient records were considered to be textbook case studies rather than medical records as we understand them today.[15-17] Hippocrates (Greece, 460 – 370 BC), well-known for the Hippocratic oath that medical doctors still take nowadays, recorded medical notes about his patients and recommended this practice.[17-19] Using paper to record patient data became more common practice in the 19th and 20th centuries. When hospitals began to become larger and more important, medical records started to contain contributions from different disciplines, e.g. medical specialists, nurses, and psychologists. Different formats of paper from various clinician authors were bound or put into folders containing the medical data from single patients.[17] It was difficult and costly to store, oversee, exchange and reuse paper records. Moreover, medical doctors have not been well-known for their good, legible handwriting. Medical errors may result from illegible handwriting and copying.[20,21] When computers became more usable and powerful, medical records were digitized into electronic health records (EHRs) that were captured in health information systems (HISs). HISs are used in hospitals, but also by general practitioners (GPs; in GP information systems) and pharmacies (pharmacy information systems) for instance. EHRs resolved some of the issues encountered with paper records.

1.3 Free text and data encoding in EHRs

Most of the data in EHRs consist of free text, such as progress notes, test results and referral letters. Free text may contain typos, abbreviations, acronyms and homonyms; notes may lack proper grammar, and might only contain some keywords, without properly phrased full sentences.[22] Therefore, it is difficult to understand and reuse free-text data for other purposes[23]; not only for patients to understand. Data recorded in EHRs, however, are reused by various stakeholders, such as other clinicians that treat the same patient, administrators for financial reimbursement, medical coders for statistics, medical registries for quality improvement and scientists for medical research. To facilitate reuse, data need to be recorded in an unambiguous, formalized and structured form. Various standardized terminology systems are used in clinical practice to encode medical data.[24] The same data, e.g. the diagnosis, of one patient may have to be registered

several times, however, in different formats: in free text and by different codes and descriptions from different terminology systems, for different purposes. Data are encoded for statistics in ICD codes (International Classification of Diseases), to monitor the mortality and prevalence of diseases, enable comparisons among countries and inform policy-makers.[25] Medical coders read full free-text notes and data from medical records to extract the reason for hospitalization and use classification rules to determine the applicable ICD codes.[26] In general practice, however, the International Classification of Primary Care (ICPC) is used.[27] For financial reimbursement, diagnoses and procedures are encoded and classified with diagnosis-related groups (DRGs).[28,29] In the Netherlands, a peculiar DRG-like system called DBC (an abbreviation for the Dutch word "diagnosebehandelcombinatie", or "diagnosis treatment combination" in English) was invented for reimbursement.[30] These either are encoded by the financial administration department or physicians are required to encode a DBC to obtain reimbursement.[31] For quality improvement and research, several medical registries exist that might have domain-specific coding systems. In intensive care, for instance, the reason for the encounter is encoded in the APACHE-IV (Acute Physiology and Chronic Health Evaluation IV) classification.[32,33] Medical literature is indexed with Medical Subject Headings (MeSH) for information retrieval.[34]

Each of these terminology systems refers to the same concepts but with different codes, descriptions and levels of detail, as illustrated in Table 1.1. Clinicians have to search for descriptions in these predefined lists. They may experience this as an administrative burden as coding might require more time than writing free-text notes. Clinicians already have to spend much time and effort on administrative tasks and documenting patient histories, treatments and progress.[35,36] Additionally, codes might limit their expressivity or be too specific.[37] Notes are most often still documented as free text. When clinicians are provided with the opportunity, they may avoid having to encode and structure their data, especially when they are under pressure.[38] Additionally, remarks may be added to encoded data or descriptions may even be modified, potentially changing the meaning of the encoded data and making it more difficult and less reliable to reuse the data if these changes are not detected.[38,39] For example, adding the free-text remark 'suspected' to a diagnosis code 'glaucoma', implies that it was not confirmed but suspected, while the underlying encoding does not contain this information about certainty and may imply that it was a confirmed diagnosis instead.[39] Therefore, in practice, free text is still abundant in medical records.

Table 1.1 Different encodings of a brain contusion of different terminology systems, with their domain, concept id and description

Terminology system	Domain	Concept id	Description
Diagnosethesaurus[40]	Dutch clinical practice	0000011788	contusio cerebri
NHG ICPC 1[41]	Dutch primary care practice	N80.04	Contusio cerebri
ICPC 3[42]	International primary care practice	ND36.00	Cerebral contusion
SNOMED CT[9]	Reference terminology	262689001	Contusion of cerebrum (disorder)
ICD-10[43]	International statistics	S06.20	Diffuse brain injury, without open intracranial wound
DBC[40]	Dutch financial reimbursement	273	Moderately severe cranial brain injury (contusio)
APACHE IV[33]	ICU registries	208 (NICE)	Head (CNS) only trauma
MeSH[34]	Medical literature	D000070624	Brain Contusion

1.4 Single entry, multiple use

There have been various efforts to structure health records and make clinicians use coding systems at the point of care by which other codes, for other purposes, can be derived. This single-entry, multiple-use paradigm aims to prevent redundant registration.[44] The Diagnosethesaurus (Dutch for “Diagnosis Thesaurus”) is an interface terminology developed in the Netherlands to help clinicians code their diagnoses once (similarly there is the Verrichtingenthesaurus, or “Procedure Thesaurus”, for procedures).[45] These codes can be reused via classification rules and the reference terminology SNOMED CT. The Diagnosethesaurus also maps to the DBC and ICD-10 aggregate terminologies, so administrative and statistical codes can be derived. SNOMED CT contains synonyms and identifiers and maps to other coding systems.[46] SNOMED CT concepts are modelled with description logic which means that logical relationships formally define the meaning (semantics) of medical concepts. For instance, pancreatitis is defined as a disorder with the associated morphology of inflammation and finding site pancreas. The logical representation enables reuse for decision support systems, comprehensive searching and data analytics.[47-49] For example, SNOMED CT can be used to retrieve all disorders that involve inflammations of the gastrointestinal tract caused by a virus or to search for all patients with pancreatic disorders, by using these relationships. The logical representation also enables automatic classification of head traumas, such as a cerebral contusion, as head trauma in the APACHE-IV reasons for encounter.[50,51]

1.5 Patient access to EHRs enabled by patient portals and personal health records

Patients have the right to view, copy, modify and delete their records or parts of them according to the General Data Protection Regulation and the Dutch law on the medical treatment agreement ("WGBO" for "wet op de geneeskundige behandelovereenkomst" in Dutch).[52,53] In some healthcare systems, it is common practice that patients take their health records home and bring them along to other providers.[54] Reading their records can help patients remember what had been discussed and learn more about their conditions and treatments. This might have a positive effect on their self-management.[55-58] Patients can access their EHRs through patient portals, which are subsystems of HISs. Patients are thus also end-users of HISs and secondary users of medical data. Patient portals have various other features as well, such as appointment scheduling, questionnaire answering, secure messaging, patient education and prescription renewal. In countries such as the USA and the Netherlands, patient portals are a common functionality offered by healthcare providers.[59] Alternatively, patients can collect health data in applications they employ themselves, called PHRs.[60] Patients can import data from HISs of their healthcare providers but also collect data they generated themselves, e.g. by their wearable devices or manually recorded in their dietary diary.[61,62] This enables them to collect all health data in one place and gain an overview of the scattered health data collected by various healthcare institutions, health apps and smart health devices. However, the adoption of PHRs by patients and hospitals is limited.[59,61]

1.6 Patient-provider communication and patient-friendly terminology

Patients are informed on their health condition in the first place by their healthcare provider during consultations. Healthcare providers may offer patient information leaflets with further information. On the other hand, patients increasingly use the internet to find health information and are not passive consumers anymore that follow everything their clinicians tell them.[63,64] However, there is also a large group of patients that is less able to process and use medical data to the benefit of their own health, a skill known as health literacy.[65] About one-third of Europeans have low health literacy.[66] Low health literacy is associated with poorer health outcomes[67] and is also a barrier to patient portal and PHR adoption.[55,68,69] Medical language is formal and might contain a mix of various contemporary and ancient languages (e.g. Latin, Greek, English and Dutch terms in the Netherlands),

acronyms and abbreviations.[22] Patients prefer clinical notes to be written in more accessible language, but clinicians are reluctant to change their writing styles and need specific medical language for medical record keeping.[70,71]

Therefore, medical terminology systems oriented to consumers have been developed that are supposed to bridge the language between patients and clinicians. Since 1993, in the Netherlands, the Thesaurus Zorg en Welzijn (TZW, meaning "Thesaurus Care and Well-being) was developed for this purpose.[72] In 1998, the proprietary Consumer Health Terminology was developed in the USA by WellMed.[73,74] Since 2005, in the USA, the Consumer Health Vocabulary (CHV) had been developed in an open collaboration to facilitate searching and understanding medical concepts for consumers.[74] These vocabularies contain consumer-oriented synonyms and definitions. In 2018, SNOMED CT Netherlands released the Patient-friendly Dutch language reference set, with patient-friendly synonyms and definitions from the TZW and the patient federation[75]. Incrementally, more terminology was added to the Patient-friendly Extension of the SNOMED CT Netherlands edition release from various sources, such as the Dutch platform for cancer patients (kanker.nl)[76] and the union for the Dutch language (taalunie.org).[77] See Table 1.2.

Table 1.2 Different encodings of a brain contusion for consumer-oriented and patient-friendly terminology systems, with their domain, concept id and description

Terminology system	Domain	Concept id	Description
TZW[72]	Consumer oriented	5214	damage to brain tissue because the brain is shaken back and forth by a sudden, violent movement (translated from Dutch)
CHV[74]	Consumer oriented	C0750971	Cortical contusion
SNOMED CT Netherlands Patient-friendly Dutch language reference set (synonym)[78]	Patient portals and personal health records	262689001	Hersenkneuzing (Dutch for "brain contusion")
SNOMED CT Netherlands Patient-friendly Dutch language reference set (definition)[78]	Patient portals and personal health records	262689001	This is damage to brain tissue because the brain is shaken back and forth by a sudden, violent movement. (translated from Dutch)

1.7 Aims and outline of this thesis

To realize patient access to EHRs and enable patients to use and understand medical data, several problems need to be addressed. Successful adoption of HISs, including patient portals and PHRs, is dependent upon a wide range of factors, such as laws, patient and provider perspectives, system usability, healthcare outcomes and productivity. Hence, primary studies and systematic reviews on patient access have been carried out from different theoretical perspectives and focus areas, which makes it difficult to get an overview of the determinants and outcomes of patient access to EHRs. The first aim of this thesis is, therefore, to provide an overview of the determinants and outcomes of patients accessing their EHRs. Thus we formulated the following research objective: to synthesize the results from the scientific literature on determinants and outcomes of patients accessing their EHRs through patient portals and PHRs. Chapter 2 contains a systematic review of systematic reviews on this topic. Chapter 3 reports how we adapted a theoretical adoption framework, to synthesize the results from the reviews, and developed guidance documentation to categorize determinants and outcomes into the framework.

Consumer-oriented and patient-friendly terms are required to clarify terms in free text and encoded data in patient portals and PHRs. However, the number of medical concepts for which consumer-oriented or patient-friendly terms were available was limited and few implementations existed in Dutch patient portals. The second aim of this thesis is therefore to develop and evaluate applications using medical terminology systems to clarify medical terms to patients. The objectives are, first, to assess to what extent providing clarifications to terms, which we call lexical clarification, helps patients read their clinical notes; second, to develop a novel approach to increase the number of diagnoses that can be clarified, by generalization to concepts with patient-friendly terms; third, to validate diagnosis clarifications and assess problems involved in generalization; and, fourth, to implement the clarifications in a hospital patient portal and evaluate the use of the clarification functionality in clinical practice. Chapter 4 describes the evaluation of a lexical clarification tool using the TZW consumer health vocabulary in a patient portal. Chapter 5 introduces the approach to generating diagnosis clarifications by generalizing them to concepts with patient-friendly terms using the SNOMED CT hierarchy. Chapter 6 reports the validation of this approach of diagnosis clarification by generalization to concepts with patient-friendly terms, but also with definitions of those concepts. Chapter 7 evaluates

the implementation of patient-friendly clarifications in a hospital patient portal with actual patient portal users.

Finally, Chapter 8 presents an overall discussion of the main findings of this thesis. It includes the methodological reflections and limitations, implications for clinical practice, and future research questions and finalizes with the conclusions of this thesis.

2

Determinants and outcomes of patient access to medical records:
Systematic review of systematic reviews

Authors:
Hugo J.T. van Mens
Ruben D. Duijm
Remko Nienhuis
Nicolette F. de Keizer
Ronald Cornet

Journal: International Journal of Medical Informatics.
Year: 2019. **ISSN:** 1386-5056. **Volume:** 129. **Pages:** 226-233.
DOI: 10.1016/j.ijmedinf.2019.05.014.

Abstract

Background Patient access to electronic health records (EHRs) is associated with several determinants and outcomes, which are interrelated. However, individual studies and the reviews summarizing them have only addressed particular aspects, such as policy, usability or health outcomes of adoption. Therefore, no comprehensive overview exists. Additionally, reviews used different theoretical frameworks, which makes results difficult to compare.

Objective We aimed to systematically review recent systematic reviews on determinants and outcomes of patient access to EHRs to create a comprehensive overview and inform policy-makers and EHR implementers about the available literature, and to identify knowledge gaps in the literature reviews.

Methods We searched MEDLINE, EMBASE and PsycINFO for systematic reviews on patient portals, personal health records, and patient access to records that addressed determinants and outcomes of adoption. We synthesized the results from these reviews into the Clinical Adoption Framework (CAF), by mapping quotes from the reviews to categories and dimensions of the CAF, starting with the most recent ones until saturation of the CAF had been reached. The risk of bias in the reviews was assessed using the AMSTAR2 checklist.

Results We included nineteen reviews from 8871 records that were retrieved until February 19th, 2018. The reviews had a median of 4 (IQR: 4-4) critical flaws according to the AMSTAR2 checklist. The reviews contained a total of 1054 quotes that were mapped to the CAF. All reviews reported on the dimension 'People' that can affect adoption (e.g. personal characteristics such as age) and the dimension 'HIS use' (health information system use). Most reviews reported the dimensions 'Organization', 'Implementation', HIS 'System quality', and outcomes of HIS use. However, gaps in knowledge might exist on macro-level determinants and outcomes, such as healthcare standards, funding, and incentives, because few reviews addressed these aspects.

Conclusions No review covered all aspects of the CAF and there was a large variety in aspects that were addressed, but all dimensions of the CAF were addressed by at least two reviews. Although reviews had critical flaws according to the AMSTAR2 checklist, almost half of the reviews did use methods to assess bias in primary studies. Implementers can use the synthesized results from this study as a reference for implementation and development when taking quality restrictions into account. Researchers should address the risk of bias in primary studies in future reviews and use a framework such as CAF to make results more comparable and reusable.

2.1 Introduction

Medical records have primarily been kept by clinicians in order to support their clinical work. Recent developments in healthcare technology have provided patients access to and control over their own medical records. Patient-held paper records have been used in different settings[79], and patient access to medical records, in general, has already been a legal right in many countries[80-82]. Increasingly, patient portals[81-84] provide patients with direct access to information in electronic health records (EHRs) of clinicians. Electronic personal health record systems[60,85] (PHRs) provide patients with their own system to manage their personal health information. The Blue Button initiative in the USA enabled patients to download their medical data first as a free text or pdf file[86,87] and later as a structured and standardized electronic format following HL7 C-CDA[88]. In the EU, a similar standard has been developed, called the Patient Summary[89]. Another initiative, OpenNotes[90,91], stimulates clinicians to share their visit notes with patients. Research on these approaches to provide patients access to personal health information addressed various aspects of adoption, such as influence of patient access to EHRs on patients' health outcomes (e.g. hospital admissions), patient engagement, but also barriers and facilitators to adoption, attitudes of patients and providers towards patients' access to EHRs, or specific patient groups such as psychiatric patients.[79-81,92-98] However, individual studies and the reviews summarizing them only addressed particular aspects, while patient access to medical records involves many interrelated aspects that transcend particular scientific paradigms, such as policy (politics), usability (software engineering) and health outcomes (medicine) of adoption. Therefore, a synthesis of these results is required, to provide a comprehensive overview, and to inform policy-makers and implementers of systems and functionalities that provide patients access to their personal health data.

The Clinical Adoption Framework (CAF) is a general evaluation framework to assess the success of healthcare information system (HIS) adoption in healthcare organizations. In this study, we used the CAF to categorize the information extracted from the literature.[99] As shown in Figure 2.1, it addresses the micro level, which encompasses the dimensions quality, use and net benefits of the HIS; the meso level, consisting of dimensions of people, organization and implementation; and the macro level, incorporating dimensions of healthcare standards, legislation, policy and governance, funding and incentives, and societal, political and economic trends. It is hence an integrated framework that covers a wide range of aspects that influence and result from HIS adoption.

We use determinants and outcomes as overarching terms to refer to what different studies call e.g. factors, barriers, facilitators, determinants, outcomes, mechanisms, problems, solutions, advantages, disadvantages, costs, or benefits. The determinants are those categories in the CAF that influence or are associated with HIS adoption and the outcomes are the 'Net benefits' of HIS adoption.

The purpose of this systematic review of systematic reviews is to provide a comprehensive overview of determinants and outcomes of patient access to and control over their personal medical data, and the adoption of patient portals and PHRs.

2.2 Methods

In this section we provide a summary of the methods; further details can be found in the protocol which was registered at PROSPERO under CRD42018084542[101,102].

2.2.1 Search strategy

We used the search interface Ovid for the databases "Epub Ahead of Print, In-Process & Other Non-Indexed Citations, Ovid MEDLINE(R) Daily and Ovid MEDLINE(R)" (referred to as "MEDLINE" hereafter), "EMBASE" and "PsycINFO". The search strategy consisted of keywords about patient access to records, patient portals, personal health records, Blue Button, and OpenNotes, and combinations of terms about access, records, reviews, and outcomes. The search queries were developed together with a clinical librarian and can be found in the PROSPERO record[103] and protocol[102].

2.2.2 Eligibility criteria

We looked for systematic literature reviews of studies with patients, informal caregivers or healthcare professionals in primary, secondary or tertiary healthcare, in any medical domain. The intervention in the studies in the reviews should have been providing patients access to or control over their own medical records, or adopting or using patient portals and personal health records. The primary outcomes of the review should have been on determinants or outcomes of patient access to and control over their own medical data. Articles were excluded if they did not have an abstract or were written in a language other than English.

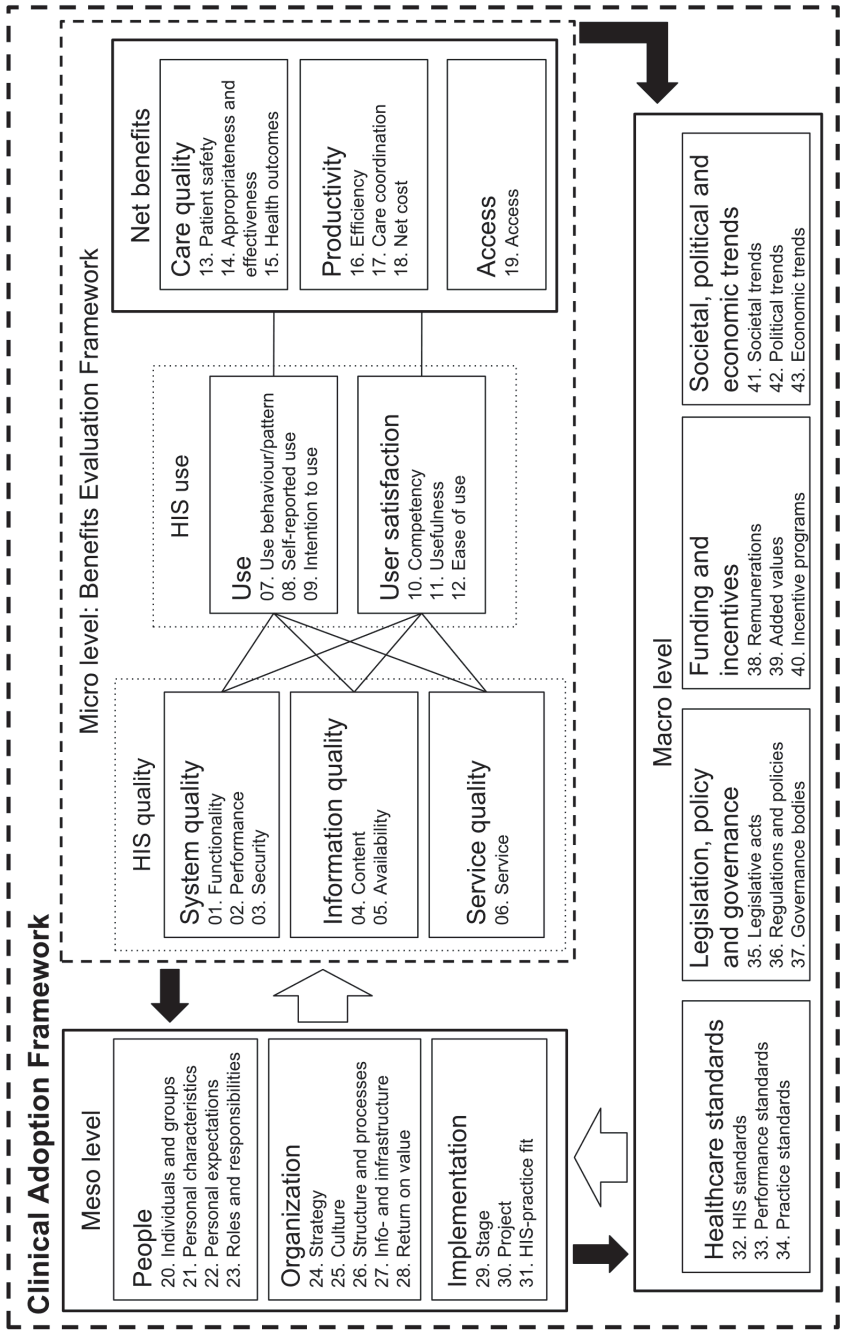


Figure 2.1 Clinical Adoption Framework (CAF) with levels, dimensions and categories, figure adapted from Lau (2012)100, by adding the CAF-categories and numerical id's (01 to 43) to the figure

2.2.3 Study selection

References were imported into Endnote X7.8 (Thompson Reuters, Toronto, ON, Canada) and duplicates were removed. The screening process was carried out with the Rayyan (Qatar Computing Research Institute, Doha, Qatar) web application, independently by two reviewers, using the eligibility criteria above. The full-text review was also carried out by two independent reviewers. Differences of opinion were discussed until a consensus was reached. Where necessary a third reviewer was involved. We limited the search to the most recent publications until saturation was achieved regarding the summary measures.

2.2.4 Data collection process

From each study, quotes were extracted about determinants and outcomes of patient access to records, patient portals, and PHRs. Additionally, any categorization of these quotes including any theoretical framework used in the original papers was also extracted together with the quotes, in order to preserve the original context and meaning. Furthermore, information about the reviews themselves was extracted, such as author, year, theoretical framework and critical appraisal method. All results and data of the reviews were extracted into a spreadsheet by the first author (HM) and verified by a second reviewer (RC, NK, RN, and RD).

2.2.5 Risk of bias

To assess the risk of bias (RoB) of the included reviews the AMSTAR2[104] checklist was filled out by two reviewers (HM and RD), and discussed until consensus was reached. This critical appraisal tool can indicate critical flaws in a systematic review that may lower the overall confidence in the review. Each question of the AMSTAR2 addresses a potential flaw in a review. No risk of bias on outcome level was assessed because this should already have been carried out by the included reviews themselves. Therefore, we did not use any statistical tools to assess RoB across studies, such as funnel plots to assess publication bias. However, the AMSTAR2 does address whether the included reviews took RoB across studies into account, e.g. publication bias and selective reporting. The number of critical and non-critical flaws in the reviews according to the AMSTAR2 were counted and a median and interquartile range (IQR) of the number of flaws per review were calculated.

2.2.6 Synthesis of results

The quotes of the reviews on determinants and outcomes were mapped to two categories of the CAF[99]: one determinant and one outcome category, see

the example in Figure 2.2. The mapping was carried out by two independent reviewers and discussed until consensus was achieved. Note that an outcome can be a determinant again of another outcome. For example, '21. Personal characteristics' can be a determinant of '07. Use behaviour/pattern', where '07. Use behaviour/pattern' will be the outcome category. '07. Use behaviour/pattern' can be a determinant again of '19. Access'. All categories in the macro and meso level, and in the 'HIS quality' and 'HIS use' sublevels are determinants of the adoption of HISs. Adoption may have several outcomes in the 'Net benefits' sublevel of the micro level (see Figure 2.1).

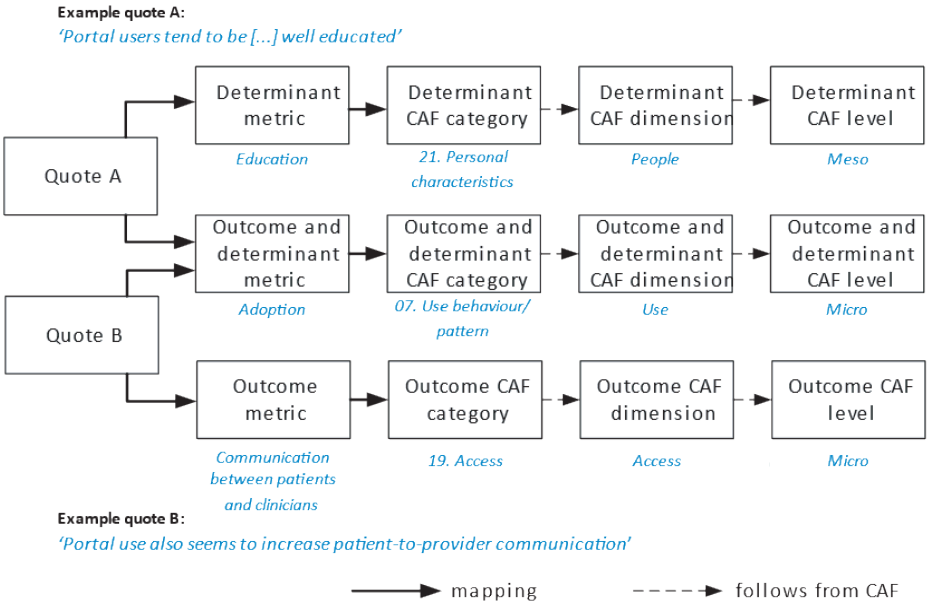


Figure 2.2 Examples of how two quotes are mapped to a determinant metric and outcome metric, which are classified as categories of the CAF, which belong to CAF-dimensions and CAF-levels. The outcome described in quote A[69] is the determinant again described in quote B[105].

2.2.7 Summary measures

For each review, the number of unique combinations of determinant categories and outcome categories was calculated. Starting with the most recent publication year we processed older literature until the categories were saturated. We defined saturation as the point at which the reviews contributed on average to less than 5% of unique combinations of categories. We calculated how many reviews reported on each dimension and category and how many reviews reported how many relationships between different levels, dimensions and categories. Statistical analysis was carried out in R version 3.5.0 (R Foundation for Statistical

Computing, Vienna, Austria) with RStudio 1.1.453 (RStudio Inc., Boston, MA, USA). See R script in Appendix A.

2.3 Results

In this section, we will first report the study selection. Secondly, we describe the study characteristics. Thirdly, we show the risk of bias of the included studies. Finally, we provide the synthesis of the results on determinants and outcomes of patient access to medical records.

2.3.1 Study selection

In total 8871 records were retrieved with the search strategy of which we screened 1862 records by title and abstract and finally included nineteen reviews[56-58,68,69,105-118], see Figure 2.3. In Appendix B the list of 37 studies that were excluded with reasons after full-text review for eligibility can be found. The nineteen included reviews had a median number of 1 (IQR: 1-3) unique mapping to combinations of determinants and outcomes, which were not found in other reviews. This is 0.9% (IQR: 0.9%-2.7%) of the total number of 112 unique combinations of categories. When we started our analyses in January 2017, first taking the most recent reviews, saturation below 5% was already achieved after mapping all reviews from 2015 to January 4th, 2017. Nevertheless, we decided to add reviews up to February 19th, 2018 to make the review more current. As saturation was already reached, we decided not to include any further studies or screen any titles and abstracts from previous years.

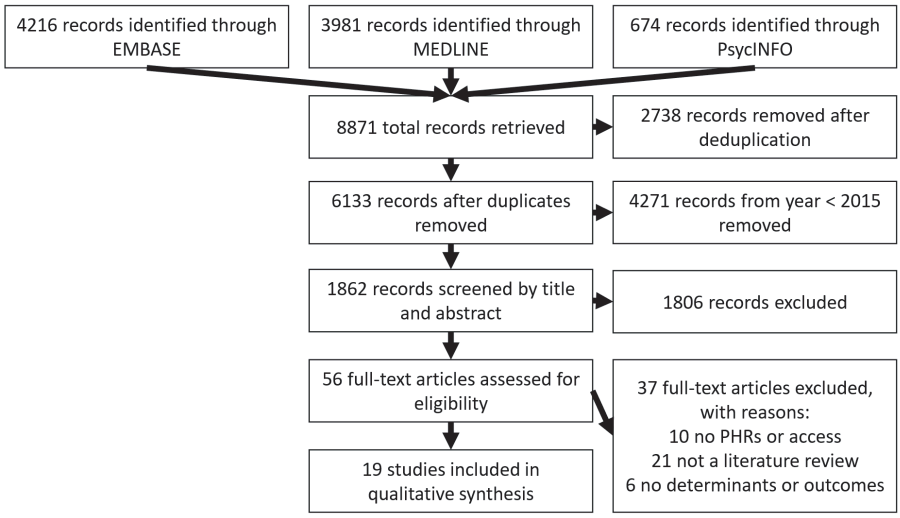


Figure 2.3 Flow chart of the study selection

2.3.2 Study characteristics

The study characteristics are described in Appendix C, Table 2.3. Seven reviews[56,68,69,105,106,110,111] were about patient portals, seven[107,109,112,115-118] about PHRs, two[57,58] about patient access to records, one [114] about patient portals and PHRs, and one[113] about patient-provider communication. Eleven reviews[56,57,68,105,107,110,111,113,115-117] were about general patient populations, two reviews[109,114] were about elderly, and five about other specific patient groups such as people with HIV[58,106,108,112,118]. Thirteen reviews did not use any theoretical framework for the analysis of results, one[107] synthesized the "The Health Literacy Skills Framework", "PHR Adoption Model" and "Integrated Model of Health Literacy", one[113] applied a conceptual framework of "Patient-centred communication functions", one[68] applied the "Problem-solving cycle", one[117] applied the "Human Factors and Ergonomics (HF/E) paradigm", one[118] applied the "Health Belief Model", and one[112] applied the "PHR Adoption Model". Ten reviews[56,68,69,105-107,110,111,113,117] did not use any method for critical appraisal, four reviews[109,114,115,118] reported the quality of reporting of the primary studies was sufficient, four reviews[57,58,108,112] used a method to assess methodological quality, and one[116] mainly addressed selection bias.

2.3.3 Risk of bias

The results of the AMSTAR2 checklist for the reviews are listed in Appendix C, Table 2.4. The median of critical flaws in the studies was 4 (IQR: 4 – 4) and the median

of non-critical flaws 5 (IQR: 4 – 5). Here we report on the critical flaws. Seventeen reviews[56,57,68,69,105-113,115-118] did not refer to a protocol that was established before the study was conducted, and that included review questions, search strategy, inclusion/exclusion criteria, and risk of bias assessment (AMSTAR2 Question 2: Q2). One review[69] did not perform a comprehensive literature review in which at least two databases were searched and the keywords or search strategy, and justified publication restrictions were provided (Q4). None of the reviews provided a list of studies that were excluded in the full-text screening for eligibility (Q7). Nine of the reviews[56,69,105,107,110,112,113,115,117] that included RCTs did not perform risk of bias assessment of unconcealed allocation and lack of blinding (Q9a). Eighteen of the reviews[56-58,68,69,105-107,109-118] that included non-randomized studies of interventions did not both assess confounding and selection bias (Q9b). Eighteen reviews[56-58,68,69,105-115,117,118] did not take the risk of bias of individual studies into account when interpreting and discussing the results of the review (Q13).

2.3.4 Synthesis of results on determinants and outcomes

The nineteen reviews contained 1054 quotes about determinants or outcomes of patients access to medical records that were mapped to the CAF. Details on each quote can be found in Appendix D: in this spreadsheet one can filter on each level, dimension, category, and metric that quotes were mapped to, filter on author, and text search for particular information. The spreadsheet is also available as a Google Sheet. In total 810 quotes were about patients, 136 about care providers, 64 about informal caregivers and 44 other. Figure 2.4 shows how many reviews reported relationships between the macro and meso level, and the sublevels of the micro level of the CAF. This shows that most reviews reported on relationships between the meso level and the sublevels of the micro level, but only three reviews[106,115,117] reported on relationships between the macro and meso level. Each level has several dimensions. The number of reviews that addressed a certain dimension in the CAF is shown in Figure 2.5. This shows that all reviews referred to ‘People’ and ‘Use’ dimensions and that all CAF dimensions were addressed by at least two reviews. One quote[68], about the infrastructure outside of an organization to exchange data between hospitals, could not be mapped into the CAF and is displayed in the dimension ‘Other’ in this figure. Table 2.1 shows relationships between CAF dimensions that were found most often in the reviews. For example, eighteen reviews contained a total of 264 quotes on relationships between the dimensions ‘People’ and ‘Use’. For the CAF categories (i.e. that belong to a certain CAF dimension), Table 2.2 shows relationships between CAF categories that were found most often in the reviews. For example, seventeen reviews contained a total of 86 quotes on a relationship between the

categories '22. Personal expectations' and '07. Use pattern/behaviour'. Appendix C, Table 2.5 shows the metrics in each category that we found and the number of reviews that referred to these categories. For example, we distinguished several specific functionalities found in seventeen reviews and mapped them as metrics of the category '01. Functionalities'.

Particular quotes about functionality such as secure messaging, medical outcomes, and other metrics can be found in the spreadsheet. For example, eight reviews[57,69,105,106,108,110,112,113] reported on the CAF category '05. Health outcomes' and one of the metrics we distinguished in this category was 'Physiological outcomes'. Two reviews[69,106] reported on how portal adoption and functionalities such as secure messaging and medication refills were associated with improvements in physiological outcomes, such as glycaemic control and blood pressure, as exemplified by the quote: 'The one study that looked at a clinical outcome found that among the entire study population 10–99 years of age, portal use was a statistically significant predictor of glycosylated haemoglobin level but not of low-density lipoprotein and total cholesterol levels'[106].

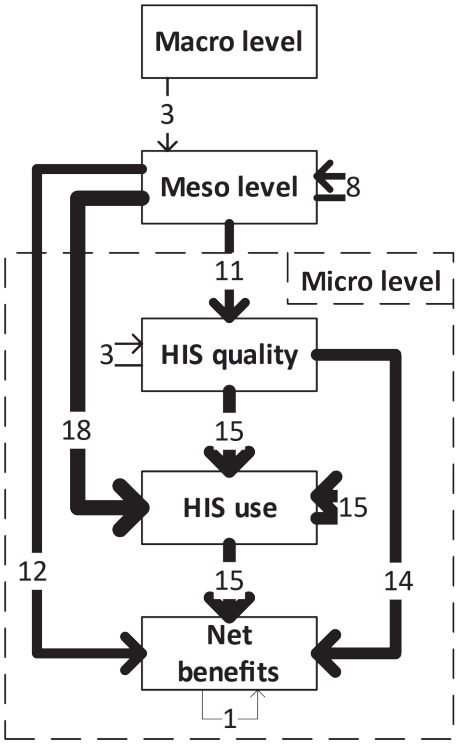


Figure 2.4 Number of reviews that found relationships between macro and meso level and the sublevels in the micro level. The width of the arrows corresponds to the number of reviews.

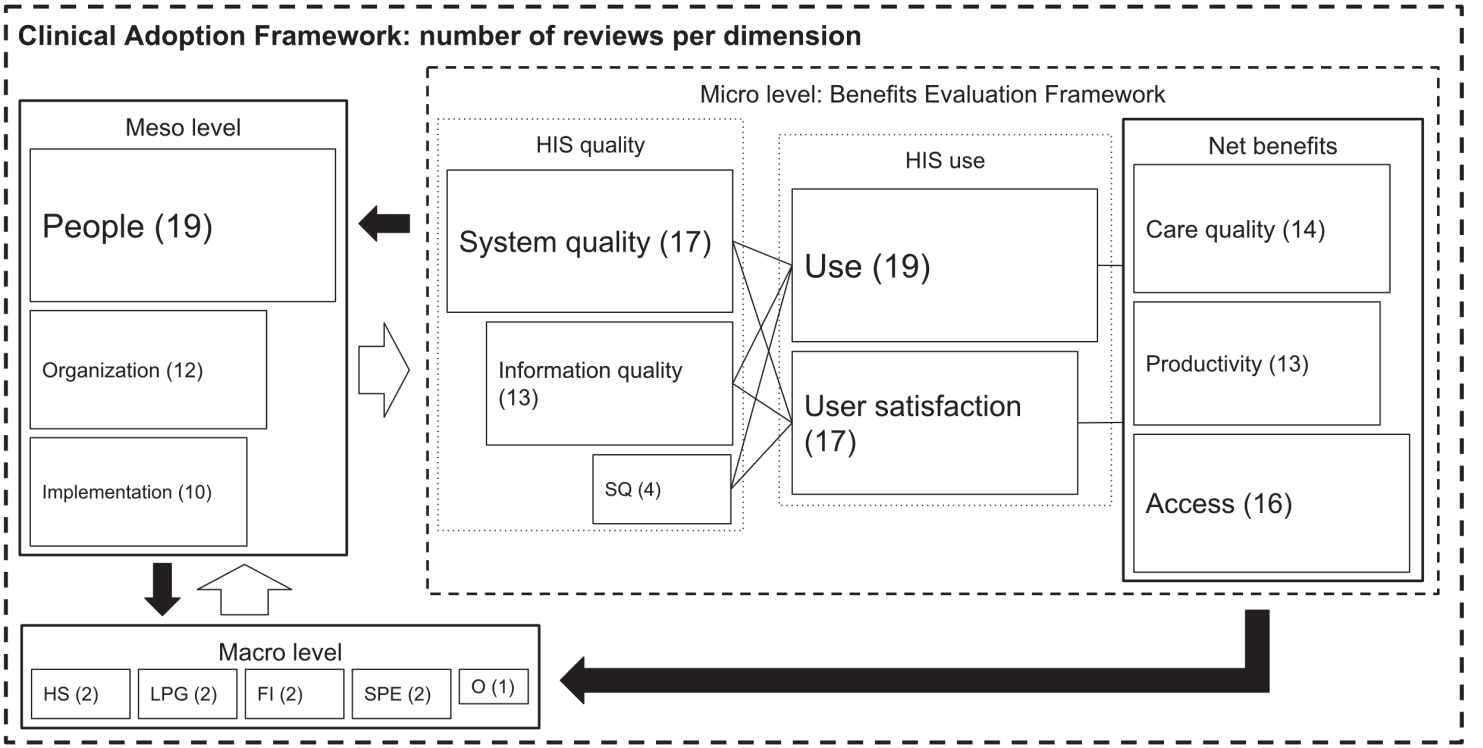


Figure 2.5 Number of reviews per dimension, the surface of the boxes of the dimensions is proportional to the number of reviews n (i.e. the height is the root of half n and the width is the twice the height). SQ = service quality, HS = healthcare standards, LPG = Legislation, policy and governance, FI = Funding and incentives, SPE = Societal, political and economic trends, O = Other

Table 2.1 Top 10 of most found relationships between determinant dimension and outcome dimension, the number of quotes (Q) where the relationship was found, the number of reviews (N) in which it was reported

Determinant dimension	Outcome dimension	Q	N
People	Use	264	18
Use	Access	41	13
System quality	Access	56	12
Use	Care Quality	53	12
System quality	Use	39	12
System quality	User satisfaction	37	12
System quality	Care Quality	64	9
Implementation	Use	35	9
People	System quality	76	8
Organization	Use	34	8

Table 2.2 Top 10 of most found relationships between determinant category and outcome category id and description, the number of quotes (Q) where the relationship was found, the number of reviews (N) in which it was reported and an example of a quote

Id	Determinant category	Id	Outcome category	Q	N	Example quote
22	Personal expectations	07	Use pattern/behaviour	86	17	'patient use was limited by patient concerns about confidentiality of their personal health data'[68]
21	Personal characteristics	07	Use pattern/behaviour	175	13	'Online record access and service users tended to be slightly older (t-test, P<0.001)'[58]
07	Use pattern/behaviour	19	Access	41	13	'Some patients used the EHR to influence physicians, for example, by using its information to challenge statements made by physicians'[113]
01	Functionality	19	Access	56	12	'PLHIV [i.e. people living with HIV] described how access to their own health information within the PHR gave them a greater sense of control over their health.'[118]
01	Functionality	07	Use pattern/behaviour	34	11	'One study enumerated why the portal is used; parents with chronically ill children enrolled in a large health organization most frequently used immunization records, secured messaging, and scheduling appointments'[106]
07	Use pattern/behaviour	14	Appropriateness and effectiveness	37	10	'Portal users also noted greater medication adherence, particularly for those individuals with chronic illnesses like diabetes.'[105]
01	Functionality	14	Appropriateness and effectiveness	49	9	'Several studies describe how patients, after reading their medical file, gained a better understanding and recollection of their health status and physician instructions,'[57]
30	Project	07	Use pattern/behaviour	27	8	'Lam et al [...] found that participants were significantly more likely to be introduced to a portal messaging system by their providers than were nonusers [...].'[114]
01	Functionality	11	Usefulness	26	8	'The theme 'practical benefits of the PHR' encompassed the usefulness of the PHR as a concrete, everyday tool for the patient. In seven studies, patients valued having their own record of their condition, and so the PHR was a good source of personal health information, such as the results from tests and scans and details of next appointments.'[115]
27	Info- and infrastructure	07	Use pattern/behaviour	17	8	'Internet access'[117] (as work system barrier for the adoption of PHRs)

2.4 Discussion

2.4.1 Principal findings

Recent reviews on determinants and outcomes of patient portal and PHR adoption, and patient access to and control over their own medical records reported on all dimensions of the CAF, and mostly on the people and organizations that influence adoption, but less on the higher level of standards, policy, funding, and society. Even though reviews applied different theoretical frameworks and focused on different medical domains, we were able to synthesize them into the overarching CAF. This enables comparison and reusability of results from the reviews. All reviews had several critical flaws, however, and therefore presumably have limited reliability according to the risk of bias assessment.

2.4.2 Strengths and weaknesses

We used novel visualization methods to present the results integrated into the CAF, using both the differing surfaces of CAF dimensions depending on the number of reviews that mentioned them and relationships between (sub) levels as visualization methods. The mapping of the results from the systematic reviews to the CAF indicated where gaps in the reviewed systematic reviews lied: on the macro level and service quality dimension. Gaps are not considered to indicate importance of dimensions, which depends on the specific research question or practical problem one wants to address.

Reviews were generally not systematic reviews of RCTs or non-randomized quantitative studies, but rather systematic reviews of qualitative or mixed methods studies. The quality of each result within the included reviews was not critically appraised in the current review, and therefore it was not possible to do a meta-analysis or assess the evidence for particular relationships between determinants and outcomes. It was not possible to provide a single negative or positive direction for each result. Because reviews may include the same primary studies, some of the categories might be overrepresented due to the reporting in the reviews rather than the primary studies.

The overall confidence in the results of each review was critically low (defined by more than one critical flaw) according to the AMSTAR2 checklist, because all reviews had around four critical flaws: most reviews did not register a protocol before starting to carry out the review, did not justify the exclusion of individual studies under full-text review for eligibility, did not carry out a proper Risk of Bias (RoB) assessment, and did not take RoB into account when interpreting their

results. However, we did find almost half of the reviews using methods, such as assessing the quality of reporting, unfortunately not covering all the aspects that were required by the AMSTAR2 for addressing RoB. Furthermore, some reviews noted it was not possible to carry out an RoB assessment because of the qualitative nature, the heterogeneity of methods, or the preliminary nature of some studies. AMSTAR2 does not address quality assessment of these types of studies, as it is mainly focused on reviews of controlled quantitative studies.

2.4.3 Practical implications

Policymakers and implementers of PHRs and patient portals can refer to relationships summarised in this systematic review synthesis. The available reviews focussing on a particular relationship can be easily selected by using the provided Excel spreadsheet in Appendix D and the Google Sheet. This supports the use of available evidence when addressing aspects that need to be tackled and prioritized during implementations and development. For example, one might be interested in the association between adoption and medical outcomes, like we illustrated in the results section, and will find eight relevant reviews and their particular quotes about these associations. Similarly, an implementer can easily select reviews that associate training and usability testing with adoption. The results of our review can also serve as a check whether most domains have been covered in implementation project definition and evaluation, and in the design of solutions. Furthermore, the results show that some possible relationships between categories have not been addressed in the original reviews, which might indicate a need for more research on these topics. We suggest addressing the relationship between the macro level and the meso level, either in reviews of other types of literature (such as in law and policy-making rather than the medical literature) or in new primary studies.

2.5 Conclusion

Determinants and outcomes were synthesized into the CAF. Reviews on patient access to medical records, patient portals, and personal health records mostly addressed people, organization, implementation, HIS quality, HIS use, and net benefits of HIS use. To a lesser extent, healthcare standards, legislation, policy, governance, funding, incentives, and social, political, and economic trends were reported. The results provide a reference when realizing patient access to health records, and suggestions for further research on HIS adoption, however, one must take the modest quality of the reviews into account when implementing their results in practice. Future reviews should, therefore, address the risk of

bias in primary studies and carry out meta-analyses on particular determinants and outcomes. Reviews and primary studies should use an integrated theoretical framework such as the CAF to make results more comparable. More reviews on relationships between the CAF macro level and the meso level are needed because these aspects were covered to a lesser extent in the included reviews.

2.6 Summary points

What was already known on the topic?

- Patient access to electronic health records (EHRs) is associated with several interrelated determinants and outcomes. However, no comprehensive overview existed.
- The Clinical Adoption Framework (CAF) is an evaluation framework to assess the success of health information systems adoption.

What this study added to our knowledge?

- We provided an overview of determinants and outcomes of patient access to EHRs, extracted from systematic reviews, by using the CAF. Systematic reviews on patient access to EHRs have several critical flaws, however, which may negatively impact their quality.
- Literature reviews indicate a gap in knowledge, as they cover few high-level aspects regarding healthcare standards, legislation, policy, governance, funding, incentives and social, political and economic trends.

2.7 Acknowledgements

Thanks to clinical librarian Joost Daams for assisting in the development of the search strategy.

2.8 Appendices

2.8.1 Appendix A

R script with syntax for analysis, is available in the online version of this paper at https://doi.org/10.1016/j.ijmedinf.2019.05.014.

2.8.2 Appendix B

List of studies excluded after full-text review for eligibility.

In the full-text review for eligibility, 10 studies[119-128] were excluded because they were not primarily about patient portals, PHRs or patient access to records, 21 studies[129-149] were excluded because these were no literature reviews or were conference proceedings, and 6 studies[150-155] because they were not about determinants or outcomes of patient portals, PHRs or patient access to records.

2.8.3 Appendix C

2.8.3.1 Study characteristics

Table 2.3 Study characteristics

RefId	Author	Year	Patient	System	Databases	Time	Theory	Appraisal
1[107]	Hemsley	2018	General	PHR	Medline, Web of Science, CINAHL, and PsychInfo	- October 2015	Synthesis of “The Health Literacy Skills Framework”, “Personal Health Record Adoption Model” and “Integrated Model of Health Literacy”	-
2[69]	Coughlin	2017	Prevention	Patient portal	PubMed	1993 - November 2016	-	-
3[108]	Kelly	2017	Inpatient	Inpatient portal	PubMed, Web of Science (including the Institute of Electrical and Electronics Engineers Xplore), Cochrane, CINAHLPlus, and Scopus databases.	January 1, 2006, to August 8, 2017	-	Methodological quality (Downs and Black)
4[109]	Kneale	2017	Older adults	PHR	Medline, CINAHL, PsycINFO and EMBASE	- July 1, 2015	-	STARE-HI
5[111]	Powell	2017	General	Patient portal	CINHAL, PsycINFO & MEDLINE-PubMed	2009 - November 2016	-	-
6[113]	Rathert	2017	General	Patient-provider communication	CINAHL, Medline, PsycINFO	- 2015	Patient-centred communication functions	-

Table 2.3 Study characteristics (**continued**)

RefId	Author	Year	Patient	System	Databases	Time	Theory	Appraisal
7[114]	Sakaguchi	2017	Older adults	Patient portals and PHRs	PubMed, EMBASE, CINAHL Complete, Compendex (includes ACM digital library and IEEE XPlore), and Inspec	January 2006 - November 2016	-	STARE-HI top 2 criteria + 2 of their own
8[116]	Showell	2017	General	PHR	PubMed, Embase, CINAHL and ProQuest	2003 -	-	Selection bias
9[57]	Vermeir	2017	General	Patient access to medical records	PubMed, Web of Science, Cinahl, and Cochrane Library	1 January 2002 - 31 January 2016	-	Methodological quality (Hawker et al)
10[68]	Otte-Trojel	2016	General	Patient portal	PubMed, ScienceDirect and LISTA	2005 - 2015	Problem-solving cycle	-
11[117]	Thompson	2016	General	PHR	PubMed, CINAHL (excluding MEDLINE records), Engineering Village (Compendex and INSPEC), IEEE Xplor, and ACM Digital Library	2000 - 2013	Human Factors / Ergonomics paradigm use barriers	-
12[118]	Turner	2016	HIV	PHR	PubMed, CINAHL, Web of Science, Scopus, EMBASE, and PsycINFO	2009 - 2015	Health belief model	Assessment of reporting quality (Mills, 2006)
13[106]	Bush	2015	Paediatric population	Patient portal	PubMed, CINAHL Plus, PsycINFO, and Academic Search Premier	1992 - August 2014	-	-
14[56]	Irizarry	2015	General	Patient portal	PubMed, Ovid Medline, and PsycInfo	2006 - 2014	-	-

Table 2.3 Study characteristics (**continued**)

RefId	Author	Year	Patient	System	Databases	Time	Theory	Appraisal
15[105]	Kruse1	2015	General	Patient portal	PubMed, CINAHL	2004 - 2014	-	-
16[110]	Kruse2	2015	General	Patient portal	PubMed, CINAHL, Google Scholar	January 1, 2011, and August 24, 2014	-	-
17[58]	Mold	2015	Primary care	Patient access to medical records	Medline, Embase, CINAHL, Cochrane Library, EPOC, DARE, King's Fund, Nuffield Health, PsycINFO, OpenGrey	1999 - 2012	-	RoB assessment, intention to use GRADE
18[112]	Price	2015	Patients with chronic diseases	PHR	Medline, CINAHL	2008 - 2014	PHR Adoption Model	Extended NHMRC evidence hierarchy
19[115]	Sartain	2015	General	PHR	through EBSCO Host in CINAHL Plus with Full Text (1982 to March week 4 2013), MEDLINE (1950 to March week 4 2013), PsycINFO (1597 to March week 4 2013), PsycARTICLES (1894 to March week 4 2013) and the Cochrane Library (1995 to present)	- 2013	-	Consolidated criteria for reporting qualitative research (COREQ)

2.8.3.2 AMSTAR2 checklist results

Table 2.4 AMSTAR2 checklists: y = yes, n = no, p = partial yes, m = no meta analysis, o = only NRSI. The criticality of a flaw when the answer to a question is no is noted in the top row, where the critical questions are marked with an asterisk *. The last two columns count the number of critical and non-critical flaws in the review. The bottom rows list the total number of reviews with answers to AMSTAR2 score with yes, partial yes, no and not applicable, the dash indicates the answer was not an option for the particular question.[a]

RefId	Review	Critical	*		*			*		*			*		*		*		Critical	Non-critical
		Year	1	2	3	4	5	6	7	8	9a	9b	10	11a	11b	12	13	14	15	16		
1	Hemsley	2018	Y	N	N	P	N	N	N	P	N	N	N	M	M	M	N	N	M	N	4	6
2	Coughlin	2017	Y	N	N	N	N	N	N	P	N	N	N	M	M	M	N	N	M	Y	5	5
3	Kelly	2017	Y	N	N	P	Y	Y	N	P	O	P	N	M	M	M	N	N	M	Y	3	3
4	Kneale	2017	Y	N	N	P	Y	N	N	P	O	N	N	M	M	M	N	N	M	N	4	5
5	Powell	2017	Y	N	N	P	N	N	N	P	O	N	N	M	M	M	N	N	M	Y	4	5
6	Rathert	2017	Y	N	N	P	Y	N	N	P	N	N	N	M	M	M	N	N	M	Y	4	4
7	Sakaguchi	2017	Y	P	N	P	N	Y	N	P	O	N	N	M	M	M	N	N	M	Y	3	4
8	Showell	2017	Y	N	N	P	N	N	N	N	O	N	N	M	M	M	Y	N	M	Y	3	6
9	Vermeir	2017	Y	N	N	P	Y	N	N	P	O	N	N	M	M	M	N	N	M	Y	4	4
10	Otte-Trojel	2016	Y	N	Y	P	Y	N	N	N	O	N	N	M	M	M	N	N	M	Y	4	4
11	Thompson	2016	Y	N	N	P	Y	Y	N	P	N	N	N	M	M	M	N	N	M	N	4	4
12	Turner	2016	Y	N	N	P	Y	N	N	N	O	N	N	M	M	M	N	N	M	Y	4	5
13	Bush	2015	Y	N	N	P	N	N	N	P	O	N	N	M	M	M	N	N	M	Y	4	5
14	Irizarry	2015	Y	N	N	P	N	N	N	N	N	N	N	M	M	M	N	N	M	Y	4	6
15	Kruse1	2015	Y	N	Y	P	N	Y	N	N	N	N	N	M	M	M	N	N	M	Y	4	4
16	Kruse2	2015	Y	N	Y	P	N	Y	N	P	N	N	N	M	M	M	N	N	M	Y	4	3
17	Mold	2015	Y	P	N	Y	Y	Y	N	N	Y	N	N	M	M	M	N	N	M	N	3	5
18	Price	2015	Y	N	N	P	Y	Y	N	P	N	N	N	M	M	M	N	N	M	N	4	4
19	Sartain	2015	Y	N	N	P	N	N	N	P	N	N	N	M	M	M	N	N	M	N	4	6

Total	1	2	3	4	5	6	7	8	9a	9b	10	11a	11b	12	13	14	15	16
Yes	19	0	3	1	9	7	0	0	1	0	0	0	0	0	1	0	0	13
Partial Yes	-	2	-	17	-	-	0	13	0	1	-	-	-	-	-	-	-	-
No	0	17*	16	1*	10	12	19*	6	9*	18*	19	0*	0*	0	18*	19	0*	6
Not Applicable	-	-	-	-	-	-	-	-	9	0	-	19	19	19	-	-	19	-

a Note that the RefId in this table refers to the RefId column in Table 2.3. The reference to the citation is followed in superscript after this RefId in Table 2.3.

2.8.3.3 Metrics and reviews per category

Table 2.5 Metrics and reviews per categoryb

Id	Level	Dimension	Category	Metrics	RefId	n
01	Micro	System quality	Functionality	Access to medical records, Add comments to notes, Adding text to pictures, Alerts, notifications and reminders, Appointment reminders, Appointment scheduling, Assistive technology, Automatic data input, Control over health information, Create own care plan, Data entry, Data exchange, Decision support, Design, Diagnostic tool, Discharge information, E-prescribing, Feature to contact multiple providers, Financial administration, Graphs, Health care services use, Health status management, Help remember daily tasks, Home monitoring, Hospital information, Immunization records, In-patient schedule, Information about daily hospital routine, Information about healthcare team, Information conveyance, Input methods, Integration of PHR with healthcare system, Journal, Lifestyle advice, Medical history, Medical renewal reminders, Medication information, Medication management, Medication renewal, Medication warnings, Medium (e.g. paper, electronic), Missing functionality, More providers in the system, Multimedia, Ordering meals, Patient education, Patient summary, Personalization, Prevention recommendations, Printing, Privacy and security settings, Prognostic tool, Provider emergency access, Proxy access, Record management, Risk assessment tool, Safety information, Secure messaging, Sharing medical information, Symptom checker, Take notes, Technology platform options, Terminology tool, Test results, Tools for health management, Tools to understand health information, Triage of messages, Video calling, Video with use instruction, View billing information, View health information, View lab results, View medication list, View notes, View operative reports, View test results, Voice recording	1, 2, 3, 4, 5, 6, 7, 9, 10, 11, 12, 13, 14, 15, 16, 17, 19	17
02	Micro	System quality	Performance	System response time, Technical issues	11	1

b Note that the RefIds in this table refer to the RefId column in Table 2.3. The reference to the citation is followed in superscript after this RefId in Table 2.3.

Table 2.5 Metrics and reviews per category (**continued**)

Id	Level	Dimension	Category	Metrics	RefId	n
03	Micro	System quality	Security	Audit trail, Authentication, Authorization, Security	3, 10, 15	3
04	Micro	Information quality	Content	Accuracy, Amount of documentation, Coded sensitive information, Consistency, Content, Intelligibility, Language, Not overwhelming, Relatedness of information, Sensitivity, View notes	3, 4, 5, 7, 9, 10, 11, 12, 13, 14, 15, 16, 17	13
05	Micro	Information quality	Availability	Availability, Timeliness	3, 10, 11, 13, 17	5
06	Micro	Service quality	Service	Clinical support, Service response time, Support, Technical support	7, 11, 15, 16	4
07	Micro	Use	Use behaviour/ pattern	Adoption, Registration, Use pattern	1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19	19
08	Micro	Use	Self-reported use			0
09	Micro	Use	Intention to use	Intention to use	3, 5, 6, 7, 8, 9, 13, 14	8
10	Micro	User satisfaction	Competency	Ability to use the system, Learnability, Skills, System feature awareness	1, 2, 4, 5, 7, 8, 10, 11, 14, 16	10
11	Micro	User satisfaction	Usefulness	Usefulness, User satisfaction	2, 3, 4, 6, 7, 9, 10, 13, 14, 16, 17, 19	12
12	Micro	User satisfaction	Ease of use	Ease of use, Navigation, Task completion	2, 3, 4, 7, 8, 10, 11, 13, 14, 15, 16, 17	12
13	Micro	Care Quality	Patient safety	Error identification, Medical error, Medication error, Patient safety	3, 5, 9, 17, 18	5

Table 2.5 Metrics and reviews per category (**continued**)

Id	Level	Dimension	Category	Metrics	RefId	n
14	Micro	Care Quality	Appropriateness and effectiveness	Appointment attendance, Being prepared for emergencies, Effective consultations, Emergency department visits, Health care services use, Immunization, Inconvenience, Knowledge and understanding about personal health, Manage administrative concerns, Medication adherence, Medication adjustment, Medication management, Number of appointments, Patient participation in health care, Preventive services use, Quality of care, Satisfaction with care, Treatment adherence	2, 3, 4, 5, 6, 7, 9, 12, 13, 14, 15, 16, 17, 18	14
15	Micro	Care Quality	Health outcomes	Confidence, Disease control, Medical outcome, Physiological measures, Psychological outcomes	2, 3, 6, 9, 13, 15, 16, 18	8
16	Micro	Productivity	Efficiency	Appointment length, Customer retention, Efficiency, Number of appointments, Number of hospitalizations, Number of messages, Number of phone contacts, Registration time, Repetition of information, Time savings, Workload (outcome)	3, 5, 6, 9, 10, 13, 14, 15, 16, 17, 19	11
17	Micro	Productivity	Care coordination	Access to information by providers, Care coordination, Continuity of care, Responsibility for record management, Workflow (outcome)	3, 7, 9, 10, 14, 18	6
18	Micro	Productivity	Net cost	Financial benefits		0
19	Micro	Access	Access	Access to personal health information, Communication between patients and clinicians, Decision making, Focus on patients, Patient-provider participation, Patient-provider relationship, Preparation for clinical visits, Responsibility for record management, Self-management, Sharing personal health information	1, 2, 3, 4, 5, 6, 7, 9, 12, 13, 14, 15, 16, 17, 18, 19	16
20	Meso	People	Individuals and groups			0

Table 2.5 Metrics and reviews per category (**continued**)

Id	Level	Dimension	Category	Metrics	RefId	n
21	Meso	People	Personal characteristics	Activeness in decision making, Address, Admission status, Adoption, Age, Age of caretaker, Autonomy, Awareness, Behaviour, Being sufficiently informed, Cognition, Computer and internet skills, Computer skills, Computer use, Conceptual knowledge, Confidence, Coping style, Education, EHR use, Employment, Engagement with PHRs, Ethnicity, Feeling responsible for their own health, Feeling sufficiently informed, Gender, Health care use, Health literacy, Health status, Insurance, Internet access, Internet use, Keeping their own medical records, Knowledge, Knowledge about right to access medical records, Language, Lifestyle, Literacy, Marital status, Medication use, Memory, Motor skills, Numeracy, Patient activation, Provider use of PHR, Receiving information from other sources, Receptiveness to information, Sensory and perceptual abilities, Skills, Socioeconomic status, Specialty, Technology use, Time constraints, Trust in healthcare provider, Use of alternative technology	1, 2, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19	17
22	Meso	People	Personal expectations	Adoption by providers, Alternative technologies, Anxiety to learn, Awareness of PHR, Beliefs, Choice for a provider, Competitiveness, Concerns, Expectations, Fear for anguish through consulting records, Fear of not understanding medical records, Intention to use, Interest in PHRs, Need to keep their own medical records, Needs, Norms, Patient safety, Perceived PHR value, Preferences, Privacy concerns, Scepticism, Willingness to pay	1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19	19
23	Meso	People	Roles and responsibilities			0
24	Meso	Organization	Strategy			0
25	Meso	Organization	Culture	Norms	11	1

Table 2.5 Metrics and reviews per category (**continued**)

Id	Level	Dimension	Category	Metrics	RefId	n
26	Meso	Organization	Structure and processes	Care process, Cooperation with healthcare provider, Involvement of healthcare team, Letting patients exchange medical data, Patient-provider relationship, Practice size, Process, Setting, Social environment, Staffing, Time constraints, Workload (determinant)	7, 9, 10, 11, 14, 17	6
27	Meso	Organization	Info- and infrastructure	Authentication mechanisms, Authorization policy, Computer access, Data exchange infrastructure, Data governance, Hardware, Infrastructure, Integration of PHR with healthcare system, Internet access, Internet and computer access, Internet speed, Interoperability, Kiosks, Security measures, Smartphone access	3, 4, 7, 8, 10, 11, 12, 14, 15, 16	10
28	Meso	Organization	Return on value	PHR value	10, 11, 16	2
29	Meso	Implementation	Stage			0
30	Meso	Implementation	Project	Communication about PHR, Demonstration, Design decisions, Family recommendation, Participatory design, Provider endorsement, Recommendations, Registration reminders, Training, Trial period, Usability testing	1, 3, 5, 7, 8, 9, 10, 11, 13, 14	10
31	Meso	Implementation	HIS-practice fit	HIS-practice fit	3, 10, 11, 13, 14	5
32	Macro	Healthcare standards	HIS standards	HIS standards, Interoperability	10, 11	2
33	Macro	Healthcare standards	Performance standards			0
34	Macro	Healthcare standards	Practice standards			0
35	Macro	Legislation, policy and governance	Legislative acts			0

Table 2.5 Metrics and reviews per category (**continued**)

Id	Level	Dimension	Category	Metrics	RefId	n
36	Macro	Legislation, policy and governance	Regulations and policies	Regulation	10, 11	2
37	Macro	Legislation, policy and governance	Governance bodies			0
38	Macro	Funding and incentives	Remunerations	Reimbursement	10	1
39	Macro	Funding and incentives	Added values	Added value	10	1
40	Macro	Funding and incentives	Incentive programs	Incentive programs	10, 16	2
41	Macro	Societal, political and economic trends	Societal trends	Education curricula, Societal trends	10, 11	2
42	Macro	Societal, political and economic trends	Political trends			0
43	Macro	Societal, political and economic trends	Economic trends			0
44	Other	Other	Other	Regional health information exchange infrastructure	10	1

2.8.4 Appendix D

Determinants and outcomes. Spreadsheet with quotes about determinants and outcomes found in each included review. This spreadsheet also contains data on the reviews, a pivot table to browse the quotes and mappings, the CAF, and the AMSTAR2 scores. See the online version of this paper at https://doi.org/10.1016/j.ijmedinf.2019.05.014. The data on the reviews and the sheet with determinants and outcomes are also available as Google Sheet https://purl.org/hjtvanmens/patientaccess.


10001011101
1100101000001111
0110011101
0110011
00000
11011

3

Towards an adoption framework for patient access to electronic health records:

Systematic literature mapping study

Authors:
Hugo J.T. van Mens
Ruben D. Duijm
Remko Nienhuis
Nicolette F. de Keizer
Ronald Cornet

Journal: JMIR Medical Informatics.
Year: 2020. **ISSN:** 2291-9694. **Volume:** 8. **Issue:** 3. **Page:** e15150.
DOI: 10.2196/15150.

Abstract

Background Patient access to electronic health records (EHRs) is associated with increased patient engagement and health care quality outcomes. However, the adoption of patient portals and personal health records (PHRs) that facilitate this access is impeded by barriers. The Clinical Adoption Framework (CAF) has been developed to analyse EHR adoption but this framework does not consider the patient as an end-user.

Objectives We aim to extend the scope of the CAF to patient access to EHRs, develop guidance documentation for the application of the CAF and assess the interrater reliability.

Methods We systematically reviewed existing systematic reviews on patient access to EHRs and PHRs. Results of each review were mapped to one of the 43 CAF categories. Categories were iteratively adapted where needed. We measured the interrater reliability with Cohen's unweighted Kappa and statistics regarding the agreements among reviewers on mapping quotes of the reviews to different CAF categories.

Results We further defined the framework inclusion and exclusion criteria for 33 of the 43 CAF categories and achieved a moderate agreement among the raters, which varied between categories.

Conclusions In the reviews, categories about people, organization, system quality, system use and the net benefits of system use were addressed more often than those about international and regional information and communication technology infrastructures, standards, politics, incentive programs and social trends. Categories that were addressed less might have been underdefined in this study. The guidance documentation we developed can be applied to systematic literature reviews and implementation studies, patient and informal caregiver access to EHRs, and the adoption of PHRs.

3.1 Introduction

Patient access to electronic health records (EHRs) is becoming increasingly common and is even a legal right in many countries. Patient access to EHRs[c] has been associated with increased patient engagement and improved health care quality outcomes.[80,81,93-97,156] However, there are also barriers to patients' access to EHRs. For example, some patients have difficulties logging in to patient portals and personal health records (PHRs) which facilitate access, due to complicated security procedures. [80,81,93-97,156] A framework is needed to assess the determinants and outcomes of PHR and EHR adoption that facilitates this access. This framework should consider patients and informal caregivers as users rather than health care providers alone. This framework would enable the comparison and aggregation of evidence, and provide an overview of any important factors involved, which can then be used as a guide in implementations and health care policies, as well as to address the gaps in knowledge.

'The Clinical Adoption Framework (CAF) is a general evaluation framework to assess the success of health information system (HIS) adoption in health care organizations.'[99,157] PHRs and EHRs are types of HISs, and thus this framework is also applicable to them. 'As shown in Figure 3.1, it addresses the micro level, which encompasses the dimensions of quality, use and net benefits of the HIS; the meso level, consisting of the dimensions people, organization and implementation; and the macro level, incorporating the dimensions health care standards, legislation, policy and governance, funding and incentives, and societal, political and economic trends'.[157] Within each dimension, several categories were distinguished, for example "01. Functionality", "02. Performance" and "03. Security" are categories of the dimension System quality at the micro level. 'It is hence an integrated framework that covers a wide range of aspects involved in HIS adoption'.[157] The CAF was developed and validated through consultation with health information technology professionals, comparisons with other survey instruments and a meta-review of 50 systematic reviews on HIS implementation.[158] Categories, dimensions and levels of the CAF were originally described by Lau, Price and Kashevjee.[99] Throughout the categories, dimensions, and levels there are feedback loops, which are indicated by the arrows in Figure 3.1, that resembles the interplay between the factors and nondeterministic characteristics of HIS adoption and outcomes of HIS use.[99,158] The CAF was applied in over 30 studies.[100,159-164]

c During copy-editing the journal replaced 'Patient access to EHRs has' for 'EHRs have' assuming EHR to be the abbreviation of 'Patient access to electronic health records (EHRs)' from the first sentence. In this thesis this formulation was corrected for legibility.

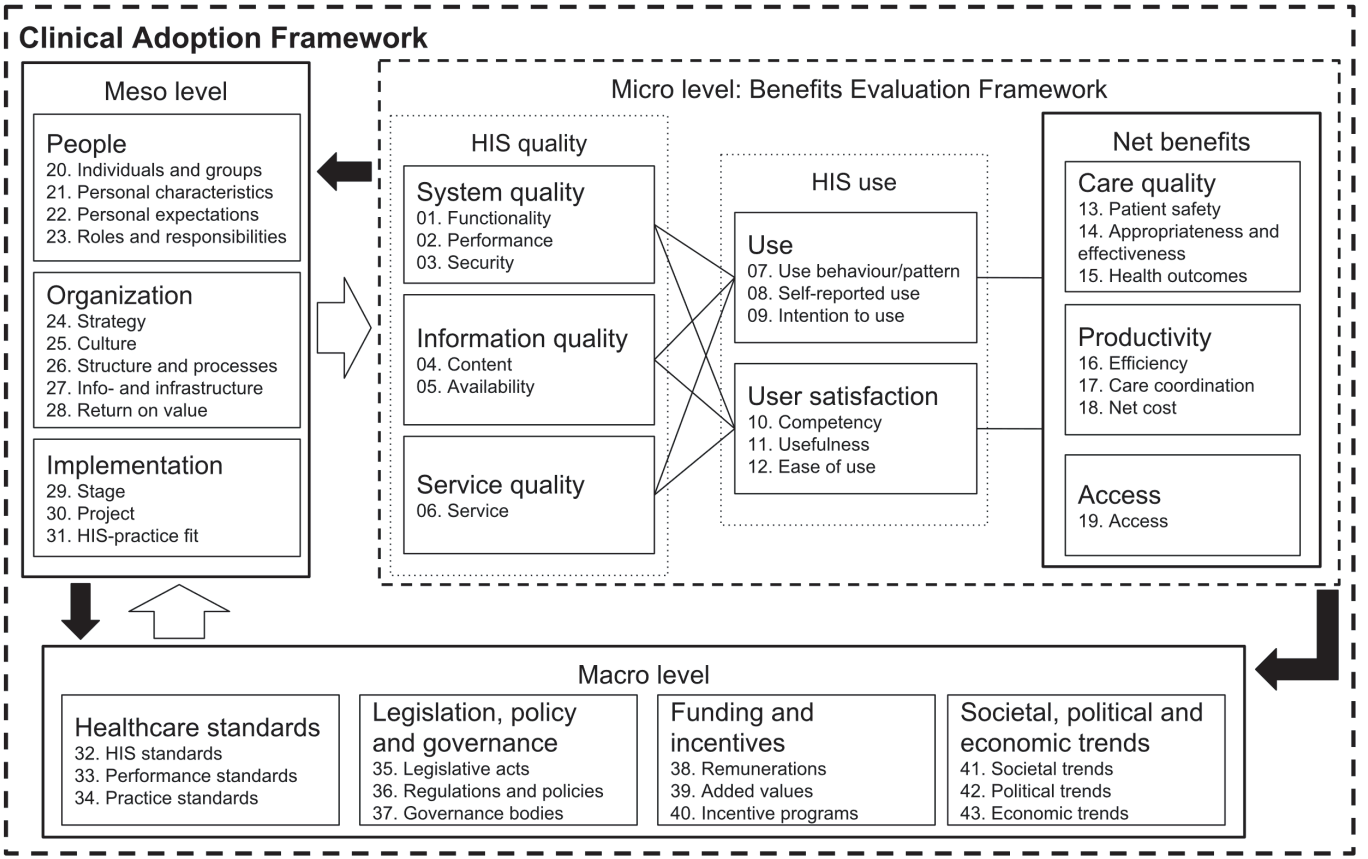


Figure 3.1 Clinical Adoption Framework with levels, dimensions and categories. HIS: health information system. Originally published in [100,157].

The CAF is a complex framework consisting of 43 categories that belong to 15 dimensions (illustrated as boxes in Figure 3.1), which are further separated into the 3 previously mentioned micro, meso, and macro levels.[99] The CAF was considered difficult to apply as there was no guidance documentation with explicit descriptions and rules regarding its use.[158] Consequently, studies[100,162-164] that have applied the CAF differed in their interpretations and applications. Furthermore, HIS adoption increasingly involves sharing medical data with patients and informal caregivers. Therefore, patients and caregivers should also be considered when understanding successful HIS implementation because they might value different factors than health care providers. This patient and caregiver perspective was not explicitly taken into account during the development of the CAF.

The primary objective of this study was to extend the CAF to make it useful for evaluating patients' access to EHRs and the adoption of PHRs. The second objective was to improve the consistent application of the CAF in literature and implementation studies. For this purpose, we aimed to assess the interrater reliability of applying the framework, discuss which areas of the CAF could be improved, and develop guidance documentation.

3.2 Methods

We systematically reviewed existing systematic review papers on determinants and outcomes of patients' access to their personal health data. Results from each review paper were mapped to categories in the CAF, which was adapted when needed to reach consensus. The protocol for this review study was developed using the first 6 review papers[56,68,106,113,117,118], which were the most recent publications at the beginning of this review study. We used 13 subsequent review papers[57,58,69,105,107-112,114-116] in this study to refine the CAF and to assess the interrater reliability. The review protocol was registered at PROSPERO under CRD42018084542.[101] We then reported the results of adapting the CAF, including its reliability, to make it suitable for an evaluation of the adoption of PHRs and patients' access to EHRs. The results of the review study on the determinants and outcomes of patients' access to medical records were reported separately.[157]

To improve the CAF and its definitions, one reviewer (HM) extracted quotes from the literature that described determinants and outcomes for the adoption of EHRs and PHRs, and another reviewer (RD) verified these extracted quotes. The two reviewers independently mapped the extracted quotes. The interrater

reliability for the agreement on the mapping was calculated with Cohen's unweighted Kappa.[165,166] Each quote was mapped to two CAF categories: one for the determinant of the quote and the other for the outcome. Within each category, the quotes were classified into metrics by thematic analysis, as illustrated in Figure 3.2. The metrics and categorizations were iteratively revised to ensure consistency and meaningful categories for summarizing results, which was similar to the process described in Bassi, Lau & Lesperance.[162] The mapping to two categories is visualized in Figure 3.2. For example, in the quote '*Online record access and service users tended to be slightly older (t-test, P<0.001)*'[58], the determinant metric could be "Age" and the outcome metric "Adoption". Age would be classified as "21. Personal characteristics", under the dimension "People" at the meso level, while adoption would be classified as "07. Use behaviour/pattern", under the dimension "Use" at the micro level. For the sake of the review, we added the category '44. Other' to denote when a quote could not be classified using the CAF.

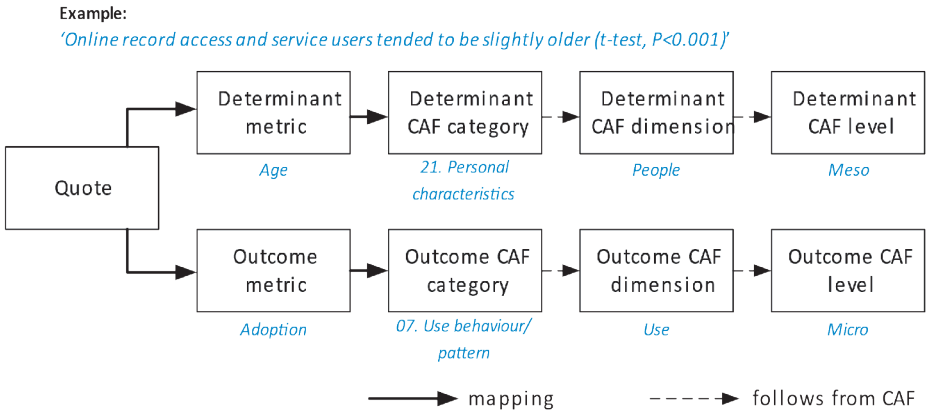


Figure 3.2 Example of how a quote[58] is mapped to a determinant metric and outcome metric. CAF: Clinical Adoption Framework.

The results of this mapping and the differences in quote interpretation and CAF categorization were discussed among the two reviewers to achieve consensus. When necessary to achieve consensus, the definitions of the CAF were adapted and extended with inclusion and exclusion criteria to make them clearer. We presented the number of definitions for categories that were introduced, extended, or unchanged in each level. For agreements and disagreements between reviewers on mapping quotes to categories, we calculated the number of times each unique combination was agreed or disagreed upon (i.e., number of

times there was agreement on one certain category or disagreement between two specific categories). We counted the number of quotes that were classified into each category by a reviewer as well as how many quotes could not be mapped to the CAF. The level of agreement between reviewers on mapping quotes to each category indicated how ambiguous or well-defined the category was. This process resulted in defined categories of the CAF with inclusion and exclusion criteria and a list of metrics that we distinguished. Statistical analysis was carried out in R version 3.5.0 (R Foundation for Statistical Computing, Vienna, Austria) with RStudio 1.1.453 (RStudio Inc, Boston, MA). The R script can be found in Multimedia appendix 1.

3.3 Results

In this section, we first list the definitions that were unchanged, extended or introduced. Second, we discuss the interrater reliability and the spread of mapping quotes to CAF categories.

3.3.1 Adaptation of CAF categories and found metrics

Definitions were introduced to the CAF for the 19 micro level categories, because they were missing in the original publication of the CAF. For example, the category “01. Functionality” of the dimension “System quality” was defined with the inclusion criteria ‘Actual or missing features/functionalities of the HIS and their quality’ and the exclusion criteria ‘If adoption or use of the HIS in general, without a particular functionality, then choose 07. Use behaviour/pattern’. Thus, the exclusion criteria were made explicit for when a quote must be classified in another category. For the 24 meso- and macro level categories, the definitions from Lau, Price, and Kashevjee[99] were used, either unchanged (9 categories) or extended (15 categories), to cover cases of patient and informal caregiver use and disambiguate the categories with refinements and exclusion criteria. For example, the definition for the category “21. Personal characteristics” of the dimension “People” was extended with “socio-economic status, ethnicity, computer skills, (health) literacy, health status” and “Behaviour”. These are factors that were found to be important for the adoption of a HIS by patients and caregivers, and were not included in the original CAF category definition. Table 3.1 shows the numbers of categories for each level and how many were introduced, extended or unchanged. Table 3.2 shows the categories that were changed and provides an example for each level. The resulting definitions for disambiguation for each category are listed in Table 3.3. The metrics of each category can be found in Table 3.4.

Table 3.1 The number of categories with introduced, extended, and unchanged definitions per level

Level	Introduced	Extended	Unchanged	Total
Micro	19	0	0	19
Meso	0	9	3	12
Macro	0	6	6	12

Table 3.2 Categories where inclusion and exclusion criteria were added

Level	Categories changed	Example [additions in brackets]
Micro	All categories from "01. Functionality" to "19. Access"	Inclusion criteria introduced for "01. Functionality": [Actual or missing features/functionalities of the HIS and their quality.] Exclusion criteria introduced for "01. Functionality": [For adoption or use of the HIS in general, not a particular functionality, use category "07. Use behaviour/pattern"]
Meso	"20. Individuals and groups", "21. Personal characteristics", "22. Personal expectations", "23. Roles and responsibilities", "25. Culture", "27. Info- and infrastructure", "28. Return on value", "30. Project", "31. HIS-practice fit"	Inclusion criteria extended for "21. Personal characteristics": "Degree to which an individual's characteristics, such as age, gender, education, [socio-economic status, ethnicity, computer skills, (health) literacy, health status,] experience and expertise can affect the adoption of an HIS"[99]. [Behaviour.]
Macro	"35. Legislative acts", "36. Regulations and policies", "39. Added values", "41. Societal trends", "42. Political trends", "43. Economic trends"	Definition extended with exclusion criteria for "35. Legislative acts": [For privacy concerns use category "22. Personal expectation."]

3.3.2 Interrater reliability and spread

From the 13 reviews[57,58,69,105,107-112,114-116], we extracted 624 quotes. Each of the 624 quotes were mapped twice (i.e., to a determinant and an outcome category) resulting in 1248 mappings. We achieved a percentage agreement of 67.0% (418) and a kappa of 0.58 for the determinant category and a percentage agreement of 62.5% (390) and a Kappa of 0.55 for the outcome category. As shown in Table 3.5 and Table 3.6 in Multimedia appendix 2, the three categories that were least ambiguous based on their high agreement score were "16. Efficiency", "21.

Personal characteristics" and "13. Patient safety". In contrast, categories "09. Intention to use", "04. Content" and "30. Project" showed low agreement scores. Some disagreements between two categories occurred more often than others. For example, a feature relating to secure messaging or access to medical records was interpreted by one reviewer as "01. Functionality" and the other reviewer as "07. Use behaviour/pattern" 94 times. This happened for instance with the quote '*Patients experienced easier communication and interactive discussion with their physician after reading the medical file.*'[57] There was one quote that did not fit into any one of the categories: '*Two articles proposed achieving data exchange by setting up (Regional) Health Information Exchanges that can standardize data and facilitate exchange among different organizations.*'[68] This result referred to infrastructure that exists outside of an organization to facilitate data exchange between organizations and would, therefore, fall into the macro level.

3.4 Discussion

3.4.1 Principal findings

The definitions of the CAF categories were extended to be applicable to patient access to EHRs and the adoption of PHRs. This was achieved by adding factors that were found in the reviewed literature on patient access to EHRs, but were not present in the CAF yet, as was illustrated in the example of "21. Personal characteristics" in the results section. In addition, we developed guidance documentation in the form of inclusion criteria, exclusion criteria, and a list of metrics found. The interrater reliability of the reviewers applying the adapted CAF was moderate. However, we found the CAF to be a highly suitable and comprehensive framework to address patients' access to EHRs, as we could achieve consensus on the mappings through discussion, and almost all results could be categorized in the CAF. The original content for the definitions of the CAF were unchanged and only extended with additional inclusion and exclusion criteria for disambiguation and for the application to patients' access to EHRs. The number of agreements and disagreements and percentage of agreements varied among the CAF categories, just like the number of quotes that were mapped to each category. Some categories were not found at all in the reviews, especially those on the macro level.

3.4.2 Strengths and weaknesses of the study

We showed how the CAF can be applied to studies evaluating patient access to EHRs and PHRs. Despite many publications on the application of the CAF, we are the first, to our knowledge, to provide measures on the interrater reliability.

However, the unweighted Cohen's kappa does not consider that categories actually reflect an order and results within each review are all correlated and come from the same study. Nonetheless, the moderate agreement indicates that the extended CAF is applicable in a consistent way. Because this study was a systematic review of systematic reviews, we have not investigated how to apply these results in primary implementation studies. The categories that were mapped to a lesser extent might have been underdefined, especially those at the macro level. It is possible that these categories may not have been reported in the literature, but also the literature may not have addressed the topics from those categories, or those categories could have been reported in other types of literature such as in policy, law, or grey literature, rather than scientific medical literature. Those categories with relatively high disagreement should also be further evaluated and redefined. Furthermore, the CAF could be used in studies to present their results in a more structured and standardized way. This will improve the ability to compare the results of different studies.

3.4.3 Results in relation to other studies

The variability in the application of the CAF categories found in previous studies[100,162-164] can be explained by ambiguities that were addressed by the inclusion and exclusion criteria of this study. In addition, we found that mapping to a determinant and an outcome CAF category, instead of only one, decreased some of the ambiguity. Only one result, concerning regional information exchange, could not be mapped in the original CAF. This shows that overall the CAF is sufficiently comprehensive. However, we believe that the infrastructure that is available in the environment of an organization forms a missing category in the framework. This category could be introduced in the framework at the macro level to incorporate regional information and communication technologies (ICT) infrastructure, which might be more advanced in some regions than in others.

3.4.4 Implications of the study

This adapted framework can be used in other reviews and in implementation studies of HISs, especially when the HIS has patients and informal caregivers as users. The definitions and metrics provided will still be of value to implementation studies by pointing out several aspects and metrics that have to be considered when carrying out HIS implementations.

Furthermore, the results of this study fulfil part of the need for more guidance documentation when applying the CAF.[158] Our definitions with inclusion and exclusion criteria as well as the metrics found may contribute to a more consistent

application of the framework. We recommend addressing specific relationships between determinants and outcomes using this framework, as we did by mapping quotes from the literature to two CAF categories.

3.4.5 Conclusions

The scope of the CAF was extended to the adoption of PHRs, in addition to EHRs, by health care providers, patients, and informal caregivers. Further definitions and inclusion and exclusion criteria disambiguate and guide the application of each category. We found moderate interrater reliability in applying the framework and variance among the categories in the framework. Future research should address the application of the CAF in primary implementation studies and studies focusing on macro level topics such as international and regional ICT infrastructures, standards, politics, incentive programs, and social trends.

3.5 Appendices

3.5.1 Multimedia appendix 1

Script.R

R script with syntax for analysis is available in the online version of this paper at https://doi.org/10.2196/15150.

3.5.2 Multimedia appendix 2

Supplementary tables.

3.5.2.1 Adapted CAF

The purpose of this table is to classify a quote as determinant or outcome factor within the CAF. To this end, we defined in- and exclusion criteria. Inclusion criteria clarify what factors are included in the particular category and the exclusion criteria contain rules on when to classify a factor into another category. This way categories were disambiguated. The CAF is only used to classify determinants and outcomes of HIS adoption. Therefore, we used the '00. Exclude' category to exclude quotes that were mistakenly taken under review, and were found not to be actual determinant or outcome factors after further evaluation and discussion. Further details on the procedure to apply the adapted CAF to literature studies can be found in the review protocol that was registered at PROSPERO under CRD42018084542.

Table 3.3 The Adapted Clinical Adoption Framework. It consists of levels, dimensions and categories. An "exclusion category" and an "other category" have been added for the mapping. For each category, inclusion and exclusion criteria have been provided. In the category column, the type of change of the in- and exclusion criteria (either introduced, extended or unchanged) is mentioned between brackets. Adapted from Lau (2011)[d]

Id	Level (change)	Dimension	Category (type of change of criteria)	Inclusion criteria	Exclusion criteria
00	Exclude	Exclude	Exclude (introduced)	Not actually a success factor (e.g. determinant, outcome, cost, benefit, barrier, facilitator, problem, solution, advantage, disadvantage, cause, effect, association) of using the HIS. Not actually a result of the review. Or it is unclear what is meant.	
01	Micro	System quality	Functionality (introduced)	Actual or missing features/functionalities of the HIS and their quality.	Adoption or use of the HIS in general, not a particular functionality --> 07. Use behaviour/pattern
02	Micro	System quality	Performance (introduced)	Reliability, downtime, response time, technical issues and performance of the HIS.	Provider or support staff response time --> 06. Service. Internet speed -> 27. Info- and infrastructure.
03	Micro	System quality	Security (introduced)	Security of the HIS, including authentication and authorization performed by the system.	Login (authenticating yourself) -> 01. functionality. Concerns about security or privacy in general by an individual who has an effect on HIS adoption, not related to the usage of the particular HIS or their records in the particular HIS --> 22. Personal expectations. Security measures on organizational level --> 27. Info and infrastructure. Concerns about privacy in general in society; not related to the particular HIS or people who directly affect HIS adoption --> 41. Societal trends.

d DOI: 10.12927/hcq.2011.22157, URL: http://www.longwoods.com/content/22157

Table 3.3 (continued)

Id	Level (change)	Dimension	Category (type of change of criteria)	Inclusion criteria	Exclusion criteria
04	Micro	Information quality	Content (introduced)	Content, completeness, consistency, accuracy, intelligibility of the information in the HIS.	Organizational decisions, discussion or solutions on these issues --> 27. Info and infrastructure.
05	Micro	Information quality	Availability (introduced)	Timeliness, reliability, accessibility of records, and accessibility of the information in the HIS.	Organizational decisions, discussion or solutions on these issues --> 27. Info and infrastructure.
06	Micro	Service quality	Service (introduced)	Technical support for the use of the HIS and clinical, medical support to understand the information in the HIS. Response time of the service (by the healthcare provider or support staff), including response time to secure messaging.	System response time --> 02. Performance. Clinical content of messaging (e.g. clinical advice) --> 14. Appropriateness and effectiveness.
07	Micro	Use	Use behaviour/pattern (introduced)	Actual usage patterns and behaviour of the HIS measured objectively (not reported by the user itself, but e.g. by log files). Adoption of the HIS.	Self-reported use or subjective measures of usage --> 08. Self-reported use. A factor that influenced the usage --> 01-06. HIS quality or 20-43. Determinants. Particular functionality of the HIS --> 01. Functionality.
08	Micro	Use	Self-reported use (introduced)	Self-reported usage patterns and self-reported usage behaviour by the user of the HIS.	Actual usage patterns and behaviour of the HIS measured in another way, or unclear whether it was measured subjectively --> 07. Use behaviour/ pattern. A factor that influenced the usage --> 01-06. HIS quality or 20-43. Determinants.

Table 3.3 (continued)

Id	Level (change)	Dimension	Category (type of change of criteria)	Inclusion criteria	Exclusion criteria
09	Micro	Use	Intention to use (introduced)	The actual intention to use or not to use the HIS or the HIS feature.	Believes and expectations about using HISs in by people who can affect the adoption of an HIS --> 22. Personal expectations.
10	Micro	User satisfaction	Competency (introduced)	Learnability, learning curve, competency and ability to understand and use the system.	Knowledge, experience or skills in general, not related to the particular HIS (implementation) used, e.g. medical knowledge or computer skills --> 21. Personal characteristics. Training provided to use the HIS --> 30. Project. User satisfaction --> 11. Usefulness. Ease of use and usability, Difficulty for the user to use the HIS --> 12. Ease of use. Awareness about HISs in general --> 22. Personal expectations.
11	Micro	User satisfaction	Usefulness (introduced)	User satisfaction about and usefulness of the HIS.	Expectations or user satisfaction about using HISs in general --> 22. Personal expectations. Ease of use --> 12. Ease of use. Degree of fit with daily life of patient or work processes --> 31. HIS practice fit.
12	Micro	User satisfaction	Ease of use (introduced)	Ease of use and usability of the HIS.	Learnability --> 10. Competency. User satisfaction --> 11. Usefulness.
13	Micro	Care Quality	Patient safety (introduced)	Patient safety, adverse events, risks or medical errors as an outcome of the usage of the HIS.	

Table 3.3 (continued)

Id	Level (change)	Dimension	Category (type of change of criteria)	Inclusion criteria	Exclusion criteria
14	Micro	Care Quality	Appropriateness and effectiveness (introduced)	Effectiveness, guideline compliance or treatment adherence, appointment attendance, healthcare utilization as an outcome of the usage of the HIS. Knowledge or awareness about the condition or treatment, preparedness for doctor visits, participation in or engagement with healthcare as an outcome of the usage of the HIS.	Patient-provider engagement, participation or communication, self-management or healthcare accessibility --> 19. Access.
15	Micro	Care Quality	Health outcomes (introduced)	Health outcomes, medical outcomes, physiological outcomes, psychological outcomes and quality of life as an outcome of the usage of the HIS.	Health knowledge or awareness, patient engagement as an outcome of the usage of the HIS --> 14. Appropriateness and effectiveness.
16	Micro	Productivity	Efficiency (introduced)	Efficiency, resource utilization, time needed, workload as an outcome of the usage of the HIS.	As an incentive for the organization that adopts the HIS --> 28. Return on value. As a general expectation on macro level --> 39. Added values. Appointment attendance --> 14. Appropriateness and effectiveness
17	Micro	Productivity	Care coordination (introduced)	Coordination of care, continuity of care, communication between providers, workflow as an outcome of the usage of the HIS.	Patient-provider communication and access to health care --> 19. Access.

Table 3.3 (continued)

Id	Level (change)	Dimension	Category (type of change of criteria)	Inclusion criteria	Exclusion criteria
18	Micro	Productivity	Net cost (introduced)	Financial costs and benefits as an outcome of the usage of the HIS.	Cost/benefit as an incentive/determinant that influences the HIS adoption on organizational level --> 28. Return on value. Remunerations available that influence the meso level (people, organization, project) dimensions that affect clinical adoption --> 38. Remunerations. Incentive programs that influence the meso level (people, organization, project) dimensions that affect clinical adoption --> 40. Incentive programs.
19	Micro	Access	Access (introduced)	Accessibility and availability of healthcare services and medication as an outcome of the usage of the HIS. Patient-provider participation or engagement, self-management, access to personal health data, shared decision making, and patient-provider communication as an outcome of the usage of the HIS.	Accessing information (as an activity) in the HIS by patients --> 07-12. Use or 01. Functionalities if it concerns a specific feature. Internet, computer or smartphone access --> 27. Info and infrastructure. Availability or accessibility of information in the system when quality of information of the HIS --> 05. Availability. Service when quality of the HIS --> 06. Service. Patient participation in or engagement with healthcare, or patient participation/engagement not further specified 14. Appropriateness and effectiveness.

Table 3.3 (continued)

Id	Level (change)	Dimension	Category (type of change of criteria)	Inclusion criteria	Exclusion criteria
20	Meso	People	Individuals and groups (extended)	'Types of individuals/groups that can affect the adoption of an HIS, including patients/clients and families, healthcare providers and managers, policy planners and stakeholder groups'[99]. For each result found we classify the type of individual it concerns in the "subgroup" column.	Characteristics of the individual itself about which the factor is reported --> 21. Personal characteristics.
21	Meso	People	Personal characteristics (extended)	'Degree to which an individual's characteristics, such as age, gender, education, [socio-economic status, ethnicity, computer skills, (health) literacy, health status,] experience and expertise can affect the adoption of an HIS'[99]. Behaviour.	Competency related to HIS use -> 10 Competency.
22	Meso	People	Personal expectations (extended)	'Degree to which an individual [who can affect the adoption of an HIS] believes HISs are important, can improve job performance, [daily life, health status or quality of life] and that infrastructures exist to support its adoption'[99], or expectation about whether they themselves or others will be able or willing to use the system. Also norms and preferences related to HIS use.	User satisfaction about a particular HIS being used --> 11. Usefulness. Expectations about the adoption of HISs of the public in general or people that cannot affect the adoption of an HIS --> 41-43. Societal, political and economic trends. Security of the particular HIS --> 03. Security. 201705: Use intention to use the particular HIS --> 09. Intention to use

Table 3.3 (continued)

Id	Level (change)	Dimension	Category (type of change of criteria)	Inclusion criteria	Exclusion criteria
23	Meso	People	Roles and responsibilities (extended)	'Position, function and obligation of an individual/group in relation to HIS adoption, e.g., being a stakeholder, leader, champion and project sponsor'[99].	Type of individual (e.g. caregiver, provider) --> 20. Individuals and group, fill out in the subgroup column. Characteristics of these individuals --> 21. Personal characteristics.
24	Meso	Organization	Strategy (unchanged)	'Set of coordinated activities designed to achieve the overall mandate and objectives of the organization, including HIS adoption'[99].	
25	Meso	Organization	Culture (extended)	'Ingrained set of shared [norms,] values, beliefs and assumptions acquired by members of an organization over time, including their views toward HISs'[99].	
26	Meso	Organization	Structure and processes (unchanged)	'Organizational functioning, including governance, configuration, reporting relationships and communication, as well as business and patient care processes such as continuity of care'[99].	
27	Meso	Organization	Info- and infrastructure (extended)	'HIS governance/management, technical architectures, information assets, level of integration and privacy/security in place or planned'[99]. Infrastructure in the environment of the consumer (patient or caregiver), including smartphone, internet, computer access.	Belief of an individual that infrastructures exist to support the adoption an HIS --> 22. Personal expectations. Usage of infrastructure --> 21. personal characteristic.

Table 3.3 (continued)

Id	Level (change)	Dimension	Category (type of change of criteria)	Inclusion criteria	Exclusion criteria
28	Meso	Organization	Return on value (extended)	'Economic return on HIS investment and use in terms of cost benefit, effectiveness, utility and avoidance, business case, return on investment, value propositions and benefits realization'[99].	Outcomes of usage of the particular HIS --> 13-19. Net benefits. Costs and benefits as an incentive/ determinant on macro level outside of the organization --> 39. Added values.
29	Meso	Implementation	Stage (unchanged)	'HIS adoption stages – initiation, building/ buying, introduction and adaptation'[99].	
30	Meso	Implementation	Project (extended)	'Planning, activities and resources for HIS adoption, including scope, objectives, constraints, targets, governance, methodology, commitment, communication, training, risks, monitoring, reporting and expectations'[99]. In the maintenance phase also uptake by new patients can be understood as part of the ongoing project implementation.	
31	Meso	Implementation	HIS-practice fit (extended)	'Degree of fit between the HIS and organizational work practices [of the care provider or the daily life of the patient or caregiver], and the extent of change from HIS adoption'[99].	
32	Macro	Healthcare standards	HIS standards (unchanged)	'Types of data, messaging, terminology and technology standards that influence the healthcare industry as a whole with respect to HIS adoption'[99].	

Table 3.3 (continued)

Id	Level (change)	Dimension	Category (type of change of criteria)	Inclusion criteria	Exclusion criteria
33	Macro	Healthcare standards	Performance standards (unchanged)	'Types of organizational performance standards in place, such as those for accreditation of healthcare facilities and performance targets'[99].	
34	Macro	Healthcare standards	Practice standards (unchanged)	'Desired level of professional competency, knowledge, skills and performance in the workplace, including HIS adoption'[99].	
35	Macro	Legislation, policy and governance	Legislative acts (extended)	'Types of HIS-related legislative acts, such as health information and privacy laws that govern the adoption of HISs'[99].	Privacy concerns --> 22. Personal expectation.
36	Macro	Legislation, policy and governance	Regulations and policies (extended)	'Types of HIS-related regulations/policies, such as data access and security/privacy guidelines'[99].	Privacy concerns --> 22. Personal expectation; Organizational security policy --> 27. Info- and infrastructure; System security --> 03. Security. Organizational policy regarding data access --> 26. Structure and process.
37	Macro	Legislation, policy and governance	Governance bodies (unchanged)	'Types of accountability and decision-making structures in place regarding the adoption of HISs'[99].	
38	Macro	Funding and incentives	Remunerations (unchanged)	'Types of compensation available, such as alternative payment schemes to entice change at the individual, practice and organizational levels'[99].	

Table 3.3 (continued)

Id	Level (change)	Dimension	Category (type of change of criteria)	Inclusion criteria	Exclusion criteria
39	Macro	Funding and incentives	Added values (extended)	'General expectations available, such as alternative payment schemes to entice change at the individual, practice and organizational levels'[99].	Outcomes of usage of the particular HIS --> 13-19. Net benefits. Cost/ benefit as an incentive/ determinant on organizational level that influences the HIS adoption --> 28. Return on value.
40	Macro	Funding and incentives	Incentive programs (unchanged)	'Types of reward programs available that entice change at the individual, practice and organizational levels'[99].	
41	Macro	Societal, political and economic trends	Societal trends (extended)	Societal trends or 'General expectations of the public toward healthcare and HISs'[99].	Expectations of particular people or groups who have a direct influence on HIS adoption --> 22. Personal expectations.
42	Macro	Societal, political and economic trends	Political trends (extended)	Political trends or 'General political climates toward healthcare and HISs'[99].	
43	Macro	Societal, political and economic trends	Economic trends (extended)	Economic trends or 'General economic investment climates toward healthcare and HISs'[99].	
44	Other	Other	Other (introduced)	Factor does not fit in CAF, other category proposed for this determinant or outcome of the adoption and use of the HIS.	

3.5.2.2 Metrics per CAF category

The purpose of this table is to provide an overview of metrics found in the literature that belong to each category. This can be used to classify results in literature reviews. It can also be used in implementation studies to help identifying factors and metrics to evaluate. Reporting results from evaluation studies in the CAF will enable comparison among different studies. This table contains the metrics we found from a thematic analysis of the quotes in the reviews and that we classified into each category. Further details on the procedure to apply the adapted CAF to literature studies can be found in the review protocol that was registered at PROSPERO under CRD42018084542.

Table 3.4 Metrics per category with the category Id, level, dimension.

Id	Level	Dimension	Category	Metrics
00	Exclude	Exclude	Exclude	
01	Micro	System quality	Functionality	Access to medical records, Add comments to notes, Adding text to pictures, Alerts, notifications and reminders, Appointment reminders, Appointment scheduling, Assistive technology, Automatic data input, Control over health information, Create own care plan, Data entry, Data exchange, Decision support, Design, Diagnostic tool, Discharge information, E-prescribing, Feature to contact multiple providers, Financial administration, Graphs, Health care services use, Health status management, Help remember daily tasks, Home monitoring, Hospital information, Immunization records, In-patient schedule, Information about daily hospital routine, Information about healthcare team, Information conveyance, Input methods, Integration of PHR with healthcare system, Journal, Lifestyle advice, Medical history, Medical renewal reminders, Medication information, Medication management, Medication renewal, Medication warnings, Medium (e.g. paper, electronic), Missing functionality, More providers in the system, Multimedia, Ordering meals, Patient education, Patient summary, Personalization, Prevention recommendations, Printing, Privacy and security settings, Prognostic tool, Provider emergency access, Proxy access, Record management, Risk assessment tool, Safety information, Secure messaging, Sharing medical information, Symptom checker, Take notes, Technology platform options, Terminology tool, Test results, Tools for health management, Tools to understand health information, Triage of messages, Video calling, Video with use instruction, View billing information, View health information, View lab results, View medication list, View notes, View operative reports, View test results, Voice recording
02	Micro	System quality	Performance	System response time, Technical issues
03	Micro	System quality	Security	Audit trail, Authentication, Authorization, Security
04	Micro	Information quality	Content	Accuracy, Amount of documentation, Coded sensitive information, Consistency, Content, Intelligibility, Language, Not overwhelming, Relatedness of information, Sensitivity, View notes
05	Micro	Information quality	Availability	Availability, Timeliness

Table 3.4 Metrics per category with the category Id, level, dimension. (**continued**)

Id	Level	Dimension	Category	Metrics
06	Micro	Service quality	Service	Clinical support, Service response time, Support, Technical support
07	Micro	Use	Use behaviour/pattern	Adoption, Registration, Use pattern
08	Micro	Use	Self-reported use	
09	Micro	Use	Intention to use	Intention to use
10	Micro	User satisfaction	Competency	Ability to use the system, Learnability, Skills, System feature awareness
11	Micro	User satisfaction	Usefulness	Usefulness, User satisfaction
12	Micro	User satisfaction	Ease of use	Ease of use, Navigation, Task completion
13	Micro	Care Quality	Patient safety	Error identification, Medical error, Medication error, Patient safety
14	Micro	Care Quality	Appropriateness and effectiveness	Appointment attendance, Being prepared for emergencies, Effective consultations, Emergency department visits, Health care services use, Immunisation, Inconvenience, Knowledge and understanding about personal health, Manage administrative concerns, Medication adherence, Medication adjustment, Medication management, Number of appointments, Patient participation in health care, Preventive services use, Quality of care, Satisfaction with care, Treatment adherence
15	Micro	Care Quality	Health outcomes	Confidence, Disease control, Medical outcome, Physiological measures, Psychological outcomes
16	Micro	Productivity	Efficiency	Appointment length, Customer retention, Efficiency, Number of appointments, Number of hospitalisations, Number of messages, Number of phone contacts, Registration time, Repetition of information, Time savings, Workload (outcome)
17	Micro	Productivity	Care coordination	Access to information by providers, Care coordination, Continuity of care, Responsibility for record management, Workflow (outcome)
18	Micro	Productivity	Net cost	Financial benefits

Table 3.4 Metrics per category with the category Id, level, dimension. (**continued**)

Id	Level	Dimension	Category	Metrics
19	Micro	Access	Access	Access to personal health information, Communication between patients and clinicians, Decision making, Focus on patients, Patient-provider participation, Patient-provider relationship, Preparation for clinical visits, Responsibility for record management, Self-management, Sharing personal health information
20	Meso	People	Individuals and groups	
21	Meso	People	Personal characteristics	Activeness in decision making, Address, Admission status, Adoption, Age, Age of caretaker, Autonomy, Awareness, Behaviour, Being sufficiently informed, Cognition, Computer and internet skills, Computer skills, Computer use, Conceptual knowledge, Confidence, Coping style, Education, EHR use, Employment, Engagement with PHRs, Ethnicity, Feeling responsible for their own health, Feeling sufficiently informed, Gender, Health care use, Health literacy, Health status, Insurance, Internet access, Internet use, Keeping their own medical records, Knowledge, Knowledge about right to access medical records, Language, Lifestyle, Literacy, Marital status, Medication use, Memory, Motor skills, Numeracy, Patient activation, Provider use of PHR, Receiving information from other sources, Receptiveness to information, Sensory and perceptual abilities, Skills, Socioeconomic status, Specialty, Technology use, Time constraints, Trust in healthcare provider, Use of alternative technology
22	Meso	People	Personal expectations	Adoption by providers, Alternative technologies, Anxiety to learn, Awareness of PHR, Beliefs, Choice for a provider, Competitiveness, Concerns, Expectations, Fear for anguish through consulting records, Fear of not understanding medical records, Intention to use, Interest in PHRs, Need to keep their own medical records, Needs, Norms, Patient safety, Perceived PHR value, Preferences, Privacy concerns, Skepticism, Willingness to pay
23	Meso	People	Roles and responsibilities	
24	Meso	Organization	Strategy	

Table 3.4 Metrics per category with the category Id, level, dimension. (**continued**)

Id	Level	Dimension	Category	Metrics
25	Meso	Organization	Culture	Norms
26	Meso	Organization	Structure and processes	Care process, Cooperation with healthcare provider, Involvement of healthcare team, Letting patients exchange medical data, Patient-provider relationship, Practice size, Process, Setting, Social environment, Staffing, Time constraints, Workload (determinant)
27	Meso	Organization	Info- and infrastructure	Authentication mechanisms, Authorization policy, Computer access, Data exchange infrastructure, Data governance, Hardware, Infrastructure, Integration of PHR with healthcare system, Internet access, Internet and computer access, Internet speed, Interoperability, Kiosks, Security measures, Smartphone access
28	Meso	Organization	Return on value	PHR value
29	Meso	Implementation	Stage	
30	Meso	Implementation	Project	Communication about PHR, Demonstration, Design decisions, Family recommendation, Participatory design, Provider endorsement, Recommendations, Registration reminders, Training, Trial period, Usability testing
31	Meso	Implementation	HIS-practice fit	HIS-practice fit
32	Macro	Healthcare standards	HIS standards	HIS standards, Interoperability
33	Macro	Healthcare standards	Performance standards	
34	Macro	Healthcare standards	Practice standards	
35	Macro	Legislation, policy and governance	Legislative acts	
36	Macro	Legislation, policy and governance	Regulations and policies	Regulation

Table 3.4 Metrics per category with the category Id, level, dimension. (**continued**)

Id	Level	Dimension	Category	Metrics
37	Macro	Legislation, policy and governance	Governance bodies	
38	Macro	Funding and incentives	Remunerations	Reimbursement
39	Macro	Funding and incentives	Added values	Added value
40	Macro	Funding and incentives	Incentive programs	Incentive programs
41	Macro	Societal, political and economic trends	Societal trends	Education curricula, Societal trends
42	Macro	Societal, political and economic trends	Political trends	
43	Macro	Societal, political and economic trends	Economic trends	
44	Other	Other	Other	Regional health information exchange infrastructure

3.5.2.3 Agreement per CAF category

Table 3.5 Number of times a category was chosen by a reviewer: with category, the total number of times n it was chosen by a reviewer, and the number of times n and percentage % agreement and disagreement

Category	n total	n (%) agree
00. Exclude	22	0 (0)
01. Functionality	376	194 (51.6)
02. Performance	0	
03. Security	1	0 (0)
04. Content	40	10 (25)
05. Availability	7	4 (57.1)
06. Service	17	8 (47.1)
07. Use behaviour/pattern	660	474 (71.8)
08. Self-reported use	3	0 (0)
09. Intention to use	10	2 (20)
10. Competency	50	22 (44)
11. Usefulness	73	52 (71.2)
12. Ease of use	38	22 (57.9)
13. Patient safety	20	16 (80)
14. Appropriateness and effectiveness	248	158 (63.7)
15. Health outcomes	34	10 (29.4)
16. Efficiency	81	70 (86.4)
17. Care coordination	9	4 (44.4)
18. Net cost	1	0 (0)
19. Access	109	58 (53.2)
20. Individuals and groups	0	
21. Personal characteristics	412	350 (85)
22. Personal expectations	208	144 (69.2)
23. Roles and responsibilities	0	
24. Strategy	0	
25. Culture	0	
26. Structure and processes	7	0 (0)
27. Info- and infrastructure	17	6 (35.3)
28. Return on value	1	0 (0)
29. Stage	0	
30. Project	48	12 (25)
31. HIS-practice fit	1	0 (0)

Table 3.5 (continued)

Category	n total	n (%) agree
32. HIS standards	0	
33. Performance standards	0	
34. Practice standards	0	
35. Legislative acts	0	
36. Regulations and policies	0	
37. Governance bodies	0	
38. Remunerations	0	
39. Added values	0	
40. Incentive programs	2	0 (0)
41. Societal trends	0	
42. Political trends	0	
43. Economic trends	0	
44. Other	1	0 (0)

3.5.2.4 Agreement per CAF category sorted by percentage agreement

Table 3.6 Number of times a category was chosen by a reviewer: with category, the total number of times n it was chosen by a reviewer, and the number of times n and percentage % agreement and disagreement. Sorted by descending percentage agreement.

Category	n total	n (%) agree
16. Efficiency	81	70 (86.4)
21. Personal characteristics	412	350 (85)
13. Patient safety	20	16 (80)
07. Use behaviour/pattern	660	474 (71.8)
11. Usefulness	73	52 (71.2)
22. Personal expectations	208	144 (69.2)
14. Appropriateness and effectiveness	248	158 (63.7)
12. Ease of use	38	22 (57.9)
05. Availability	7	4 (57.1)
19. Access	109	58 (53.2)
01. Functionality	376	194 (51.6)
06. Service	17	8 (47.1)
17. Care coordination	9	4 (44.4)
10. Competency	50	22 (44)
27. Info- and infrastructure	17	6 (35.3)
15. Health outcomes	34	10 (29.4)

Table 3.6 (continued)

Category	n total	n (%) agree
04. Content	40	10 (25)
30. Project	48	12 (25)
09. Intention to use	10	2 (20)
00. Exclude	22	0 (0)
03. Security	1	0 (0)
08. Self-reported use	3	0 (0)
18. Net cost	1	0 (0)
26. Structure and processes	7	0 (0)
28. Return on value	1	0 (0)
31. HIS-practice fit	1	0 (0)
40. Incentive programs	2	0 (0)
44. Other	1	0 (0)
02. Performance	0	
20. Individuals and groups	0	
23. Roles and responsibilities	0	
24. Strategy	0	
25. Culture	0	
29. Stage	0	
32. HIS standards	0	
33. Performance standards	0	
34. Practice standards	0	
35. Legislative acts	0	
36. Regulations and policies	0	
37. Governance bodies	0	
38. Remunerations	0	
39. Added values	0	
41. Societal trends	0	
42. Political trends	0	
43. Economic trends	0	


Inflammation of the stomach and intestines due to the coronavirus
mnesis
as been ve ll and suffered from dyspnea and
rgia H stroenteritis due to SARS-CoV-2
cal
simulatio quency 73/minute. QRS and QT/QTc
liac decompensation Complaints
dmill t

4

Evaluation of lexical clarification by patients reading their clinical notes:

A quasi-experimental interview study

Authors:
Hugo J.T. van Mens *
Mirte M. van Eysden *
Remko Nienhuis
Johannes J.M. van Delden
Nicolette F. de Keizer
Ronald Cornet
* Contributed equally

Journal: BMC Medical Informatics and Decision Making.
Year: 2020. **ISSN:** 1472-6947. **Volume:** 20. **Issue:** Suppl 10. **Pages:** 278.
DOI: 10.1186/s12911-020-01286-9.

Abstract

Background Patients benefit from access to their medical records. However, clinical notes and letters are often difficult to comprehend for most lay people. Therefore, functionality was implemented in the patient portal of a Dutch university medical centre (UMC) to clarify medical terms in free-text data. The clarifications consisted of synonyms and definitions from a Dutch medical terminology system. We aimed to evaluate to what extent these lexical clarifications match the information needs of the patients. Secondarily, we evaluated how the clarifications and the functionality could be improved.

Methods We invited participants from the patient panel of the UMC to read their own clinical notes. They marked terms they found difficult and rated the ease of these terms. After the functionality was activated, participants rated the clarifications provided by the functionality, and the functionality itself regarding ease and usefulness. Ratings were on a scale from 0 (very difficult) to 100 (very easy). We calculated the median number of terms not understood per participant, the number of terms with a clarification, the overlap between these numbers (coverage), and the precision and recall.

Results We included 15 participants from the patient panel. They marked a median of 21 (IQR: 19.5 – 31) terms as difficult in their text files, while only a median of 2 (IQR: 1 – 4) of these terms were clarified by the functionality. The median precision was 6.5% (IQR: 2.3% – 14.25%) and the median recall 8.3% (IQR: 4.7% – 13.5%) per participant. However, participants rated the functionality with median ease of 98 (IQR: 93.5 – 99) and a median usefulness of 79 (IQR: 52.5 – 97). Participants found that many easy terms were unnecessarily clarified, that some clarifications were difficult, and that some clarifications contained mistakes.

Conclusions Patients found the functionality easy to use and useful. However, in its current form it only helped patients to understand few terms they did not understand, patients found some clarifications to be difficult, and some to be incorrect. This shows that lexical clarification is feasible even when limited terms are available, but needs further development to fully use its potential.

4.1 Background

Patient access to electronic health records (EHRs) is facilitated by patient portals and personal health records. Patients benefit from reading their clinical notes, as it helps them to remember more from what was discussed during consultations and supports them to take care of themselves.[167-170] However, medical data and jargon are difficult to comprehend for most people without a medical background.[4,56,106,107,171] Previous research on lexical simplification (replacing difficult terms with easier terms) and lexical clarification (providing synonyms and definitions to terms) has shown that minimising medical jargon and providing clarifications in medical records may increase comprehension.[172-175] Lexical clarification works similar to infobuttons that are inserted into the EHR to provide additional information.[176,177] However, rather than retrieving external information resources to aid decision-making, our work is aimed at patients, clarifying medical terminology with a short textual explanation or definition. Therefore, the Dutch university medical centre UMC Utrecht developed functionality in its patient portal to help clarify medical concepts in free-text data sources, such as discharge letters. The functionality used synonyms and definitions from a Dutch thesaurus for care and wellbeing (in Dutch: "Thesaurus Zorg & Welzijn"). Nonetheless, this thesaurus has not been tailored yet to low literacy levels and is not developed to clarify medical concepts to laymen. Previously the terminology had been used as a thesaurus for search functionalities on healthcare websites. Previous research has evaluated what difficulties patients experience when reading their medical records[4,171], but has not assessed functionality that provides clarifications of difficult terms to patients personal medical records.

We aimed to evaluate to what extent the lexical clarifications match the information needs of patients. First, we assessed whether the right terms were explained, i.e. terms that patients considered difficult. Second, we evaluated whether the terms were explained in the right way, and third, we evaluated how the clarifications and functionality can be improved.

4.2 Study context

The study was carried out at UMC Utrecht, the Netherlands. The functionality was developed by the university hospital[e] itself and implemented in the hospital-wide

e During copy-editing 'university hospital' was replaced by 'university'. For clarity, in this thesis we used the original formulation.

patient portal in January 2019. The functionality matches free text with terms and synonyms from the thesaurus by text matching and provides a preferred synonym with a definition as a clarification for the matched term. Abbreviations were excluded. The functionality underlines terms that could be matched to the thesaurus, which users can click to view a pop-up window with the clarification. The functionality was activated for treatment reports, medical letters, and test results. No formal evaluation had yet been carried out.

4.3 Methods

4.3.1 Study design

We carried out an exploratory quasi-experimental before and after interview study. Participants were first asked to read their notes without the functionality and then again with the functionality activated.

4.3.2 Participants

Participants were invited through the patient panel of the hospital, which included 80 patients willing to be contacted for research on diverse topics related to the quality of care. We included a convenience sample of the first fifteen positive respondents for a 1.5-hour interview. The participants received reimbursement for their travel expenses and a gift voucher of 20 euro for their participation.

4.3.3 Study flow

The interviews were carried out in October and November 2019. The test environment and acceptance environment of the patient portal were used for the study, the first without the functionality, the second with the functionality. During the interviews the participants were asked to read free-text notes from their own EHR aloud. We included medical correspondence between clinicians, discharge summaries, and treatment reports less than one year old and routinely available in the patient portal. We excluded notes that were addressed to the participant. Test results were excluded as well, because we did not want to potentially confront the participants with unfamiliar test results. We asked the participants to mark the terms not understood or for which they wished to see an explanation during reading, which we denote as “difficult terms” hereafter. Then, the participants rated the ease of these terms. Next, the functionality was activated and we asked the participants to read the letter again. For each clarification, participants were asked about their thoughts on the clarification and how it could be improved, and to rate the ease and usefulness of the clarification. The terms not marked as difficult, that did get a clarification we denominate as “easy terms” throughout

the text. Furthermore, we asked the participants to provide feedback on the functionality, and to rate the ease and usefulness of the functionality. Finally, we removed directly identifying data, such as years and names of the patient or clinicians, and stored the letters including the participants' terms selection and ratings for further analyses.

4.3.4 Outcome measures

We collected the following background data from the participants: gender, age, education level, treatment duration in the UMC, whether they had work experience in healthcare, and their health literacy using the validated Dutch version of the Set of Brief Screening Questions (SBSQ)[178]. The primary outcome measure was the number of terms that the participants deemed difficult and that were provided with a clarification by the functionality. The secondary outcome measures were the usefulness of clarifications of the difficult terms compared to the easy terms, the ease and usefulness of the clarification functionality, and the feedback the users provided on the clarifications and functionality. Measurements were carried out on a 100 mm visual analogue scale (VAS, from 0 to 100) and collected with background data on paper case report forms.

4.3.5 Methods for data acquisition and measurement

The pseudonymized notes were stored. Interviews were audio-recorded, transcribed, and pseudonymized. After the interviews the quantitative data were entered into the electronic data-capture system Castor EDC v2019.3.10 (Ciwit B.V., Amsterdam, The Netherlands).

4.3.6 Methods for data analysis

We reported the numbers and percentages of the participant characteristics. We calculated the precision and recall of the functionality for each participant. Precision in this study context was defined as the number of difficult terms clarified by the functionality divided by the number of clarifications provided. Recall was defined as the number of difficult terms clarified divided by the number of difficult terms. For each participant we calculated the median number of difficult terms, clarifications provided by the functionality, difficult terms clarified, and the VAS score of the ease and usefulness of the terms, of the clarifications, and of the functionality. We calculated the median and interquartile range (IQR) of the medians per participant. Statistical analysis was carried out in R version 3.6.1 (R Foundation for Statistical Computing, Vienna, Austria) with RStudio 1.2.1335 (RStudio Inc., Boston, MA, USA). The script can be found in Additional file 1.

4.4 Results

4.4.1 Demographic and other study coverage data

Table 4.1 lists the characteristics of the fifteen participants. Participants had a median age of 57 (ranging from 34 to 70), eight participants had received higher education, seven were treated at the UMC for more than ten years, and eight had worked in healthcare in the past. None of the participants had inadequate health literacy. Participants read a median of 2 (IQR: 2 – 3) letters during the interviews, with a median of 214 (IQR: 144 – 395) words per letter, including fourteen outpatient clinic letters, two discharge summaries, and sixteen treatment reports. The letters covered a wide range of medical specialties: angiology, cardiology, dermatology, dietetics, endocrinology, gynaecology, infectiology, nephrology, neurology, nursing, oncology, ophthalmology, physiatry, pulmonology, surgery, and urology. More detailed data can be found in Additional file 2.

Table 4.1 Participant characteristics with the statistic, category, number n and percentage.

Statistic	Category	n	(%)
Gender	Male	8	(53)
	Female	7	(47)
	Other	0	(0)
Age group	0 - 18 years	0	(0)
	19 - 29 years	0	(0)
	30 - 39 years	1	(7)
	40 - 49 years	0	(0)
	50 - 59 years	8	(53)
	60 - 69 years	4	(27)
	70 - 79 years	2	(13)
	≥ 80	0	(0)
Education	No education	0	(0)
	Elementary school	0	(0)
	Lower secondary education	2	(13)
	Preparatory vocational secondary education	2	(13)
	Vocational education and training	3	(20)
	Senior general or university preparatory secondary education	0	(0)
	Higher professional education	7	(47)
	Research-oriented higher education	1	(7)

Table 4.1 (continued)

Statistic	Category	n	(%)
Treatment duration	< 3 years	3	(20)
	3 - 10 years	5	(33)
	> 10 years	7	(47)
Works in healthcare	Yes, currently	0	(0)
	Yes, in the past	8	(53)
	Never	7	(47)
Health literacy	Inadequate (SBSQ <= 2)	0	(0)
	Adequate (SBSQ > 2)	15	(100)

4.4.2 Study findings and outcome data

Participants marked a median of 21 (IQR: 19.5 – 31) terms in their notes as difficult during the interviews. The functionality provided clarifications for a median of 26 (IQR: 22 – 44) terms per participant, and a median of 2 (IQR: 1 – 4) of these clarifications was provided to terms that participants had also marked as difficult. The median precision per participant was 6.5% (IQR: 2.3% – 14.25%) and the median recall per participant 8.3% (IQR: 4.7% – 13.5%). Two participants did not find any of the terms they deemed difficult clarified. See Figure 4.1.

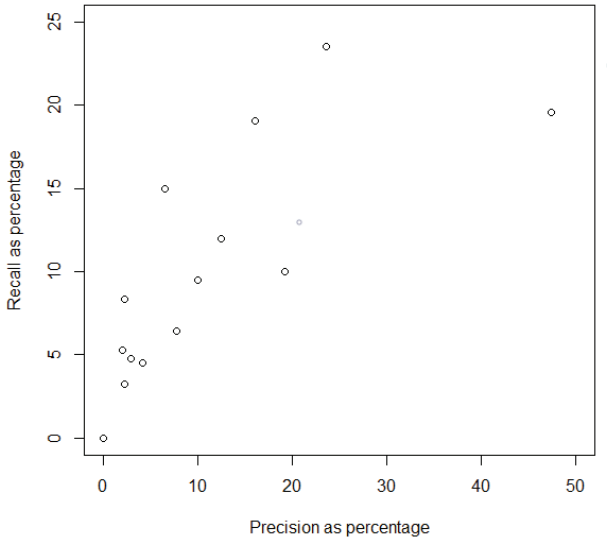


Figure 4.1 Precision and recall per participant. Two participants had zero difficult terms that were clarified and thus the precision and recall was zero in these cases.

Participants rated difficult terms with a median ease of 8.5 (IQR: 6.5 – 20.75), from 0, very difficult, to 100, very easy, and easy terms with a median ease of 99 (IQR: 97 – 100). Difficult term clarifications were rated a median ease of 93 (IQR: 72 – 96) and easy term clarifications 96 (IQR: 94.25 – 99.5). See Figure 4.2 and Figure 4.3. Difficult term clarifications were rated with a median usefulness per participant of 90 (IQR: 76 – 97), from 0, not useful at all, to 100, very useful, while clarifications of easy terms were rated with a usefulness of 4 (IQR: 1.5 – 18.75). See Figure 4.4. The density plot in Figure 4.5 shows the distribution of overall ratings of the usefulness of clarifications. Participants mostly rated clarifications of easy terms as not useful, because they were not necessary for them personally. However, in some cases, participants thought the clarification was useful somehow anyway because it provided new information and the participants could provide feedback on the clarifications of terms they already knew. They rated difficult term clarifications as useful when it helped them understand the term, but not when the clarification itself was too difficult or incorrect. For example, in a cardiological context, the plaque of blood vessels was clarified with dental plaque. Examples of difficult terms, the most common terms that were clarified, and errors in clarifications are listed in Table 4.2, Table 4.3, Table 4.4 respectively. For detailed data, see Additional file 2 and Additional file 3. For translations of the examples from Dutch, see Additional file 4.

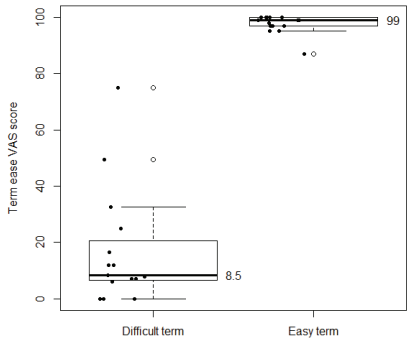


Figure 4.2 Median ease (from very easy to very difficult) per participant of difficult terms compared to easy terms.

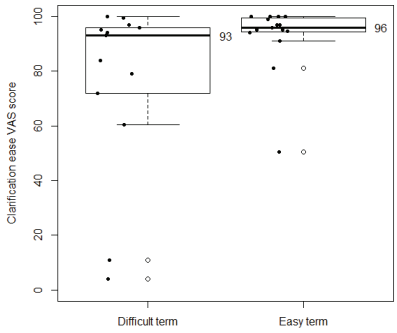


Figure 4.3 Median ease per participant (from very easy to very difficult) of clarifications of difficult terms compared to clarifications of easy terms.

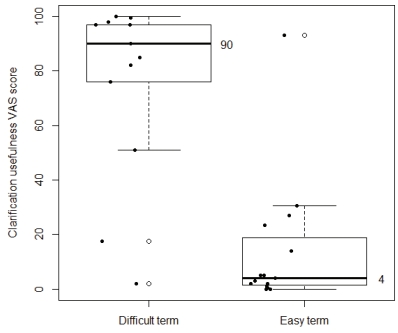


Figure 4.4 Median usefulness (from not useful at all to very useful) of clarifications of difficult terms compared to clarifications of easy terms.

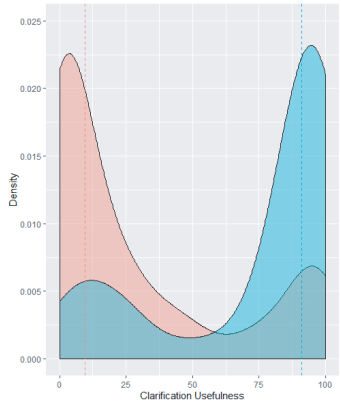


Figure 4.5 Density plot of usefulness (from not useful at all to very useful) ratings for all terms.

Table 4.2 Terms marked as difficult by two or more participants, with n as the number of participants that encountered the term

Difficult term	n
eGFR (CKD-EPI)	3
CNS	2
Endocrinology	2
HNP	2
Immune serology	2
Proximal	2
RR	2

Table 4.3 Most common terms clarified by the functionality, with n as the number of participants that encountered the term

Clarification	n
Outpatient clinic	13
Anamnesis	9
Medicine	8
Endocrinology	7
Physical examination	7

Table 4.4 Examples of problems found with some clarifications

Problem with clarification	Example term	Clarification provided to example
Unnecessary	Belly	**Belly** Part of the trunk between the midriff and the pelvis
Too difficult	Intoxication	**Poisoning** Distortion of the life functions by a too high concentration of a certain substance in the body
Circular	Neurologists	**Neurologists** Medical specialists who are specialized in neurology
Homonym (context was about plaque in blood vessels)	Plaque	**Plaque** White, sticky substance on the teeth and molars in which may occur living and dead bacteria, released tissue cells and food scraps
Related term	Peristaltic	**Digestive system** Process by which food taken in by the mouth can be made ready for absorption in the blood and the residual products are excreted and the food is then digested

Most participants found the functionality easy to use, with a median ease of 98 (IQR: 93.5 – 99). Two outliers found the functionality not easy or difficult (scores 40 and 50). One of these participants commented it was not clear that the terms were underlined at first and the other found they had to scroll as the clarifications sometimes appeared outside of the window. We observed both issues with other participants as well. The majority of participants found the functionality to be useful, with a median of 79 (IQR: 52.5 – 97), even though participants reported that most clarifications were not useful and the coverage was very low. In general, participants commented that the functionality was fast, easy, inviting to click, well-designed, added value, and liked that it allowed you to do something with the notes, and that one could choose to click or not. They did not like that misspelled words were not taken into account and found a lack of consistency, experienced anxiety, would not use the functionality, thought too many words were underlined, or did not like the design. Participants suggested to add links to further information on the UMC website, enable asking questions, make clarifications more personalized, make the colour of the underlining clearer, and to add more clarifications.

4.5 Discussion

The functionality demonstrated a low precision and recall, which indicates that it does not match the information needs of the patients. However, the patients found the clarifications of the terms they considered difficult to be useful, with

some reservations for incorrect and difficult clarifications. Overall, most patients considered the functionality to be easy to use and useful. We observed variance among patients in precision, recall, ease, and usefulness.

The patients were not fully representative for patient portal users in general, as they were actively involved in the patient panel, half had worked in healthcare before, and none of them had inadequate health literacy. We expect the precision and recall to be higher for patients with lower health literacy, and for patients who are still unfamiliar with the topic of their disease and treatment. However, the actively engaged patients from the sample were relatively more knowledgeable about their own health status, and were hence more critical about the functionality. Therefore, the patients were already familiar with many of the terms the functionality clarified, that other persons might not have known, and could provide feedback for improvement from their personal experience and knowledge.

Provider notes are among the most difficult sections of medical records [171]. We have not measured whether the functionality improved the comprehension of patients, but this first requires a further increase of the recall and quality of the clarifications. A strong point of our study is that we read medical correspondence from personal EHRs of the patients. Earlier studies did not use the records from patients themselves [173-175] and have not reported the precision and recall of the functionality that was evaluated. It can be expected that they had a similarly low performance that varied among different patients and notes, and that the increase in comprehension might be lower, when these studies would have used the actual records from patients themselves.

The variance observed between patients is due to multiple factors. On the one hand this includes the (health) literacy of the patient, and his or her familiarity with medical terminology. On the other hand, this might vary according to the medical specialty, writing style of the clinician, and type of free-text source (i.e. treatment reports or medical letters). Further research should address how clarifications can be tailored to the literacy of patients, and how different types of free-text sources can be improved. For example, parts of the free text originate from coded and structured fields in the medical record, such as lab tests and diagnoses, but have lost their underlying coding by being converted to text. It will be easier to clarify the coded data rather than free text, because it is less ambiguous. The difficulty of some clarifications and feedback provided by the patients indicates that the definitions from the thesaurus have a high level of reading difficulty. We

thus recommend to make the definitions easier to read. Rather than providing definitions that unambiguously define concepts like dictionaries do, a terminology for lexical clarification should provide explanations of the terms that clarify the meaning to the reader in an appropriate reading level [175]. Further research should therefore address tailoring the definitions to patients' health literacy levels. Additionally, evaluation studies on lexical clarification functionality should assess the precision and recall of their solutions for different users.

Our results show that in spite of the low recall patients found the clarifications and functionality useful. This is promising for smaller languages where little content for consumer health vocabularies is available. It indicates that it is possible to develop functionality for lexical clarification, starting with a small set of terms and basic text-matching functionality, and to improve it gradually. The results were reported to the developers of the thesaurus and the functionality and will be used for further improvement. This process needs to address the trade-off between introducing more clarifications and having less unnecessary clarifications. More clarifications might increase the recall and usefulness, but will also decrease the precision and may increase the number of incorrect clarifications. For example, many of the unknown terms were abbreviations, which are difficult to disambiguate, even for clinicians. More advanced techniques from natural language processing are required in order to resolve these challenges, that take the context and semantics into account.

4.6 Conclusion

The lexical clarification functionality helped patients to understand terms they did not understand, although the coverage was low. Patients found some clarifications to be difficult or incorrect. Despite low coverage and some problems with available clarifications, patients still found the functionality easy to use and useful. This shows that lexical clarification is feasible and of added value even with limited terminology and coverage. However, incorrect clarifications should be limited to prevent confusion and anxiety.

4.7 Declarations

4.7.1 Acknowledgements

The authors would like to thank David Jongsma, Annelies Hetharia, Linda Minnen, Christina de Bie, Rebecca Holman, Marije Wolvers, and Rudy Scholte for their support in carrying out the study.

4.7.2 Ethics approval and consent to participate

A waiver from the Medical Research Ethics Committee of UMC Utrecht was obtained and filed under reference number WAG/mb/19/033611. It confirmed that the Medical Research Involving Human Subjects Act (in Dutch: WMO) does not apply to the study and that therefore an official approval of this study by the ethics committee was not required under Dutch law. An approval from the quality assurance board of UMC Utrecht was obtained regarding the research protocol, invitation letter, consent forms, research collaboration agreement, and data management plan. Participants provided written informed consent to participate in this study.

4.7.3 Availability of data and materials

The aggregate data are available in the attachments of the manuscript. The Dutch transcripts of the interviews, medical letters, paper case report forms, and castor data are available from the corresponding author on reasonable request. The Dutch interview guide can be obtained from https://purl.org/hjtvanmens/lexicalclarification/interviewguide.

4.8 Appendices

4.8.1 Additional file 1

The R-script used for the analysis is available in the online version of this paper at https://doi.org/10.1186/s12911-020-01286-9.

4.8.2 Additional file 2

The spreadsheet with detailed statistics from the R-script output is available in the online version of this paper at https://doi.org/10.1186/s12911-020-01286-9.

4.8.3 Additional file 3

Additional figures with term and clarification ease and usefulness on term level and clarification ease and usefulness per participant.

4.8.3.1 Term and clarification ease and usefulness on term level

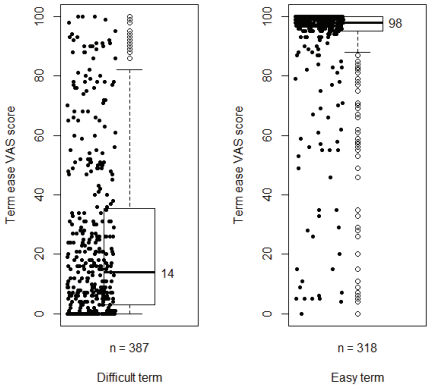


Figure 4.6 Term ease of difficult terms compared to terms not found difficult (Easy term).

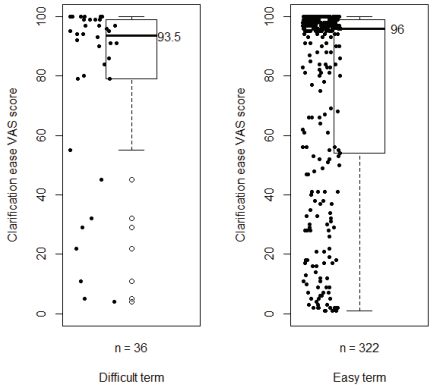


Figure 4.7 Clarification ease of terms found difficult compared to terms not found difficult.

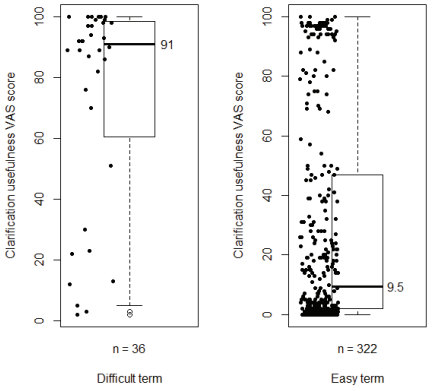


Figure 4.8 Clarification usefulness of terms found difficult (Difficult term) compared to not found difficult (Easy term).

4.8.3.2 Clarification ease and usefulness boxplots per patient

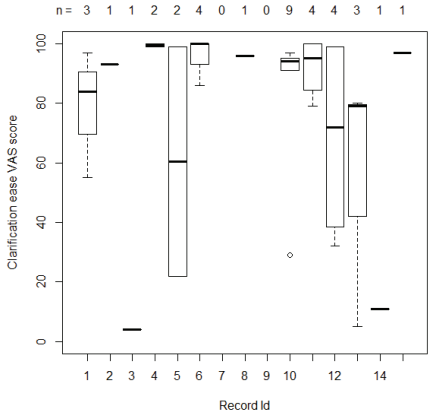


Figure 4.9 Clarification ease of terms found difficult per participant.

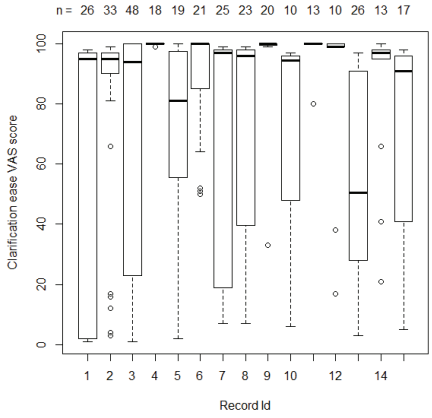


Figure 4.10 Clarification ease of terms not found difficult per participant.

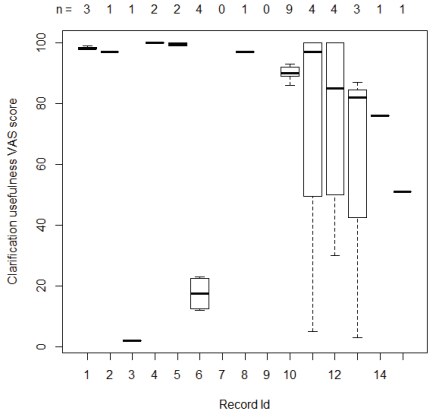


Figure 4.11 Clarification usefulness of terms found difficult per participant.

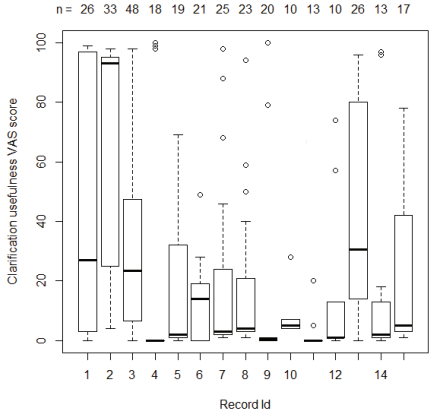


Figure 4.12 Clarification usefulness of terms not found difficult per participant.

4.8.4 Additional file 4

Additional tables with the original Dutch and translated examples of terms marked difficult (Table 4.2 in the manuscript), most common terms clarified (Table 4.3 in the manuscript) and problems found (Table 4.4 in the manuscript).

Table 4.5 Terms marked as difficult by two or more participants with original Dutch terms

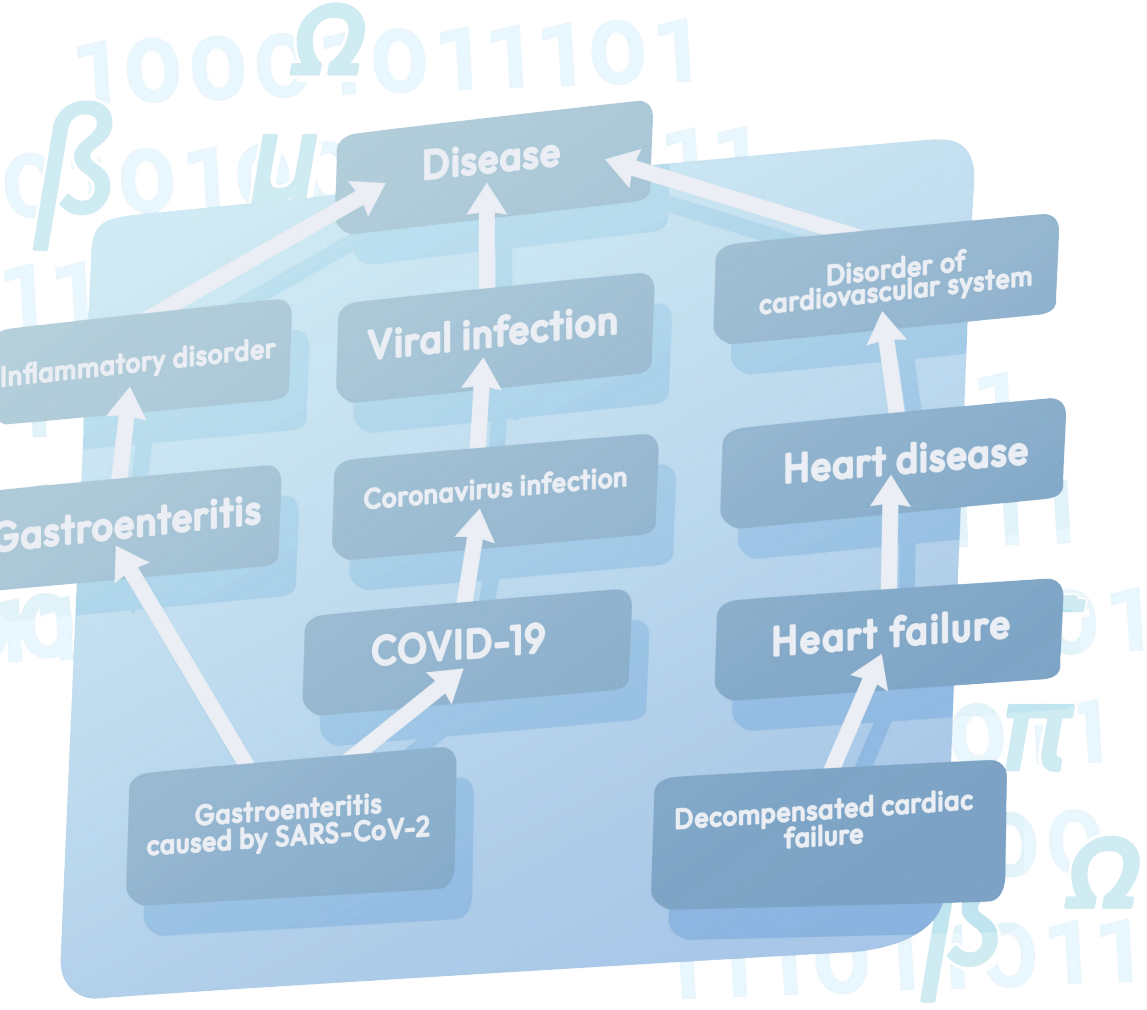
Difficult term English	Difficult term Dutch	n
eGFR (CKD-EPI)	eGFR (CKD-EPI)	3
CNS	CNS	2
Endocrinology	Endocrinologie	2
HNP	HNP	2
Immune serology	Immuunserologie	2
Proximal	Proximaal	2
RR	RR	2

Table 4.6 Most common terms clarified with original Dutch terms

English	Dutch
Outpatient clinic	Polikliniek
Anamnesis	Anamnese
Medicine	Medicijnen
Endocrinology	Endocrinologie
Physical examination	Lichamelijk onderzoek

Table 4.7 Examples of problems found with some clarifications with original Dutch terms and clarifications

English term	Dutch term	English clarification	Dutch clarification
Belly	Buik	**Belly** Part of the trunk between the midriff and the pelvis	**Buik** Deel van romp tussen middenrif en bekken
Intoxication	Intoxicatie	**Poisoning** Distortion of the life functions by a too high concentration of a certain substance in the body	**Vergiftiging** Verstoring van de levensfuncties door een te hoge concentratie van een bepaalde stof in het lichaam
Neurologists	Neurologen	**Neurologists** Medical specialists who are specialized in neurology	**Neurologen** Medisch specialisten die gespecialiseerd zijn in de neurologie
Plaque	Plaque	**Plaque** White, sticky substance on the teeth and molars in which may occur living and dead bacteria, released tissue cells and food scraps	**Plaque** Wittige, kleverige substantie op de tanden en kiezen waarin levende en dode bacteriën, losgelaten weefselcellen en voedselresten voorkomen
Peristaltic	Peristalstiek	**Digestive system** Process by which food taken in by the mouth can be made ready for absorption in the blood and the residual products are excreted and the food is then digested	**Spijsvertering** Proces waarmee door de mond opgenomen voedsel geschikt wordt gemaakt voor opname in het bloed en de restproducten worden uitgescheiden- het voedsel is dan verteerd


Disease
Inflammatory disorder
Viral infection
Disorder of cardiovascular system
Gastroenteritis
Coronavirus infection
Heart disease
COVID-19
Heart failure
Gastroenteritis caused by SARS-CoV-2
Decompensated cardiac failure

5

Clarifying diagnoses to laymen by employing the SNOMED CT hierarchy

Authors:
Hugo J.T. van Mens
Nicolette F. de Keizer
Remko Nienhuis
Ronald Cornet

Journal: Studies in Health Technology and Informatics.
Year: 2018. **ISSN:** 0926-9630. **Volume:** 247. **Pages:** 900 – 904.
DOI: 10.3233/978-1-61499-852-5-900.

Abstract

Patient access to electronic health records (EHRs) is associated with improved efficiency, self-management, and patient engagement. However, the EHR contains medical language that can be difficult to comprehend by patients. In Dutch hospitals, the Diagnosethesaurus (DT) is used as an interface terminology to register diagnoses, but it does not contain patient-friendly terms. Fortunately, the DT is partly mapped to SNOMED CT and there is a proportionately small set of patient-friendly terms available in the Dutch SNOMED CT release. The purpose of this study was, therefore, to investigate if SNOMED CT can be used to generate clarifications of diagnoses for patients. Only 1.2% of the DT diagnoses that were already mapped to SNOMED CT had patient-friendly synonyms that were different from the diagnoses descriptions. However, by generalizing diagnoses to SNOMED CT concepts with patient-friendly terms, this number could be increased to 71%. In conclusion, we showed that a high percentage of diagnoses could be clarified to at least some extent with the relatively small set of patient-friendly terms. Future research will involve the further optimization of the clarifications, and evaluation with clinicians and patients.

5.1 Introduction

There is an increased attention to patient access to electronic health records (EHRs), because of associated benefits such as improved patient satisfaction, patient-provider communication, patient engagement, self-management, efficiency, and patient safety.[58,105] However, EHRs contain medical language that can be difficult to comprehend by patients. [56,106,107] A patient-friendly terminology could help patients to better understand their medical records. [106,107] Zeng et al. [175,179] showed that the Unified Medical Language System (UMLS) and the Consumer Health Vocabulary (CHV) could be used to find synonyms, and generate explanations of medical concepts for laymen. However, the UMLS is not available for the Dutch medical language. Fortunately, the National Release Center of SNOMED CT in the Netherlands published the SNOMED CT Patient-Friendly Extension Release [180] (PFE) with 301 patient-friendly terms to describe 288 SNOMED CT concepts. This small initial set was based on the terms that can be found on a website of the Dutch patient federation where patient experiences about healthcare provider can be shared. [75] SNOMED CT and the PFE can be used to clarify diagnoses in the same way as the UMLS and the CHV. Furthermore, the Diagnosethesaurus (DT) [181] is becoming the standard interface terminology to register diagnoses at the point of care in Dutch hospitals, and this interface terminology is currently partly mapped to SNOMED CT (the full mapping was not yet finished at the time of writing). Therefore this provides opportunities to make descriptions of diagnoses more comprehensible for patients.

A diagnosis can be clarified by providing a patient-friendly synonym or by generalizing it to one or more generic concepts that can be described by a patient-friendly term. [175,179] Table 5.1 illustrates these clarifications methods. For example, “agranulocytosis is a type of immune system disorder” is a generalization of a diagnosis to a more general concept, where the diagnosis is called the subtype and the more general concept the supertype. UMLS and SNOMED CT both contain hierarchical relationships that can be used for this purpose. There are also other relationships in these terminology systems that can be used to clarify medical concepts, such as the “finding site” and “part of” relationship. This enables clarifications such as “aortic valve is a part of the heart”. These latter relationships were also used in the studies with the UMLS, [175,179] but were out of the scope of the current study, where we focused on the hierarchical subtype-supertype relationship.

The purpose of this study was thus to investigate if the PFE and the SNOMED CT hierarchy can be used to generate Dutch clarifications of diagnoses for patients.

Table 5.1 Three methods to clarify diagnoses illustrated by an example (freely translated)

Clarification method	Example of the clarification method
Synonym	Trigeminal neuralgia is another word for facial pain.
Generalization to one concept	Agranulocytosis is an immune system disorder.
Generalization to multiple concepts	Papillon-Lefèvre syndrome is a heritable disorder of bone, tooth, and skin.

5.2 Methods

We combined the DT (version of June 2017) with the SNOMED CT Netherlands edition and the PFE (version of March 2017) and investigated how many of the diagnoses can be clarified with how many patient-friendly terms. When a diagnosis was mapped to a SNOMED CT concept that had a patient-friendly term in the PFE, this term was provided as the synonym of the diagnosis. We verified whether this term was actually different from the diagnosis description, using text matching.

The SNOMED CT hierarchy was used to determine if a diagnosis is a subtype of one or more concepts that can be described by a patient-friendly term (see Figure 5.1). In order to find the supertypes of a concept in SNOMED CT, the transitive closure table was used. [182] This table contains all subtype-supertype relationships. This way all supertypes with patient-friendly terms could be found for each diagnosis. However, this might provide redundant explanatory terms, because supertypes of the supertypes would be found as well. For example, “Aortic valve stenosis”, is a “Heart valve disorder”, “Heart disease”, and “Disorder of cardiovascular system”. In these cases, we only used the most specific generalizations. This would simplify the clarification to “Aortic valve stenosis is a heart valve disorder”. In case a diagnosis had multiple supertypes, as is illustrated with “Telangiectasia macularis eruptiva perstans” which is an “immune system disorder”, “skin disorder” as well as a “disorder of cardiovascular system”, the most specific generalizations were used in combination. We counted the number of unique combinations of these generalizations, i.e., the number of unique clarifications that could be generated with subtype-supertype relationships.

When a supertype was contained in the diagnosis description, it was not regarded to add any new information. For example, that “Salmonella infection” is an

"infectious disease" is already implied by the word "infection" in "Salmonella infection". For this reason, we disregarded supertypes that are not informative and filtered them out by text matching, e.g. matching on "infect" to match both "infection" as well as "infectious" when the supertype was "infectious disease".

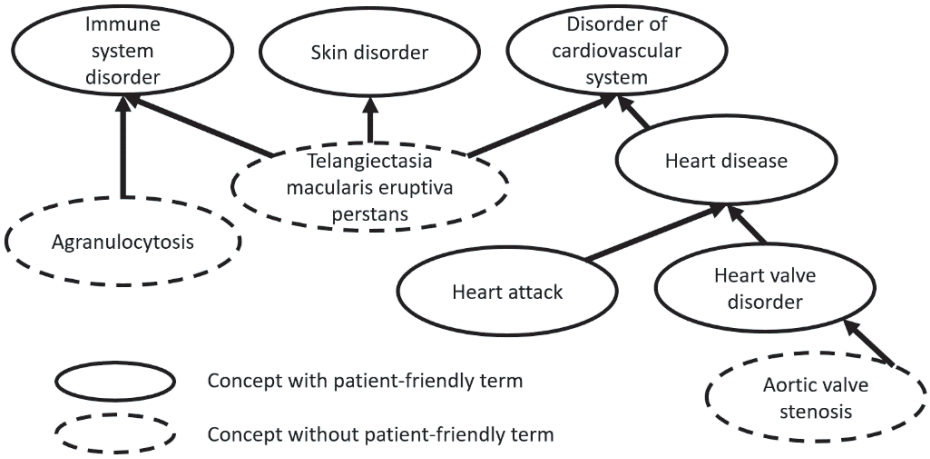


Figure 5.1 Simplified example of SNOMED CT concepts in the SNOMED CT hierarchy

5.3 Results

Of the 21,426 diagnoses in the DT, only 12,453 diagnoses (58.1%) had already been mapped to SNOMED CT. In total 288 SNOMED CT concepts had a patient-friendly term in the PFE, and these described 225 of the diagnoses with a SNOMED CT mapping. 75 diagnoses descriptions were already equal to the patient-friendly terms, hence, the diagnosis description was actually different from the patient-friendly term for 150 diagnoses. Therefore the patient-friendly terms could be used to clarify 1.20% of all diagnoses that have a SNOMED CT mapping with a synonym.

By generalizing diagnoses to supertypes with patient-friendly terms, 8,797 diagnoses (70.6% of all that are mapped to SNOMED CT) could potentially be clarified with 211 concepts that have patient-friendly terms. These diagnoses were described by 1 to 6 different supertypes; as shown in Table 5.2. The 211 supertypes formed 735 unique combinations of supertypes that can be used for clarifications.

There were 110 diagnoses with a patient-friendly term that could also be clarified using supertypes. Analogously, 128 patient-friendly terms were also used as a

supertype. See Figure 5.2. The total numbers are thus 40 + 110 + 8,687 = 8,837 diagnoses and 22 + 128 + 83 = 233 concepts with patient-friendly terms. As a result, 8,837 diagnoses (71.0% of all mapped diagnoses) could be clarified with a synonym and/or supertypes using 233 concepts with patient-friendly terms and 735 unique combinations these terms.

Table 5.2 Number of diagnoses that have a certain number of patient-friendly supertypes to clarify the diagnoses (excluding the redundant ones) and the number of unique clarifications with a certain number of supertypes

Supertypes	Diagnoses	Unique clarifications
1	6,329	192
2	2,050	376
3	366	132
4	49	32
5	2	2
6	1	1
Total	**8,797**	**735**

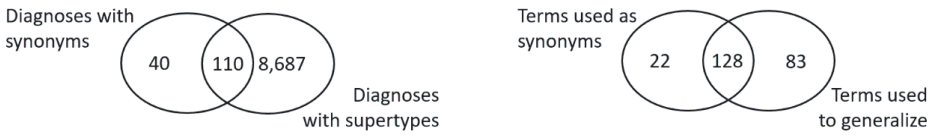


Figure 5.2 The overlap between diagnoses that can be clarified with synonyms and supertypes, and the overlap between patient-friendly terms used as synonyms or to generalize, and their numbers.

5.4 Discussion

With the relatively small set of SNOMED CT concepts that have one or more patient-friendly terms in the PFE, a high percentage of diagnoses could be clarified to at least some extent, by using hierarchical subtype-supertype relationships of SNOMED CT, and by providing synonyms.

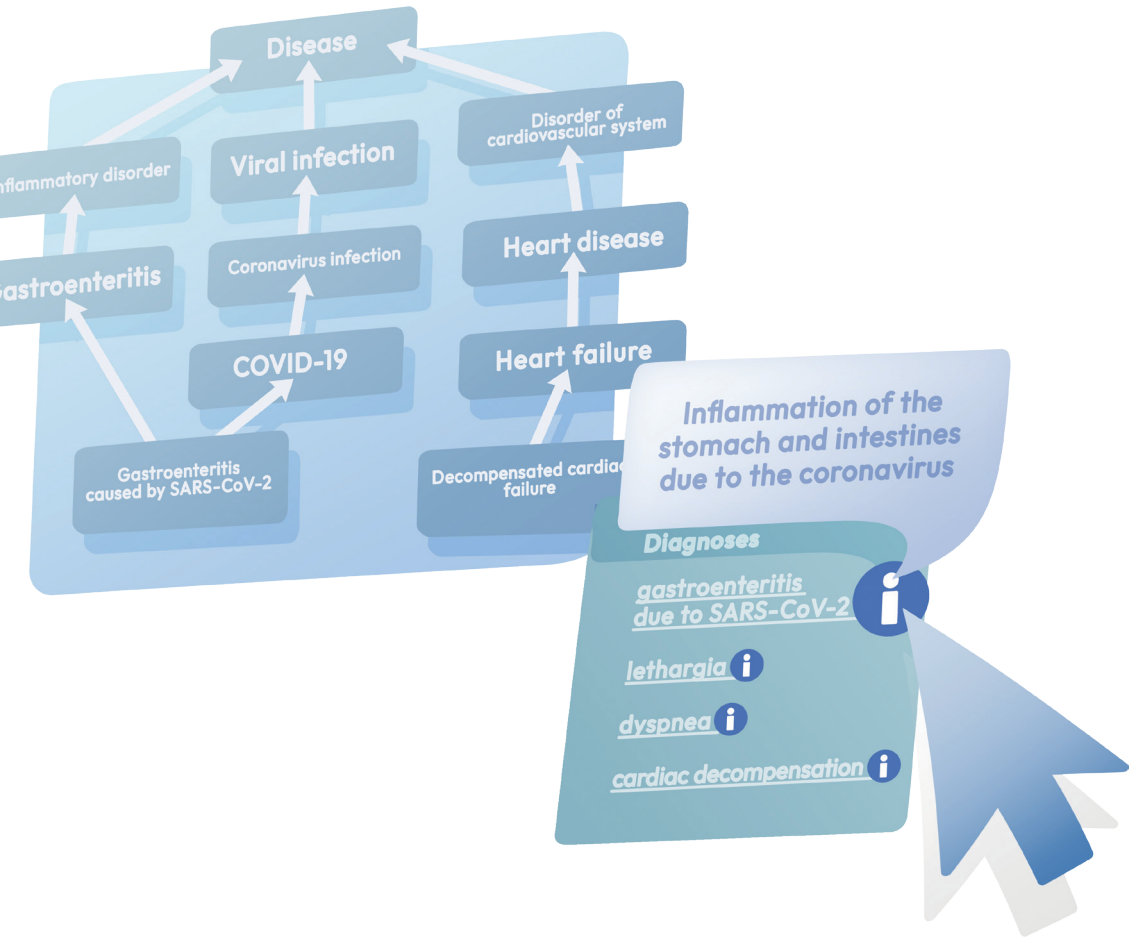
The DT was not completely mapped to SNOMED CT yet, but this is expected to be completed in coming releases. [181] We will have to repeat the analysis from this paper after this is finished. The earlier studies [175,179] did not mention how much content of the terminology systems used could be clarified with their system, which makes it difficult to compare their results with those of our study.

The PFE is quite small compared to the number of diagnoses in the DT, and, particularly, compared to the number of concepts in SNOMED CT. Perhaps some concepts that could be informative to patients, or could clarify the 29% of the diagnoses that were not clarified by the method presented in this study, could relatively easily be translated to patient-friendly terms to obtain even better results. However, finding patient-friendly synonyms for *all* diagnoses and other types of medical concepts is a costly process and might not be possible for some complex diagnoses. Compared to English language, only limited tools are available to perform natural language processing for the Dutch language as well as many other languages. [183] Therefore our result is of particular interest, showing that text simplification might be feasible with a relatively small set of patient-friendly terms. We believe that this method could be applied to other languages.

Whether the clarifications will increase the comprehension of patients and other users, such as caregivers, remains to be evaluated. The clarifications are not yet validated by clinicians, although it can be assumed the hierarchical relationships are correct, because SNOMED CT is clinically validated. We want to first improve the clarification method by utilizing other types of relationships in SNOMED CT, such as the "finding site" relationship or "associated morphology" relationship. We expect this might be more useful for certain diagnoses, such as malignant neoplasms and infections, where the clarification could be e.g. that it is a form of cancer (associated morphology) with a certain finding site. The method presented in this paper can also be used to determine which concepts could be translated to patient-friendly terms to result in a maximum increase of the number of concepts that could potentially be clarified.

5.5 Conclusion

We showed that a relatively small patient-friendly terminology and the SNOMED CT hierarchy can be used to generate clarifications of a large proportion of diagnoses for patients. Future research will involve the further optimization of the clarifications, the utilization of other types of relationships, and evaluation with clinicians and patients. Additionally, research should focus on which parts of SNOMED CT could be translated to patient-friendly terms in order to clarify the highest number of concepts in the most comprehensive manner.


Disease
nflammatory disorder
Viral infection
Disorder of cardiovascular system
astroenteritis
Coronavirus infection
Heart disease
COVID-19
Heart failure
Gastroenteritis caused by SARS-CoV-2
Decompensated cardiac failure
Inflammation of the stomach and intestines due to the coronavirus
Diagnoses
gastroenteritis due to SARS-CoV-2
lethargia
dyspnea
cardiac decompensation

6

Diagnosis clarification by generalization to patient-friendly terms and definitions:
Validation study

Authors:
Hugo J.T. van Mens
Savine S.M. Martens
Elisabeth H.M. Paiman
Alexander C. Mertens
Remko Nienhuis
Nicolette F. de Keizer
Ronald Cornet

Journal: Journal of Biomedical Informatics.
Year: 2022. **ISSN:** 1532-0464. **Volume:** 127. **Pages:** 104071.
DOI: 10.1016/j.jbi.2022.104071.

Abstract

Background Now that patients increasingly get access to their healthcare records, its contents require clarification. The use of patient-friendly terms and definitions can help patients and their significant others understand their medical data. However, it is costly to make patient-friendly descriptions for the myriad of terms used in the medical domain. Furthermore, a description in more general terms, leaving out some of the details, might already be sufficient for a layperson. We developed an algorithm that employs the SNOMED CT hierarchy to generalize diagnoses to a limited set of concepts with patient-friendly terms for this purpose. However, generalization essentially implies loss of detail and might result in errors, hence these generalizations remain to be validated by clinicians. We aim to assess the medical validity of diagnosis clarification by generalization to concepts with patient-friendly terms and definitions in SNOMED CT. Furthermore, we aim to identify the characteristics that render clarifications invalid.

Results Two raters identified errors in 12.7% (95% confidence interval – CI: 10.7-14.6%) of a random sample of 1,131 clarifications and they considered 14.3% (CI: 12.3-16.4%) of clarifications to be unacceptable to show to a patient. The intraclass correlation coefficient of the interrater reliability was 0.34 for correctness and 0.43 for acceptability. Errors were mostly related to the patient-friendly terms and definitions used in the clarifications themselves, but also to terminology mappings, terminology modelling, and the clarification algorithm. Clarifications considered to be most unacceptable were those that provide wrong information and might cause unnecessary worry.

Conclusions We have identified problems in generalizing diagnoses to concepts with patient-friendly terms. Diagnosis generalization can be used to create a large amount of correct and acceptable clarifications, reusing patient-friendly terms and definitions across many medical concepts. However, the correctness and acceptability have a strong dependency on terminology mappings and modelling quality, as well as the quality of the terms and definitions themselves. Therefore, validation and quality improvement are required to prevent incorrect and unacceptable clarifications, before using the generalizations in practice.

6.1 Background

6.1.1 Scientific background

Patients can increasingly access their healthcare data in patient portals and personal health records. However, these records are full of jargon with which most persons without a medical background are unfamiliar.[4,56] Therefore, patient-friendly terms and definitions might help patients and their significant others understand the medical concepts in their records.[172-175] There are, however, myriads of terms that are used in the medical domain. SNOMED CT, a comprehensive medical terminology system in healthcare, contains terms for 352,568 concepts (January 2020 release). It is thus not practical, and very costly, to manually develop patient-friendly terms and definitions for each concept. Furthermore, specialists use very detailed terms that a general practitioner might not use. Laypersons may also require fewer details than specialists, general practitioners, or chronic patients. Hence, a higher level, general description might already be sufficient to clarify the meaning to a layperson.

In a former study, we showed that more than 70% of 12,453 diagnoses in the Dutch Diagnosethesaurus ("Diagnosis Thesaurus" in English, referred to as DT hereafter) can be generalized to a small set of 211 concepts with patient-friendly synonyms from the SNOMED CT Netherlands Patient-Friendly Extension (PFE), by employing the SNOMED CT "is-a" hierarchy.[184] The PFE language reference set 'states which descriptions are appropriate to show to patients, caregivers and other stakeholders who have not received care-related training' (definition of concept '15551000146102 |Patient-friendly Dutch language reference set|'. For example, 'reactive perforating collagenosis' can be generalized to the concept of 'skin disorder', which we consider more comprehensible to laypersons. Similarly, there were 1316 different types of other skin disorders in the DT that can be clarified in this manner. The PFE also contains text definitions that can be used to further explain the diagnoses with a textual description. For example, 'familial periodic paralysis' can be generalized to 'myopathy'. Myopathy is not a common term, but the definition of myopathy can be provided to further explain the concept. See Table 6.1 for examples of diagnosis clarifications based on SNOMED CT.

Table 6.1 Examples of types of clarifications with supertypes, synonyms and definitions. The supertypes are bolded.

Clarification type	Example diagnosis	Example clarification
Synonym and definition	Phlebitis	Another word for "phlebitis" is inflammation of the vein: Inflammation of a vein that makes it red, swollen, and painful.
Definition	Blepharospasm	This involves the involuntary contraction of the eyelids as a cause of a disorder in muscle tension.
Supertype synonym	Reactive perforating collagenosis	A **skin disorder**.
Supertype synonym and definition	Familial periodic paralysis	A type of **myopathy**. Myopathy: This is an abnormality in the structure of the muscle tissue.
Multiple supertypes	High altitude retinopathy	A type of **disorder of retina** and **barotrauma**. Barotrauma: Damage or pain, mainly in the middle ear, because of changes in pressure.
Multiple supertypes with text filter	Congenital cyst of adrenal gland	A type of **inborn abnormality** and **hormonal disorder**. **Cyst**: cavities in the body filled with liquid.
Multiple supertypes with seven supertypes	Lowe syndrome	A type of **inborn abnormality**, **mental disorder**, and disorder of **brain**, **kidney**, **eye**, and **metabolism**. It is **hereditary**.

Previous efforts had focused on developing and enriching lists of patient-friendly terms, by text mining consumer vocabulary and using frequency measures and term similarity to find easier terms. [185-189] However, few studies have focused on generating explanations. They were primarily focused on increasing coverage for lexical simplification. Zeng et al[190] used relationships in the UMLS to generate an explanation, such as 'a type of lung disease' as the explanation of 'pulmonary emboli', derived from the "is-a" relationships. They found some problems with hierarchical relationships, for instance, that a cyst should not be generalized to a type of tumour, because it 'may falsely alarm or unnecessarily alienate the reader'. In a later study elaborating upon these results, Kandula et al[175] reported 32% of hierarchical explanations to be incorrect and therefore evaluated an alternative method, using other types of semantic relationships from the UMLS, such as the 'part of' and 'a device used in' relationship. In a similar application that linked medical texts to existing lay synonyms and definitions, Chen et al[191] found generalizing drugs to higher level drug class definitions was sometimes not accurate to define a drug, because drugs may be used in several treatments that were not always applicable to subclasses. However, to our knowledge, no extensive evaluations have been published about the issues involved in

generalization to concepts with patient-friendly terms and definitions. Moreover, these prior studies were primarily focused at lexical simplification of texts rather than clarification of concepts with generated clarifications.

6.1.2 Rationale for the study

Generalization essentially implies a loss of detail, and thus a loss of information. The technique we have developed of generalization to concepts with patient-friendly terms and definitions can also lead to errors. For example, errors arise if there are mistakes in the mapping of patient-friendly synonyms and text definitions to SNOMED CT concepts, the mapping of diagnoses from the DT to SNOMED CT, or in SNOMED CT "is-a" relationships. Our technique has not been validated yet by medical doctors and terminologists. Systematic validation is necessary before implementing the generalizations in patient portals and personal health records. It can also shed light on for which terms, concepts and clarifications this technique works and for which it does not, and how it can be improved.

6.1.3 Objectives of the study

We aim to assess the medical validity of diagnosis clarification by generalization to supertype concepts with patient-friendly terms and definitions in SNOMED CT. Furthermore, we want to identify the characteristics of the clarifications that render them invalid.

6.2 Study context

6.2.1 Terminology systems

We used the SNOMED CT Netherlands Edition Release (SCT-NL) together with the PFE from March 2020 (SNOMED CT Netherlands National Release Center, Nictiz, The Hague, The Netherlands; referred to as NRC hereafter), and the DT version 3.10 from April 2020 (DHD, Utrecht, The Netherlands). The SCT-NL and PFE are dependent upon the SNOMED CT International Release from January 2020. Table 6.2 shows the number of synonyms, definitions, and diagnoses in these terminologies. Only 714 diagnoses in the DT had synonyms or definitions in the PFE.

Table 6.2 Terminology statistics: the number of definitions, preferred and acceptable synonyms in the SNOMED CT International Release (SCT-Int) and SCT Netherlands Release (SCT-NL), SNOMED CT Netherlands Patient-Friendly Extension Release (PFE), the total number of diagnoses and the number of diagnoses with terms or definitions in PFE from the Diagnosethesaurus (DT).

Terminology	Statistic	n
SCT-Int	Active concepts	352,568
SCT-NL	Active concepts	361,835
PFE	Preferred synonyms	1,409
PFE	Acceptable synonyms	138
PFE	Definitions	700
DT	Diagnoses	24,966
DT	Diagnoses with SNOMED CT mapping	18,579
DT	Diagnoses with PFE term	714
DT	Diagnoses with PFE definition	361

6.3 System details

For diagnoses from the DT that did not have a definition in the PFE, we generated clarifications by generalizing them to supertypes in SNOMED CT with a patient-friendly synonym and definition. See Figure 6.1 for an illustration of the algorithm. Examples of clarifications are provided in Table 6.1. To find the supertypes, we traversed the SNOMED CT "is-a" hierarchy.[184] Because SNOMED CT is polyhierarchical, this method can produce generalizations to a single supertype (e.g., 'Reactive perforating collagenosis' in Table 6.1) or to multiple supertypes (e.g., 'High altitude retinopathy' in Table 6.1). The algorithm removed all redundant supertypes that were also supertypes of other supertypes of the concept to be clarified. For example, disease is a supertype of cardiovascular disorder, which is a supertype of heart disorder, which is a supertype of heart valve disorder, which is a supertype of aortic valve stenosis. All these supertypes have PFE terms. The clarification of aortic valve stenosis, however, only included the direct supertypes (that it is a heart valve disorder), leaving out the redundant supertypes of supertypes. The synonyms and definitions of the supertypes of the concept were concatenated into a short description that served as the clarification. When available, the clarification consisted of the synonym, an enumeration of supertypes and an enumeration of the definitions of the supertypes. For diagnoses with multiple supertypes that were disorders of body parts, the body parts were aggregated to remove redundancy (e.g., 'Lowe syndrome' in Table 6.1). If a supertype term was already part of the medical description, the algorithm

did not use the term as a clarification, but did add its definition (e.g., 'Congenital cyst of adrenal gland' in Table 6.1). We developed an ASP Core .NET (Microsoft, Redmond, WA, USA) web application with a Neo4j graph database (Neo4j, San Mateo, CA, USA) to generate the clarifications.

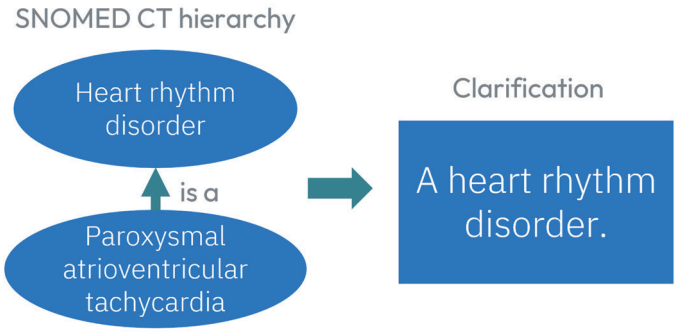


Figure 6.1 The algorithm generates a clarification by generalizing a medical concept to a more general, supertype concept with a patient-friendly description from the PFE

6.4 Materials and methods

6.4.1 Study design

We performed a validation study with a randomized sample of diagnosis clarifications that was evaluated by two raters (SMa and EP). The raters have a medical background and had been involved in the translation of SNOMED CT concepts to Dutch for the SCT-NL release, at the NRC.

6.4.2 Study flow

6.4.2.1 Validation instrument development

We developed a validation instrument to evaluate the quality of the diagnosis clarifications. The validation questions were developed in three iterations. Firstly, with two medical doctors (MDs) other than the raters in the current study, we tested on 45 cases whether the instructions and validation questions were clear. Secondly, with one of the MDs other than the raters of the current study, we measured the intra-rater reliability. In this second iteration, the MD completed twelve cases after receiving the full instructions. Subsequently, he completed the same cases again a few days later. The MDs and the first author (HvM) discussed the differences in ratings and HvM improved the validation questions accordingly to disambiguate them. Additionally, the validation questions were discussed and refined after the first 60 cases with the two raters in the current study. Three cases were repeated in these first 60 cases to assess the intra-rater reliability. We

sampled them randomly from the first 30 cases. We included the three repeated cases again and repeated six extra cases in the final 1140 cases (1131 unique cases).

6.4.2.2 Sampling and validation

We took a random sample of 1200 clarifications by supertypes. This was constrained by the estimated amount of time of 40 hours of validation work made available by the NRC. For each patient-friendly term, we randomly selected one subtype that only had that concept as supertype in its clarification (when available), and one subtype that also had other supertypes that were used in the clarification (when available). In other words, in this way, we could validate all concepts with patient-friendly terms or definitions, both as a single supertype in the clarification, as well as one of multiple supertypes (see Table 6.1). To get a representative set of combinations of supertypes, we sampled by unique combinations of supertypes, covering up to one-third of all unique combinations of supertypes. Additionally, we sampled from concepts with patient-friendly synonyms that had no patient-friendly definitions, to see whether clarifications with supertypes can be useful for these cases. Two raters evaluated the clinical validity and provided feedback from the perspective of how they would explain the diagnosis to a patient. The raters were first provided with instructions on carrying out the validation and were provided with six example clarifications and how to validate them. They then validated 60 clarifications independently and discussed the results with the first author (HvM) to clarify the approach of the study and to achieve consensus. Where necessary, the validation questions were further refined and made explicit. After that, the raters continued with the rest of the clarifications independently. Both raters validated all of the clarifications. Finally, we discussed the clarifications in general. See Figure 6.2 for the steps in the study flow, Appendix A for the sampling script, Appendix B for the sampling procedure, Appendix C for the rating guide including the validation questions, and Appendix D for the data dictionary.

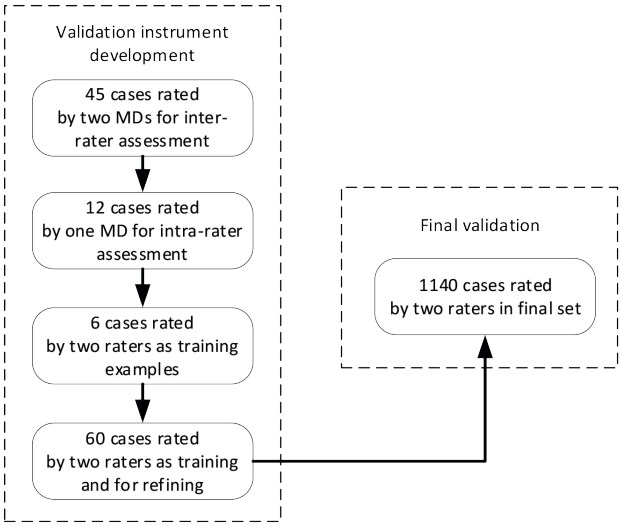


Figure 6.2 Study flow

6.4.3 Outcome measures and evaluation criteria

The primary outcome measures were the percentage of clarifications with identified errors and the percentage of clarifications that were considered unacceptable by the raters. Secondary outcomes were the completeness, relevance, and clarity of the clarifications and the interrater reliability of correctness and acceptability. See Table 6.3 for the description and scale of the outcome measures. The acceptability was based on the errors, completeness, relevance and clarity. A clarification should not contain errors (correctness), generalization should not leave out important information (completeness), the information in the clarification should be relevant (relevance) and it should be clear (clarity). The raters motivated their ratings in their own words in free text. Clarifications could be considered unacceptable without containing any mistakes, for example, if they provided irrelevant information. Conversely, mistakes could be considered acceptable, if they were considered minor issues to patients. Additionally, we categorized problems found and feedback for improvement.

Table 6.3 Outcome measures with a description and the scale used

Outcome measure	Description	Scale
Error	Whether the clarification of the diagnosis contains any mistakes.	Boolean (0 or 1)
Completeness	The extent to which important information is present that you would use to provide a short clarification of the diagnosis.	Likert (1 – 5)
Relevance	The extent to which the information that is provided in the clarification is relevant to clarify a diagnosis to a patient.	Likert (1 – 5)
Clarity	The extent to which the clarification is clear, not vague, not confusing and not difficult to understand the diagnosis.	Likert (1 – 5)
Acceptability	The extent to which it is acceptable to show the clarification to a patient in the diagnosis list of his or her patient portal.	Likert (1 – 5)

6.4.4 Data entry forms

A case report form developed by SMe used for the mapping of code systems to SNOMED CT was extended and used to validate the diagnoses. It was developed using Python (Python Software Foundation, Delaware, DE, USA), Django (Django Software Foundation, Atlanta, GA, USA), and PostgreSQL (The PostgreSQL Global Development Group, Berkeley, CA, USA). The results were exported into a spreadsheet.

6.4.5 Methods for data analysis

6.4.5.1 Quantitative analysis

The intra-rater and interrater reliability were calculated using a two-way, agreement, single-measures intraclass correlation coefficient (ICC)[192]. We calculated the ICCs based on the Boolean and Likert rating scales used to score the outcome measures, to assess the agreement between the raters on those scales. The percentages of clarifications that were considered to contain errors and to be unacceptable were reported with their 95-percent confidence intervals (CI), similarly for their completeness, relevance, and clarity. We dichotomized the Likert scales taking a score of 1 or 2 as falsified (unacceptable, incomplete, irrelevant or unclear, respectively). We carried out a subgroup analysis by testing whether there were significant differences between clarifications that had single or multiple supertypes and compared clarifications with patient-friendly synonyms with those that did not have a patient-friendly synonym. We used the CIs and the Mann-Whitney test to determine whether the differences were significant. For this quantitative analysis, we excluded the first 60 cases, because they were still used to refine the validation instrument. Statistical data analysis was carried out in R (R Foundation for Statistical Computing, Vienna,

Austria), using RStudio (RStudio Inc., Boston, MA, USA). The script can be found in Appendix E.

6.4.5.2 Qualitative analysis

Problems found regarding correctness, completeness, relevance, clarity, and acceptability, as well as suggestions for improvement were thematically analysed, categorized and summarized narratively by the first author (HvM). A qualitative subgroup analysis of clarifications was carried out by analysing patterns in the correctness, completeness, relevance, clarity, and acceptability. For this qualitative analysis, we included all cases (also the first 60). The qualitatively derived themes were based on the feedback provided during the validation. The themes were discussed with the raters to achieve consensus.

6.5 Results

6.5.1 Sampling

We generated clarifications with 2,690 unique combinations of 620 supertypes for the 16,124 diagnoses from the DT for which a SNOMED CT mapping was available. From these diagnoses, we randomly sampled 1,188 cases for validation (7.4% of all cases) with 1,103 unique combinations of supertypes (41.0% of all unique combinations). With this sample, we covered all unique 620 supertypes, both as single (531) and multiple (512) supertypes. See Table 6.4 for the inclusion reasons and types of cases from the random sample.

Table 6.4 Inclusion reasons and the number of included cases n with totals m per inclusion type, for clarifications consisting of a single supertype, those consisting out of multiple supertypes, and those with synonyms. If there was only one clarification with a particular supertype (only single) we took that one, if there were more than one, we sampled one of them. We also took extra samples to have a total of 1,188 cases.

Inclusion reason	n	Inclusion type	m
Only single	115	Total single	543
Sampled single	416		
Extra single	12		
Only multiple	108	Total multiple	522
Sampled multiple	260		
Sampled unique	144		
Extra multiple	10		
Sampled synonyms	123	Total synonyms	123
Total			**1,188**

6.5.2 Ratings

6.5.2.1 Errors

The raters identified errors in 12.7% (CI: 10.7-14.6%) of the clarifications. The raters agreed on the correctness of 90.7% of the cases and the ICC of the interrater reliability was 0.34 for correctness. Most errors were related to the PFE synonyms (36.8%) and definitions (51.3%), but also some were related to the mappings between DT and SNOMED CT (5.9%) and the modelling of SNOMED CT (16.5%), or were inherent to the algorithm (9.2%). Splitting into supertypes by the algorithm introduced errors with metastases and sometimes the "type of"-explanation was considered to lead to erroneous clarifications. The categorization of errors with examples is listed in Table 6.6 and Appendix F.

6.5.2.2 Acceptability

The raters found 14.3% (CI: 12.3-16.4%) of the clarifications to be unacceptable for a patient portal diagnosis list (defined as one or both raters providing a score of 1 or 2 for acceptability). The raters agreed on 48.3% of the Likert scale acceptability ratings with an ICC of 0.43 and 87.9% of the dichotomized acceptability. Table 6.5 shows the percentage of cases that were rated with a certain score. Figure 6.3 shows the majority of the clarifications were considered complete, relevant, clear and acceptable. A categorization of issues with examples is listed in Table 6.7.

6.5.3 Subgroup analysis

Clarifications with a single supertype were rated significantly higher than those with multiple supertypes on correctness, relevance, and acceptability. For completeness, this difference was not significant. Clarifications with patient-friendly synonyms were rated significantly higher on the completeness, but there was no significant difference on the other measures, see Appendix G. However, in 59% of these cases, a synonym itself, without any supertypes, would have been sufficient in clarifying the diagnosis and providing synonyms or definitions of supertypes did not add any value. In the other 41%, the synonym itself was too difficult and needed further clarification.

6.5.4 General remarks

Raters suggested adding lexical clarifications, e.g., what malignant or benign means. Definitions should not use enumeration with examples, because they are only applicable to some of the subtypes and may cause unnecessary worry, e.g., the definition pancreatic disorder contains the example of pancreatic cancer.

Additionally, raters noted some synonyms and definitions were too difficult to be used as a clarification.

Table 6.5 Ratings of dichotomous measures, as a percentage with 95% confidence interval (CI), and the interclass coefficient (ICC). * Correctness (whether a clarification contained errors) was only measured dichotomously. For the other measures, the dichotomous percentage was understood as a score of 1 or 2 (falsified).

Measure	ICC
Correctness	0.34
Completeness	0.45
Relevance	0.39
Clarity	0.45
Acceptability	0.43

Measure	Percentage	
	%	CI
Incorrect	12.7	10.7-14.6
Incomplete	14.2	12.1-16.2
Irrelevant	16.7	14.6-18.9
Unclear	18.0	15.7-20.2
Unacceptable	14.3	12.3-16.4

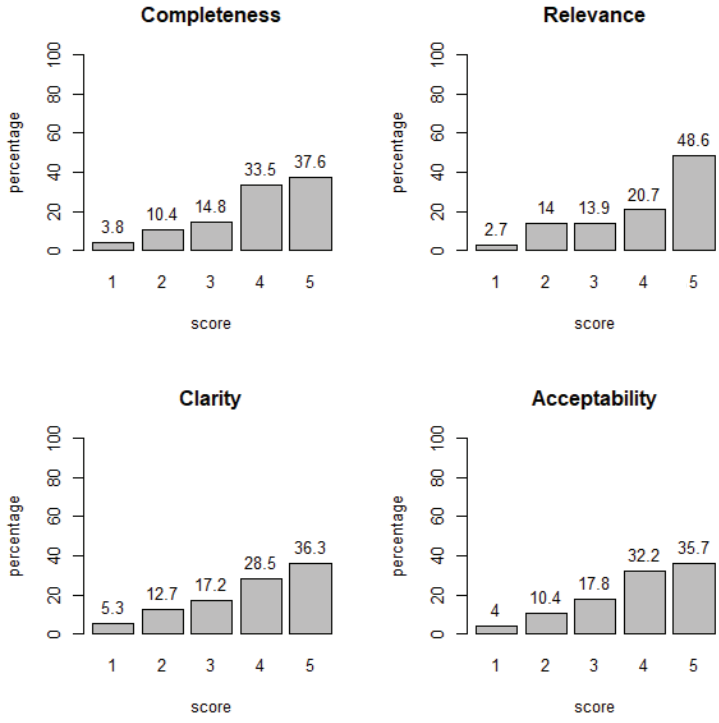


Figure 6.3 The percentage of clarifications with scores for completeness, relevance, clarity, and acceptability. The strictest ratings given by the raters are shown.

Table 6.6 Error category, with the number and percentage of clarifications with errors that had an error of that category, an example case, a description of the error in the example, and a description of the error category explaining what the category means. One clarification can contain multiple errors.

Error category	n	%	Example case	Example error	Error category description
DT-SNOMED CT mapping	9	5.9	'Duchenne female carrier A type of birth defect and limb-girdle muscular dystrophy. It's hereditary. 'limb-girdle' muscular dystrophy: This is a hereditary condition in which the muscles of the shoulders, upper arms, hips and thighs in particular do not function or function insufficiently.'	'Duchenne female carrier' from the DT was mapped on '240050008 |Manifesting female carrier of X-linked muscular dystrophy|' and thereby being clarified with several manifestations.	DT diagnoses that are not mapped correctly to SNOMED CT concepts
SNOMED CT modelling	25	16.3	'40425004 |Postconcussion syndrome| A type of dementia'	'40425004 |Postconcussion syndrome|' was modelled as being a '230282000 |Post-traumatic dementia|'	Defining relationships that are modelled incorrectly in SNOMED CT
PFE synonym	56	37.3	'32644009 |Braxton Hicks contractions| A type of childbirth'	Childbirth was the patient-friendly synonym of '118216002 |Labor finding|'	PFE synonyms that are not correctly mapped to SNOMED CT concepts
PFE definition	78	51.6	'20751000146109 |Recurrent nasopharyngeal carcinoma| A type of chronic illness and cancer in the throat. Chronic illness: Conditions with no prospect of full recovery and with a relatively long duration of illness, often requiring long-term care.'	A recurrent disease is a chronic disease in SNOMED CT. However, the definition of chronic disease 'with no prospect of recovery' does not apply to this recurrent disease.	PFE definitions that are not correctly mapped or not applicable to subtypes
Algorithm	14	9.2	'94360002 |Secondary malignant neoplasm of kidney| A type of cancer in kidney and metastases'	It seems like kidney cancer with metastases, but it rather is a metastasis in the kidney.	Splitting up into supertypes and word order by the algorithm changes the meaning incorrectly.

Table 6.7 Examples of issues concerning acceptability with the number and percentage of clarifications with that issue that were considered unacceptable, an example case, and a description of the issue in the particular example and a description of the issue category explaining what the category means. One clarification can have multiple issues.

Issue	n	%	Example case	Example issue	Issue category description
Error	93	54.1	See Table 6.6 for examples of errors		One or both raters found an error. Note that a clarification may be considered unacceptable, even if it does not contain an error (e.g., the example below).
Incomplete	102	59.3	'296495003 |Methocarbamol overdose| A type of disease. Disease: This is a deviation of the healthy state of the body or mind, which manifests itself in symptoms and disorders.'	There is no clarification of the disease itself, but rather the term disease is defined, which is too general.	Ranked as incomplete by one or both raters.
Irrelevant	147	85.5	'20781000146104 |Recurrent supratentorial primitive neuroectodermal tumor| A type of chronic disease, glioma, brain tumor and cancer. Chronic disease: Disorders with no prospect of full recovery and with a relatively long duration of illness, which often means that care is required for a long time. Glioma: This is a malignant tumor that develops in the supporting cells (glia) of the brain.'	The explanation of this recurrent disease as a chronic disease is incorrect, not relevant and unacceptable.	Ranked as irrelevant by one or both raters.
Unclear	154	89.5	'232357009 |Post-surgical epistaxis| A type of bleeding, complication and disease of the heart and vessels and nose. Bleeding: Loss of blood, such as bruises, arterial bleeding, etc. Complication: Unintended and unwanted outcomes during or following the actions of a healthcare provider, which are so harmful to the health of the patient that adjustment of the medical treatment is necessary is, or that there is irreparable damage; the cause of a complication may be unavoidable and lie in the underlying disease, in comorbidity or characteristics of the patient, or be avoidable and lie in the (non) action of a care provider and/or in the care system.'	This case contains a very complex clarification for a nosebleed after surgery.	Ranked as unclear by one or both raters.

Table 6.7 (continued)

Issue	n	%	Example case	Example issue	Issue category description
Algorithm	12	7.0	'254486004 |Benign neoplasm of nasal vestibule| A kind of benign tumor of airways and nasal disease.'	Airways is redundant when nose is mentioned already and now it might seem like multiple types of cancer	Splitting up into supertypes and word order by the algorithm changes the meaning unacceptably.
Examples in definition	7	4.1	'262830005 |Pancreatic duct injury| A pancreatic disease. Pancreatic disease: Diseases of the pancreas, such as pancreatic infection and pancreatic cancer.'	The example of pancreatic cancer is not applicable and may cause unnecessary worry.	Examples that might cause worry.
PFE definition missing	21	12.2	'267611002 |Degeneration of macula and posterior pole| A type of macular degeneration.'	Macular degeneration is a difficult synonym that needs further clarification.	The synonym is not patient-friendly and requires further clarification

6.6 Discussion

6.6.1 Principal findings

We validated diagnosis clarifications that were automatically generated by generalizing the diagnoses to patient-friendly synonyms and definitions of more general concepts in the PFE. We found the majority of clarifications were considered correct, acceptable, complete, relevant, and clear by both raters. However, we identified errors and unacceptable results in some clarifications that need to be addressed. Clarifications that were considered to be the most unacceptable were incorrect and seemed to cause worry. The interrater reliability of correctness (ICC: .342) was poor and of acceptability (ICC: .429) was fair[193], however, which shows that raters had different interpretations. Most problems that we found were associated with the patient-friendly synonyms and definitions from the PFE themselves, SNOMED CT modelling, and terminology mappings. Some identified problems related to the algorithm itself, creating unnecessarily complicated clarifications or changing the semantics by splitting the concept into its supertypes.

6.6.2 Results in relation to other studies

Zeng et al[190] used the Unified Medical Language System to provide generalizations for concepts and encountered similar problems: too general explanations that were not helpful 'nulligravida (a type of finding)' and some incorrect explanations. Zeng et al used a readability score to assess whether a term needed to be translated, while we assumed that terms and definitions in the PFE were at the right difficulty level. Their system was used to clarify free-text documents and was evaluated with a sample of EHR notes. They didn't address the validity of the clarifications on the level of the terminology but rather assessed the performance on the level of the text. Research on text simplification[175,190,191,194,195], with a focus on free-text documents and records, has to address named entity recognition and medical concept normalization, to determine which concepts are represented in the terms that might contain ambiguity and spelling errors. We avoided this by starting from diagnosis data that are coded with terminology systems. This way, we could focus on the task of clarifying the specific medical concept, using knowledge already available in SNOMED CT ontology. However, we still encountered issues in the underlying terminologies our approach was based on.

Studies[196-198] on ontology verbalization have aimed to convert SNOMED CT ontology into text. This can be used to enable users to understand the formal SNOMED CT representations for terminology auditing and to access definitions

of concepts. These studies have not yet provided information on a patient-friendly level of detail. Our approach can also be considered a type of ontology verbalization, by retrieving knowledge represented in the SNOMED CT hierarchy and terminology mappings. Additionally to the challenges of verbalization, we found issues in the SNOMED CT hierarchy and the level of detail required to clarify concepts to laypersons.

6.6.3 Implications of the study and recommendations

SNOMED CT logically specifies disorders well that are fully defined, for example by morphology and finding site. However, some domains are underspecified, such as genetic and metabolic disorders, because SNOMED CT does not contain formal specifications of genes and proteins. Furthermore, concepts of syndromes and uncertainty are not fully defined in SNOMED CT. The algorithm splits concepts along the anatomic hierarchy, which can lead to unexpected results. These problems imply that the ontology verbalization approach will work better with fully defined concepts, rather than the generalization approach. We will investigate a combination of these approaches in further research. In our study, we focused on Dutch diagnoses and the PFE, but the algorithm can also be applied to other parts of SNOMED CT, and other languages and terminology systems. We expect that one would encounter similar issues.

Most problems we identified can be resolved by improvements in the algorithm, terms and definitions. They should be taken into account and improvements should be made accordingly where possible before implementing the clarifications in a patient portal. Most of the errors related to PFE synonyms and definitions arose because they did not apply to subtypes. We therefore recommend that, when terminologists add synonyms and definitions to SNOMED CT concepts in the PFE, the applicability to subtypes is taken into account to ensure the synonym or definition is also applicable to its subtypes. This implies excluding examples in definitions because those are not always applicable to subtypes. Additionally, the PFE should provide terms and definitions at a patient-friendly, low reading difficulty level; currently, the PFE still contains some difficult definitions that are not easy to read. A readability measure could also be considered, for example using frequency scores, to determine whether a term needs a definition. In focus groups with several Dutch terminology organizations[199], there was indeed disagreement on whether the terms and definitions in the PFE should be considered “patient-friendly”. Additionally, “patient-friendliness” itself was a controversial concept. In one of the focus groups it was agreed that terms and definitions should be written at B1 level (in the European Reference Framework,

i.e., using more common, plain language terms, and more simple formulations) or should be tailored to the user. In future additions and modifications of terms and definitions in the PFE this will be taken into account. In our study, we found that the experts did not always identify the same errors. The PFE, and results from future similar studies, should therefore be validated by more than one evaluator, preferably with a background in the particular medical specialty in which the concept is used. Although this was not the initial objective of our study, the problems with terminology mappings and modelling we found show that this method can be used to evaluate the quality of the underlying terminologies.

6.6.4 Strengths and weaknesses

The selection of clarifications covers all patient-friendly synonyms and definitions from the PFE that are supertypes of diagnoses from the DT and a large part of the unique combinations of supertypes used in the diagnosis clarifications. We, therefore, consider it to be representative of other diagnoses. However, the diagnosis clarifications that were not included due to the limited sample size still have to be validated. Additionally, clarifications should be revalidated when new terminology versions are released that change the relationships, terms and definitions used to generate the clarifications. We assumed the patient-friendly terms and definitions were included in SNOMED CT on the right level of detail and written in a patient-friendly form, but some terms and definitions were too difficult. Additionally, the definitions were not created to generalize more specific concepts to more general ones. We expect that results can be improved by taking this into account when improving or adding terms and definitions to the PFE.

The poor interrater reliability shows that experts identify different errors and differ in their opinion on whether patient-friendly clarifications are acceptable, which indicates that the validation instrument could be improved and emphasizes that more than one rater is needed to identify issues in the clarifications. One rater systematically identified more errors and provided lower ratings than the other rater. While the agreement between raters on correctness was high (91%), the ICC was affected by the skewed distribution of the correctness: there was a high a priori probability of agreement because 87% of clarifications were considered correct by both raters, which led to a higher penalty by the ICC in case of disagreement. The divergent interpretations seemed to be related to having a different medical specialization and clinical experience. For example, ‘steroid responder’ is an ophthalmologic disorder in the DT, but may seem not to be a diagnosis from the point of view of another medical specialty.

6.6.5 Future work

The clarifications have only been medically evaluated in this validation study. Ultimately, patients and their significant others are supposed to benefit from clarifications and further evaluation by these end-users is planned. However, a medical evaluation is a necessary first step to assess the validity of the patient-friendly terms, as patients are incapable of this assessment. The future validation instrument should be improved to increase the reliability and to specifically gauge the experience and understanding of the patients. While generalization increases the content coverage and reuse of terms and definitions, the dependence upon the quality of underlying terminology systems introduces uncertainty about the correctness and acceptability of the clarifications. In further research, we thus have to address how this approach should be used in practice: whether the automatic clarifications can be used in practice directly, with some disclaimer or warning, or whether it should only be used as a first draft clarification to develop patient-friendly terminology.

6.7 Conclusion

We found the majority of diagnosis clarifications to be considered correct, complete, relevant, clear and acceptable but also identified some problems in diagnosis generalization to patient-friendly synonyms and definitions. Most of the found problems can be resolved, as they were related to the underlying terminologies. This shows generalization using the SNOMED CT hierarchy can be a useful and adequate method to increase the content coverage of the PFE. Before implementing the clarifications in practice some modifications have to be made to the terms and definitions, and this needs to be updated when new terminology versions are released. Further research should focus on improving the quality of the underlying terminologies, the relevance and clarity of the clarifications, and on evaluation with end-users.

6.8 Declarations

6.8.1 Acknowledgements

Thanks to Pim Volkert, Annemarieke Homan, Feikje Hielkema, Natasha Krul, Eef Verbeek, Daniël Huliselan, Theo van Mens and Frank van Oosterhout, for their contributions to this study.

6.8.2 Consent to participate and for publication

Written informed consent to participate and informed consent for publication was obtained from the authors SMa and EP who rated the clarifications.

6.8.3 Availability of data and materials

Data are available from the corresponding author on reasonable request.

6.9 Appendices

6.9.1 Appendix A

Randomization.R sampling script is available in the online version of this paper at https://doi.org/10.1016/j.jbi.2022.104071.

6.9.2 Appendix B

Background about terminologies and sampling procedure.

6.9.2.1 Terminologies

The Diagnosethesaurus is used in Dutch hospitals to register diagnoses and contains multiple synonymous terms for each diagnosis concept. One of the terms is the preferred term of the concept, which is shown in the electronic health record (EHR). The Diagnosethesaurus is mapped to SNOMED CT concepts. Each SNOMED CT concept has multiple descriptions. A description may be a synonym or a text definition. Descriptions are referenced again by language reference sets, that state whether the description of the term is acceptable or preferred in a reference set. SNOMED CT Netherlands particularly releases a Patient-Friendly Extension that contains the SNOMED CT Dutch module for patient-friendly descriptions and Patient-friendly Dutch language reference set. In this study we aimed to clarify diagnoses registered with the Diagnosethesaurus by providing patient-friendly terms and definitions from the SNOMED CT Patient-friendly Dutch language reference set. Figure 6.4 visualizes the relationships between these terminologies.

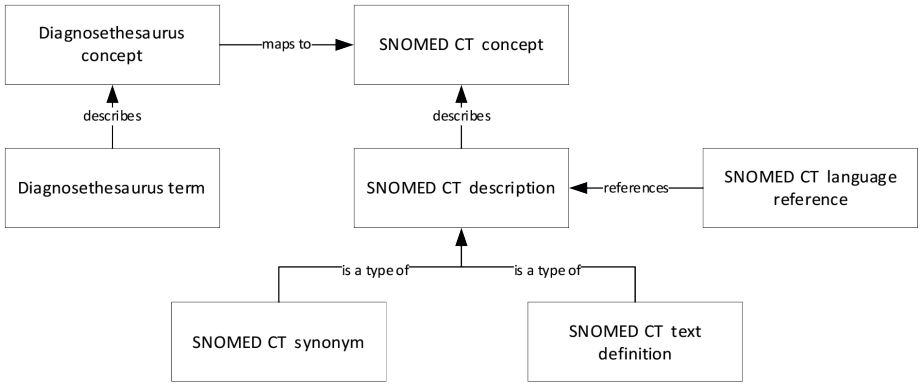


Figure 6.4 Relationships between Diagnosethesaurus concepts and terms, and SNOMED CT concepts, descriptions, and references.

6.9.2.2 Clarification sampling

For each SNOMED CT concept that was referenced by the Patient-friendly Dutch language reference set, we sampled the subtypes that are registered in clinical practice in the Netherlands with the Diagnosethesaurus. When available we took one random subtype that had a clarification with only that supertype. Additionally, we took another random subtype that had a clarification of that supertype, but also with other supertypes referenced by the patient-friendly reference set. In total we selected 1200 clarifications this way. We excluded concepts that already had a patient-friendly definition, and concepts that had no supertype concepts that were referenced by the patient-friendly reference set. We registered the inclusion reasons for each diagnosis sampled. Some supertypes only have one subtype that use the supertype as the only supertype. Some supertypes only have one subtype that use supertype together with multiple other supertypes. These were all included as 'Only single' or 'Only multiple'.

Table 6.8 Inclusion reasons with their description and rationale, with the number of supertypes s, the number of diagnoses sampled n and the number of diagnoses from which was sampled N.

#	Reason	Description	Rationale	s	n	N
1	Only single	The only diagnosis with the supertype as single supertype	There is only one diagnosis with the supertype as the only supertype. In sampling it is hence included automatically.	115	115	115
2	Sampled single	Randomly sampled the diagnosis from all diagnoses that had the supertype as single supertype		416	416	9571

Table 6.8 (continued)

#	Reason	Description	Rationale	s	n	N
3	Only multiple	The only diagnosis with the supertype as one of multiple supertypes	There is only one diagnosis with the supertype as one of multiple other supertypes. In sampling it is hence included automatically.	111	108	108
4	Extra multiple	Contains combinations of the other multiple supertypes from the included diagnoses in 'Only multiple'	Diagnoses from 'Only multiple' also had unique supertypes, other than the sampled ones (because they have multiple). These thus already have been sampled in the previous step.	72	0	1725
5	Sampled multiple	Randomly sampled the diagnosis from all diagnoses that had the supertype as one of multiple supertypes, for those supertypes that were not already included earlier during sampling multiples		320	255 - 264	4605
6	Sampled uniques	Sampled extra diagnoses with 2 to 7 supertypes, by increasing their coverage to one-third of all unique combinations, for each number of supertypes (2 to 7)	There are few diagnoses with many supertypes (2, 3, 4, 5, 6 or 7) supertypes. During sampling they might be underrepresented otherwise, while they result in the most complicated clarifications.		146 -154	
7	Sampled synonyms	Sampled from diagnosis with synonyms so all unique combinations of supertypes were covered	There are few diagnoses with synonyms, but fewer with a text definition.		118 - 119	
8	Sampled extra	Sampled from diagnoses not yet included	To reach the sample size some random extra diagnosis are sampled		9- 22	
9	Sampled first intra-rater	Sampled from the first 30 samples	To calculate the intra-rater reliability for the first 60 (- 3 = 57) cases		3	
10	Sampled second intra-rater	Sampled from samples 99 to 999	To calculate the intra-rater reliability for the last 1140 (- 6 = 1134) cases		6	
	Total				**1200**	

6.9.3 Appendix C

6.9.3.1 English translation of Dutch guide for participants

Diagnosis clarifications validation instructions

Background

Patient-friendly terms and definitions can help patients and their significant others understand their medical data. There is, however, a myriad of terms that is used in the medical domain and it costs a lot of time and effort to make patient-friendly descriptions for all these terms, in all languages. A description in more general terms might already be sufficient for a layperson, leaving out some of the specialist details. In this study, we want to ask you to validate clarifications of diagnoses. The clarifications are based on the generalization of diagnoses to more general terms and definitions. We generate the clarifications with the SNOMED CT hierarchy, by generalizing diagnoses to concepts with terms and definitions from the patient-friendly extension. The clarifications are to be used to clarify the meaning of the diagnoses in patient portals or personal health records, where patients and their significant others can view the medical descriptions of the diagnoses on their problem list.

Content research

We will start the study with some background questions. Then we illustrate the procedure with six examples and discuss the instructions. After completing the first 60 clarifications independently from the other participant, we will discuss the results together to reach consensus. Then you can continue validating the rest of the clarifications. We ask you to keep the time it took to execute the study.

Throughout this research it is important to take the perspective of you as a medical doctor. You have to imagine that you registered the diagnoses in the medical record of the patient and that the patient can later find the diagnoses in the problem list in the patient portal. Next to the medical description of the diagnoses as you registered, there will be the clarification to briefly clarify what the diagnosis means. This way, the patient and its significant other know what the diagnosis is about. The diagnosis has usually already been discussed with the patient. The clarification is not a replacement for patient education. The patient can view the diagnosis and the clarification after a consultation, or months or years later, in the patient portal.

Background questions

We will ask you to provide this data during the instructions and they will be reported in the paper.

- Gender
- Age group
- What is your medical specialty?
- How many years of work experience do you have in healthcare, including internships?
- How many years of work experience do you have with formal terminology systems, such as the Diagnosethesaurus and SNOMED CT?

Questions about each diagnosis clarification

You will get to see a diagnosis term, as it is registered in the medical record, and a description that contains the clarification that the patient will be able to see. You have to answer the questions below about each diagnosis clarification. You only get one task at the same time, so you cannot go back to the previous task after a task has been completed. You are allowed to look up anything you deem necessary to assess the diagnosis clarification. If you have any questions or something goes wrong, please contact us. Below you find some explanations of the questions. Please read it carefully and use it while answering the questions. The last four questions are on a scale from 1 to 5, where, e.g., 1 is very unacceptable, 2 unacceptable, 3 not unacceptable or acceptable, 4 acceptable and 5 very acceptable. In the last question you evaluate if the clarification can be displayed in the patient portal according to your opinion, or if you have objections against that.

Question	Explanation
Cannot evaluate the clarification?	If you believe you cannot evaluate the clarification, because something went wrong, because you do not have the expertise, or because of some other reason.
Why not?	Here you can write why it could not be evaluated.
Does this clarification contain mistakes from a medical perspective?	The clarifications may not contain clinical mistakes that would provide wrong information to a patient, significant other or layperson. Also small errors or corrections are relevant. This does not address the completeness, relevance, clarity or whether it is acceptable, but only mistakes in the clarification of the diagnosis.
If it contains mistakes, what are the mistakes exactly?	If there are mistakes, we ask you to specify what is wrong with the clarification and how it should be improved.

Question	Explanation
When you register this diagnosis in the medical record of the patient and the diagnosis is displayed on the patient portal to your patient or its significant others, what do you think about that? When not complete, relevant, clear or acceptable, can you describe what makes it incomplete, irrelevant, unclear or unacceptable and how it can be improved?	Here you can provide feedback on the clarification. Here is when we ask you to evaluate the completeness, relevance, clarity and how acceptable it is.
How complete is the clarification from 1 very incomplete to 5 very complete, where 3 is not incomplete or complete?	Up to what extent you miss important information, that you would use as a medical doctor in a short clarification of the diagnosis.
How relevant is the clarification from 1 very irrelevant to 5 very relevant?	Up to what extent the information that is being provided in the clarification is relevant at all, to clarify the diagnosis to a patient or significant other.
How clear is the clarification from 1 very unclear to 5 very clear?	Up to what extent you consider the clarification to be clearly formulated for the understanding of the diagnosis, because it should not be too vague, confusing or difficulty.
When you register this diagnosis in the medical record of the patient and the diagnosis is displayed on the problem list of your patient or his or her significant others: How acceptable is this clarification from 1 very unacceptable to 5 very acceptable, where 3 is not unacceptable or acceptable?	Up to what extent you consider that this clarification is allowed to be displayed for this diagnosis in the problem list of the patient. If you really have an objection that this clarification is displayed when a patient or significant other clicks on the info button, then it should be 1 very unacceptable or 2 unacceptable.

Questions afterwards about the clarifications in general

Afterwards we will discuss the following points.

- How many hours did it take in total?
- What is or is not acceptable to implement into a patient portal?
- What has to be improved to use the clarifications in practice?

6.9.3.2 Original Dutch guide for participants

Instructies validatie diagnosetoelichtingen

Achtergrond

Patiëntvriendelijke termen en definities kunnen patiënten en hun naasten helpen om hun medische gegevens te begrijpen. Er zijn echter ontzettend veel termen die in het medische domein gebruikt worden en het kost veel tijd en moeite om patiëntvriendelijke beschrijvingen voor al deze termen te maken, in alle talen. Een beschrijving in meer algemene termen zou al genoeg kunnen zijn voor een leek, waarbij sommige specialistische details weggelaten worden. In dit onderzoek willen we u vragen om toelichtingen op diagnoses valideren. De toelichtingen zijn gebaseerd op de generalisering van diagnoses naar meer algemene termen en definities. We genereren de toelichtingen met de SNOMED CT hiërarchie, door de diagnoses te generaliseren naar concepten met termen en definities uit de patiëntvriendelijke extensie. De toelichtingen zouden gebruikt kunnen worden om de betekenis van de diagnoses te verduidelijken in patiëntenportalen en persoonlijke gezondheidsomgevingen, waarin patiënten en hun naasten de medische beschrijvingen van de diagnoses op hun probleemlijst kunnen inzien.

Inhoud onderzoek

We beginnen het onderzoek met enkele achtergrondvragen. Dan illustreren we de procedure met zes voorbeelden en bespreken we de instructies daarbij. Na het afronden van de eerste 60 toelichtingen onafhankelijk van de andere deelnemer, bespreken we de resultaten gezamenlijk om consensus te bereiken. Daarna kunt u verder met het valideren van de rest van de toelichtingen. We vragen u om steeds de tijd bij te houden die het kostte om het onderzoek uit te voeren.

Tijdens dit onderzoek is het van belang het perspectief van u als arts in te nemen. U moet zich voorstellen dat u de diagnose heeft geregistreerd in het medisch dossier van de patiënt en dat de patiënt daarna de diagnose op het patiëntportaal kan vinden in de probleemlijst. Naast de medische omschrijving van de diagnose, zoals u die heeft geregistreerd, staat dan de toelichting om in het kort te verduidelijken wat de diagnose betekent. Op deze manier weten de patiënt en zijn naasten waarover het gaat. De diagnose is dan doorgaans al besproken met de patiënt. De toelichting is geen vervanging voor patiëntvoorlichting. De patiënt kan de diagnose en de toelichting na een consult, of maanden of jaren later, nog terugzien in het patiëntenportaal.

Achtergrondvragen

Deze gegevens vragen we u tijdens de instructies op te geven en komen ook in het paper.

- Geslacht
- Leeftijdsgroep
- Wat is uw medisch specialisme?
- Hoeveel jaar werkervaring heeft u in gezondheidszorg, inclusief coschappen?
- Hoeveel jaar werkervaring heeft u met formele terminologiesystemen, zoals de Diagnosethesaurus en SNOMED CT?

Vragen over elke diagnosetoelichting

U krijgt een diagnoseterm te zien, zoals die in het dossier wordt geregistreerd, en een beschrijving waarin de toelichting staat die de patiënt erbij te zien krijgt. Bij deze uitleg van de diagnose moet u de onderstaande vragen beantwoorden. U krijgt steeds maar één taak tegelijk, u kunt niet meer teruggaan naar de vorige taak nadat een taak is voltooid. U mag wel alles opzoeken wat u nodig acht om de diagnosetoelichting te beoordelen. Mocht u vragen hebben of er iets misgaan, neem dan even contact op. Hieronder de uitleg bij de vragen. Neemt u deze goed door en gebruik deze ook tijdens de beantwoording van de vragen. De laatste vier vragen zijn op een schaal van 1 tot 5, waarbij dus bijvoorbeeld het gaat om 1 heel onacceptabel, 2 onacceptabel, 3 niet onacceptabel of acceptabel, 4 acceptabel en 5 heel acceptabel. In de laatste vraag beoordeelt u of de toelichting naar uw inzicht in het patiëntenportaal getoond mag worden of dat u daar bezwaren tegen heeft.

Vraag	Uitleg
Kan toelichting niet beoordelen?	Als u vindt dat u de toelichting niet kan beoordelen, omdat er iets mis is gegaan, omdat u de expertise niet heeft, of om een andere reden.
Waarom niet?	Hier kunt u aangeven waarom het niet beoordeeld kon worden.
Staan er fouten in deze toelichting vanuit een medisch perspectief?	De toelichtingen mogen geen klinische onjuistheden bevatten die verkeerde informatie zouden geven aan een patiënt, naaste of leek. Ook kleine fouten of correcties zijn relevant. Het gaat hierbij niet over de compleetheid, relevantie, duidelijkheid of dat het acceptabel is maar echt over fouten in de toelichting op de diagnose alleen.

Vraag	Uitleg
Wanneer er fouten in staan, wat zijn dan precies de fouten?	Als er fouten in staan, dan vragen we u te specificeren wat er mis is met de toelichting en hoe het verbeterd zou moeten worden.
Wanneer u deze diagnose registreert in het medische dossier van de patiënt en de diagnose op de probleemlijst in het patiëntenportaal aan uw patiënt en zijn of haar naasten wordt getoond, wat vindt u daarvan? Wanneer niet compleet, relevant, duidelijk of acceptabel: kunt u beschrijven wat het incompleet, irrelevant, onduidelijk of onacceptabel maakt en hoe zou het verbeterd kunnen worden?	Hier kunt u feedback geven op de toelichting. Hier vragen we u pas om de compleetheid, relevantie, duidelijkheid en hoe acceptabel het is te beoordelen.
Hoe compleet is deze toelichting, van 1 heel incompleet tot 5 heel compleet, waarbij 3 niet incompleet of compleet is?	In hoeverre belangrijke informatie mist, die u als arts zou gebruiken in een korte toelichting op de diagnose.
Hoe relevant is deze toelichting, van 1 heel irrelevant tot 5 heel relevant, waarbij 3 niet irrelevant of relevant is?	In hoeverre de informatie die wel gegeven wordt in de toelichting überhaupt relevant is om de diagnose toe te lichten aan een patiënt of naaste.
Hoe duidelijk is deze toelichting, van 1 heel onduidelijk tot 5 heel duidelijk, waarbij 3 niet onduidelijk of duidelijk is?	In hoeverre u de toelichting duidelijk vindt geformuleerd voor het begrip van de diagnose, omdat het niet te vaag, verwarrend of moeilijk mag zijn.
Wanneer u deze diagnose registreert in het medische dossier van de patiënt en de diagnose op de probleemlijst in het patiëntenportaal aan uw patiënt en zijn of haar naasten wordt getoond: Hoe acceptabel is deze toelichting, van 1 heel onacceptabel tot 5 heel acceptabel, waarbij 3 niet onacceptabel of acceptabel is?	In hoeverre het acceptabel vindt dat deze toelichting getoond wordt bij de diagnose op de probleemlijst van de patiënt. Als u echt een bezwaar heeft dat deze toelichting getoond wordt wanneer een patiënt of naaste op een infoknop klikt dan zou het dus 1 heel onacceptabel of 2 onacceptabel zijn.

Vragen achteraf over de toelichtingen in het algemeen

Achteraf bespreken we de volgende punten.

- Hoeveel uur kostte het in totaal?
- Wat is wel of niet acceptabel om te implementeren in een patiëntenportaal?
- Wat moet er verbeteren om de toelichtingen in de praktijk te gebruiken?

6.9.4 Appendix D

Data dictionary.

Table 6.9 Data to be imported into the case report form

Id	Variable name	Data type
1	SortIndex	Int
2	SctConceptId	Int (64 bit)
3	DtConceptId	Int
4	DtPreferredTerm	String
5	Clarification	String

Table 6.10 Participant characteristics

Id	Variable name	Data type	Input type
1	ParticipantId	Int	Autonum
2	Gender	Int	Radio buttons [Male, female, other]
3	AgeGroup	Int	Radio buttons [20-29, 30-39, 40-49] etc.
4	MedicalSpecialty	String	Text area
5	YearsHealthcare	Int	Number
6	YearsTerminology	Int	Number

Table 6.11 Clarification validation data

Id	Variable name	Data type	Input type
1	ParticipantId	Int	Display value
2	SortIndex	Int	Display value
3	SctConceptId	Int (64 bit)	Display value
4	CannotValidate	Boolean	Checkbox
5	WhyNotValidate	String	Text Area
6	ContainsErrors	Boolean	Radio buttons [Yes, No]
7	Errors	String	Text area
8	Completeness	Int	Radio buttons from 1 to 5
10	Relevance	Int	Radio buttons from 1 to 5
11	Clarity	Int	Radio buttons from 1 to 5
12	IsAcceptable	Boolean	Radio buttons from 1 to 5
13	NotAcceptableNotes	String	Text area
14	Suggestion	String	Text area

Table 6.12 Tool evaluation data

Id	Variable name	Data type	Input type
1	ParticipantId	Int	Autonum
2	HoursTaken	Int	Number
3	ToolFeedback	String	Text area

6.9.5 Appendix E

Script for statistical analysis is available in the online version of this paper at https://doi.org/10.1016/j.jbi.2022.104071.

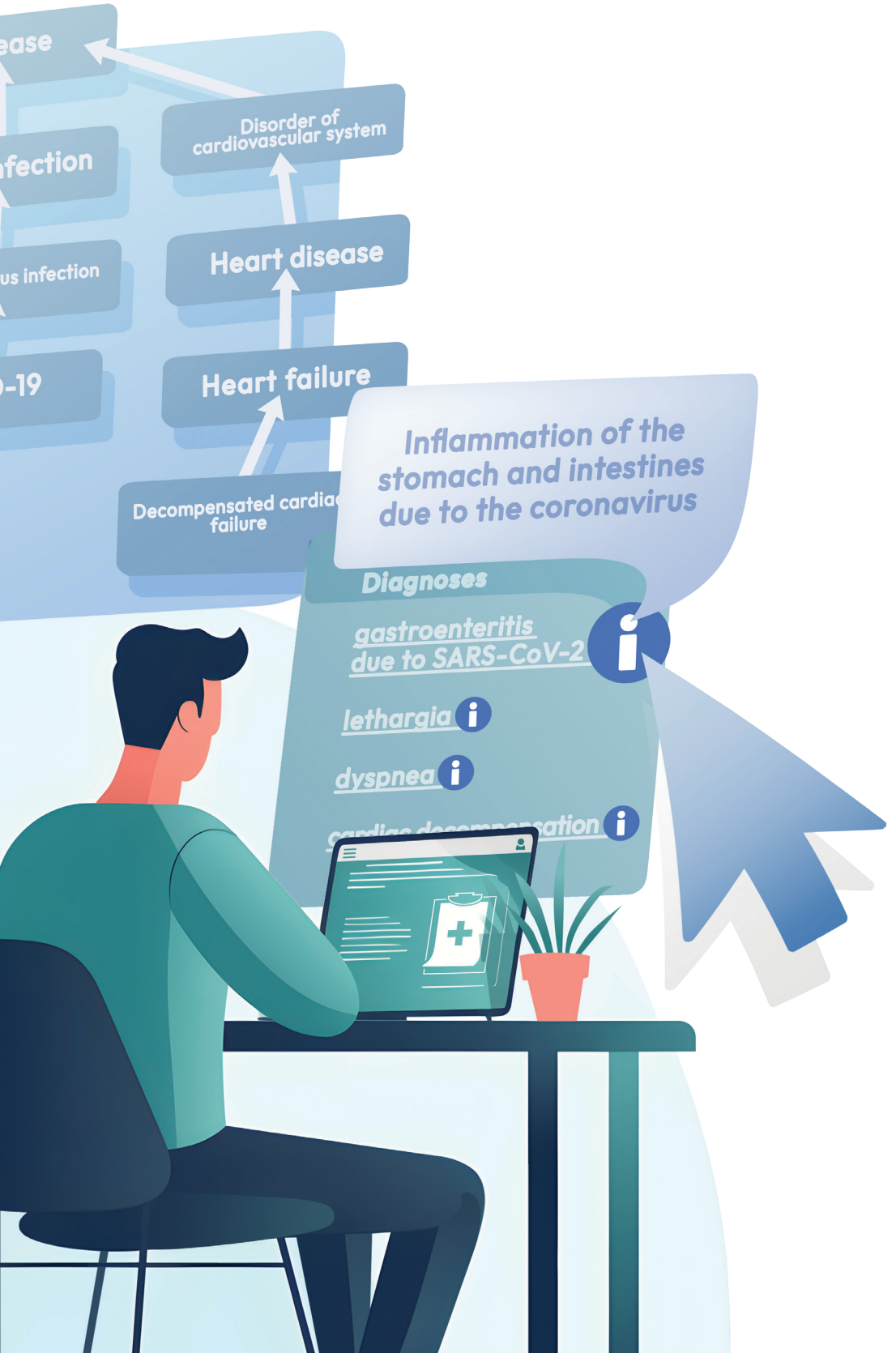
6.9.6 Appendix F

Validation results categorization is available in the online version of this paper at https://doi.org/10.1016/j.jbi.2022.104071.

6.9.7 Appendix G

Table 6.13 Comparison of outcomes between rater 1 and 2, between concepts with a single supertype or multiple supertypes, and between concepts with a patient-friendly synonym and those that did not have a patient-friendly synonym. Percentages with lower and upper value of the confidence interval, total number of cases, whether the CI was different and the p-value of the Mann-Whitney test, with column s containing an asterisk where p-value < 0.05.

from	measure	%	lower	upper	total	from	measure	%	lower	upper	total	p	s
Rater 1	Error	11.1	9.3	12.9	1125	**Rater 2**	Error	4.5	3.3	5.8	1123	0	*
	Not Complete	9.2	7.5	10.8	1125		Not Complete	8.5	6.8	10.1	1123	0.29	
	Not Relevant	14.0	12.0	16.1	1125		Not Relevant	6.3	4.9	7.7	1123	0	*
	Not Clear	12.1	10.2	14.0	1125		Not Clear	9.7	8.0	11.4	1123	0	*
	Not Accept	11.2	9.4	13.0	1125		Not Accept	5.3	3.9	6.6	1123	0	*
Singles	Error	7.4	5.1	9.7	502	**Multiples**	Error	18.9	15.4	22.4	486	0	*
	Not Complete	16.3	13.1	19.6	502		Not Complete	14.0	10.9	17.1	486	0.83	
	Not Relevant	13.3	10.4	16.3	502		Not Relevant	21.4	17.8	25.0	486	0	*
	Not Clear	16.9	13.7	20.2	502		Not Clear	21.2	17.6	24.8	486	0	*
	Not Accept	12.0	9.1	14.8	502		Not Accept	18.7	15.3	22.2	486	< 0.01	*
Friendly	Error	10.9	5.3	16.5	119	**Unfriendly**	Error	12.9	10.8	14.9	1010	0.55	
	Not Complete	3.4	0.1	6.6	119		Not Complete	15.4	13.2	17.7	1010	< 0.01	*
	Not Relevant	12.6	6.6	18.6	119		Not Relevant	17.2	14.9	19.6	1010	0.87	
	Not Clear	10.1	4.7	15.5	119		Not Clear	18.9	16.5	21.3	1010	0.05	
	Not Accept	7.6	2.8	12.3	119		Not Accept	15.1	12.9	17.4	1010	0.15	


ease
Disorder of cardiovascular system
fection
Heart disease
us infection
-19
Heart failure
Decompensated cardiac failure
Inflammation of the stomach and intestines due to the coronavirus
Diagnoses
gastroenteritis due to SARS-CoV-2
lethargia
dyspnea

7

Evaluation of patient-friendly diagnosis clarifications in a hospital patient portal

Authors:
Hugo J.T. van Mens
Gaby E.G. Hannen
Remko Nienhuis
Roel J. Bolt
Nicolette F. de Keizer
Ronald Cornet

Status: Accepted.
Journal: Applied Clinical Informatics.
Year: 2023. **ISSN:** 1869-0327.
DOI: 10.1055/a-2067-5310.

Abstract

Background Medical data can be difficult to comprehend for patients, but only a limited number of patient-friendly terms and definitions are available to clarify medical concepts. Therefore, we developed an algorithm that generalizes diagnoses to more general concepts that do have patient-friendly terms and definitions in SNOMED CT. We implemented the generalizations, and diagnosis clarifications with synonyms and definitions that were already available, in the problem list of a hospital patient portal.

Objective We aimed to assess the extent to which the clarifications cover the diagnoses in the problem list, the extent to which they are used and appreciated by actual patient portal users in a real-life setting, and to explore differences in viewing problems and clarifications between subgroups of users and diagnoses.

Methods We measured the coverage and usage of the clarifications, and user and diagnosis characteristics with aggregated, routinely available EHR and log file data. Additionally, quantitative and qualitative feedback from patient portal users was collected about the quality of the clarifications.

Results Of all patient portal users that viewed diagnoses on their problem list (n = 2,660), 89% had one or more diagnoses with clarifications. We found that 55% of patient portal users viewed the clarifications. Users that rated the clarifications (n = 108) considered the clarifications to be of good quality on average, with a median rating per patient of 6 (interquartile range: 4 – 7; from 1 very bad to 7 very good). Users commented that they found clarifications to be clear and recognized the clarifications from their own experience, but sometimes also found the clarifications incomplete or disagreed with the diagnosis itself.

Conclusions This study shows that the clarifications are used and appreciated by patient portal users. Further research and development will be dedicated to the maintenance and further quality improvement of the clarifications.

7.1 Background and Significance

Medical data can be difficult to comprehend for patients[4,56,70], but only a limited number of patient-friendly terms and definitions are available to clarify medical concepts. Patients understand medical data on a more general level, in less detail than clinicians.[200] To clarify what a medical term means, we hypothesize that a short description in more general terms might thus be sufficient if it is at the right level of detail. While this does not replace the need to inform patients thoroughly during consultations and to provide patient information resources, this can help patients understand data in their medical records.

Therefore, we developed a method to generalize diagnoses to more general concepts that do have patient-friendly terms and definitions in the SNOMED CT Netherlands Patient-Friendly Extension (denoted as "PFE" hereafter), by employing the SNOMED CT hierarchy[184,201]. We showed that this method increases the number of diagnoses that can be clarified significantly[184]. Additionally, more than 85% of clarifications were regarded as correct and acceptable to use in practice[200]. We further improved the clarifications based on the input from the validation study and updated the clarifications with the latest version of the PFE. The final set of clarifications contained clarifications consisting of direct synonyms and definitions available in the PFE (e.g. 'phlebitis' in Table 7.1) and clarifications that were generated by the generalization to concepts that do have PFE synonyms and definitions (e.g. 'pulmonic valve regurgitation' in Table 7.1).

Table 7.1 Examples of diagnoses registered in Dutch problem lists of medical records and their corresponding clarifications that can be displayed after clicking on the diagnosis description or info button.

Medical diagnosis description	Clarification
Phlebitis	Another word for "phlebitis" is inflammation of vein: Inflammation of a vein, which makes it red, swollen, and painful.
Pulmonic valve regurgitation	A type of leaky heart valve. Leaky heart valve: This is a heart valve that closes poorly so that oxygen-rich blood no longer flows properly through the body. This causes complaints such as shortness of breath, fatigue after exertion and dizziness.
Congenital cyst of adrenal gland	A type of inborn abnormality and hormonal disorder. Cyst: cavities in the body filled with liquid.
Lowe syndrome	A type of inborn abnormality, mental disorder, and disorder of brain, kidney, eye, and metabolism. It is hereditary.

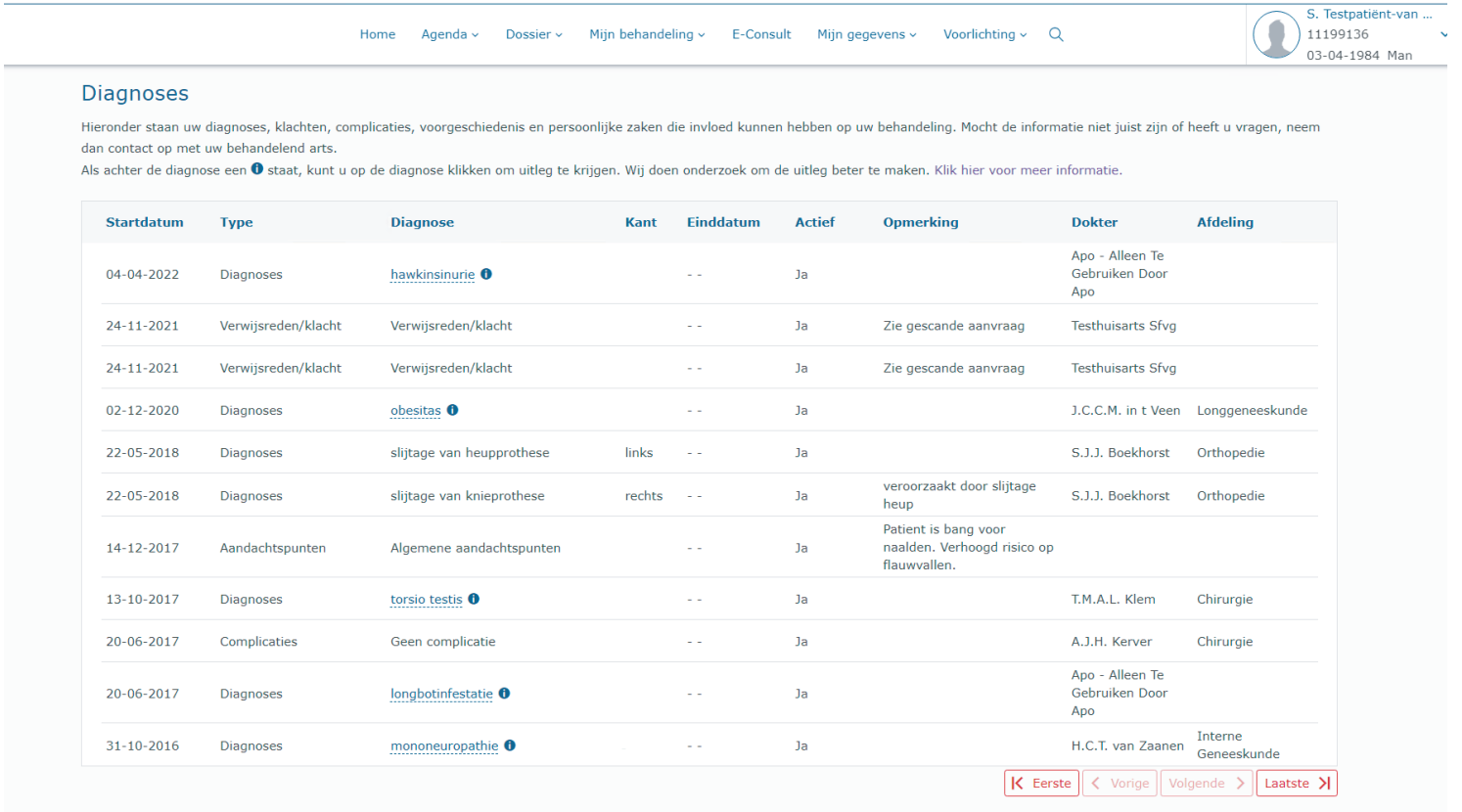
7.2 Objective

The clarifications had not been evaluated by actual patient portal users in a real-life setting. The current study aims to evaluate the implementation of these diagnosis clarifications in a patient portal problem list. First, we aimed to assess the coverage of the clarifications and evaluate patient portal users' information needs by analysing to what extent they use the clarification functionality when they view their problem list. Second, we evaluated the quality of the clarifications from the perspective of the users and explored differences in user and diagnosis characteristics between those that only view their problem list and those that use the additional clarification functionality.

7.3 Study Context

7.3.1 System Details

The study was carried out at the teaching hospital Franciscus Gasthuis & Vlietland (Franciscus).[202] The hospital used the health information system HiX and its patient portal (version 6.2; ChipSoft B.V., Amsterdam, The Netherlands). Patients, or their authorized proxies, use the patient portal, for instance, to view their medical data, schedule appointments, securely message their health care provider, and complete questionnaires. Proxy users can be anyone authorized by the hospital or the patients (depending on their age), such as informal caregivers, case managers or the parents of a child. The diagnosis clarifications were implemented in the problem list, which contains diagnoses, complications and attention notes. See Figure 7.1. The description of the diagnosis was highlighted, underlined and provided with an info icon if a clarification was available. When clicked, the diagnosis description and a clarification of the diagnosis were displayed. A warning was displayed for the clarifications with supertypes, stating that the clarification was generated automatically and might contain errors. For questions about their diagnosis, patients were referred to their doctors. Figure 7.2 and Figure 7.3 illustrate the clarifications, feedback, information and warnings.



Startdatum	Type	Diagnose	Kant	Einddatum	Actief	Opmerking	Dokter	Afdeling
04-04-2022	Diagnoses	hawkinsinurie		- -	Ja		Apo - Alleen Te Gebruiken Door Apo	
24-11-2021	Verwijsreden/klacht	Verwijsreden/klacht		- -	Ja	Zie gescande aanvraag	Testhuisarts Sfvg	
24-11-2021	Verwijsreden/klacht	Verwijsreden/klacht		- -	Ja	Zie gescande aanvraag	Testhuisarts Sfvg	
02-12-2020	Diagnoses	obesitas		- -	Ja		J.C.C.M. in t Veen	Longgeneeskunde
22-05-2018	Diagnoses	slijtage van heupprothese	links	- -	Ja		S.J.J. Boekhorst	Orthopedie
22-05-2018	Diagnoses	slijtage van knieprothese	rechts	- -	Ja	veroorzaakt door slijtage heup	S.J.J. Boekhorst	Orthopedie
14-12-2017	Aandachtspunten	Algemene aandachtspunten		- -	Ja	Patient is bang voor naalden. Verhoogd risico op flauwvallen.		
13-10-2017	Diagnoses	torsio testis		- -	Ja		T.M.A.L. Klem	Chirurgie
20-06-2017	Complicaties	Geen complicatie		- -	Ja		A.J.H. Kerver	Chirurgie
20-06-2017	Diagnoses	longbotinfestatie		- -	Ja		Apo - Alleen Te Gebruiken Door Apo	
31-10-2016	Diagnoses	mononeuropathie		- -	Ja		H.C.T. van Zaanen	Interne Geneeskunde

Figure 7.1 Problem list with diagnoses, complications, and attention notes. Diagnoses with clarifications have info buttons: these are highlighted, underlined and followed by an information icon. Users can click on the info button to view a pop-up with a diagnosis clarification.

Figure 7.2 Example of diagnosis clarification for the diagnosis 'Osteoarthritis of knee'. This clarification defines the supertype osteoarthritis. Users can provide a rating of the clarification from (1) very bad to (7) very good. Below, a warning is provided that the clarification was generated automatically, and questions can be addressed to the clinician. Additionally, information is provided that the feedback is used to improve the clarifications and a link is provided to further information about the research.

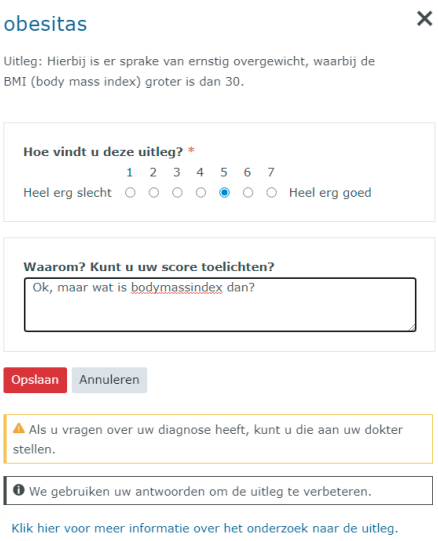


Figure 7.3 Example of diagnosis clarification for the diagnosis 'Obesity'. This clarification consists of a definition from the Dutch SNOMED CT patient-friendly extension. Users can motivate their rating in free text.

7.3.2 Terminologies

Diagnoses on the problem list were encoded by the Diagnosethesaurus ("DT" hereafter; Dutch for "Diagnosis Thesaurus"; DHD, Utrecht, The Netherlands). We used the modified supertype clarifications (see Appendix B) from the April 2020 DT version and the descriptions from the February 2022 DT version. The coverage of the diagnoses by clarifications was 9.4 times higher with generalization compared to the coverage of diagnoses by clarifications based on descriptions from the PFE only. See Table 7.2.

Table 7.2 Diagnosethesaurus coverage

Statistic	n	% of DT	% of SCT
Diagnoses	25,199	100.0	
Diagnoses with SNOMED CT id	22,762	90.3	100.0
Diagnoses with clarifications by synonyms and definitions	1,576	6.3	6.9
Diagnoses with clarifications by generalization	13,304	52.8	58.5
Total diagnoses with clarifications	14,880	59.1	65.4

7.4 Methods

7.4.1 Study Design

We performed a post-implementation evaluation study with the reuse of routinely collected data and prospectively collected quality improvement feedback. The protocol was approved by the university hospital medical ethics review board (reference W21_259 # 21.285), and the local hospital privacy officer and scientific research bureau (reference 2021-109), before commencing the data collection and registered at the ISRCTN[203].

7.4.2 Participants and Study Flow

During the nine-week study period from Monday, April 4th to Monday, June 6th, 2022 all patient portal users were included. We analysed usage data about the logins on the patient portal, problem list views, which diagnoses were displayed when users view their problem list, the number of diagnoses with clarifications and which info buttons were clicked on by users. Users were free to provide feedback or not. Users could log in, view the problem list, display clarifications and provide ratings and feedback multiple times. This thus resulted in a convenience sample with those users that logged in, viewed their problem list, clicked on the info buttons to view the clarifications, and took the effort to provide feedback on the clarifications. We refer to these steps as “conversion steps” and we call the percentage of users that convert from one step (e.g. logging in) to another (e.g. view the problem list) conversion rates.

Conversion steps:

1. Login into the patient portal
2. View the problem list
3. Click on the info button to view the clarification
4. Provide feedback on the clarification

7.4.3 Outcome Measures and Evaluation Criteria

The coverage of the clarifications was measured as the diagnosis clarification recall: the number of diagnoses with a clarification divided by the total number of diagnoses viewed on the problem list. We distinguish this from the problem clarification recall: the number of diagnoses with clarifications divided by the total number of problems (including diagnosis, complications and attention notes) viewed. The use of the clarification functionality was measured as the info button click precision: the number of info buttons clicked on divided by the total number of info buttons viewed. For each conversion step from login to rating the clarifications, we reported the percentage of users that converted to that step, the number of actions (i.e. logins, views, clicks, ratings) they carried out and the number of unique problems, diagnoses and info buttons where the actions were performed on. We aggregated user and diagnosis characteristics for each step, to compare differences between subgroups in the conversion rates. User characteristics were user type (patient user or proxy account user), age group, gender, the patient's latest diagnosis year, and the number of diagnoses. We aggregated diagnoses by DT concept, clarification type and medical specialty.

7.4.4 Data Acquisition and Measurement

We reused EHR and audit trail data to derive which diagnoses were viewed by patients, for which diagnoses users clicked on the info button, and what other actions were taken on the patient portal. We also reused diagnoses, age and gender already registered in the EHR to explore differences in user and diagnosis characteristics.

The two feedback questions asked were simple and minimally invasive: '(1) Please rate this explanation (1. very bad – 7. very good)? (2) Can you motivate your score?' The questionnaire functionality of the EHR was used for this purpose. The feedback was monitored by the hospital staff, to assess whether it contained questions that needed to be addressed or whether any issues arose. The hospital was able to contact the patients to address their questions and where necessary a clarification could have been corrected or removed, or the functionality might have been turned off completely if deemed necessary.

Aggregated data on all patient portal users during the study period were exported from the EHR. To protect the privacy of the patients, variables such as gender and age were aggregated in separate tables so they could not be combined. Free text from the feedback provided was anonymized by an authorized hospital functionary. Anonymization was carried out by removing directly identifying data,

such as dates, places, names of patients, clinicians or others. EHR, audit trail and free-text data were made available by the hospital without any directly identifying personal information.

7.4.5 Statistical Analysis

Conversion rates were calculated and aggregated by the different outcome measure levels. For the number of actions, the interquartile ranges (IQRs) and the maximum number of actions per user were reported. We calculated the diagnosis clarification recall and info button click precision for each patient and took the median and IQR. For the clarification quality ratings, we used the median and interquartile range of the median rating per patient and clarification. The difference in ratings for clarifications with synonyms and definitions compared to clarifications with generalizations to concepts with synonyms and definitions was tested with the Mann-Whitney U test[204]. We analysed the feedback thematically and summarized it narratively. Thematic analysis was carried out by two authors (HJTvM and GEGH) and differences were discussed until consensus was achieved. The Fisher Exact test[205] was used to test differences among users and diagnosis characteristics in the proportions of problem list views and info button clicks. Odds ratios (ORs) were calculated post hoc of each variable to estimate the associations between the characteristics and the views and clicks, comparing the odds of the particular variable (e.g., patients of female gender or with age between 0 to 9) with a reference group (e.g., male gender or age 30 to 39). We took the largest group as the reference group. The p-values were corrected for false discovery rate with the Benjamini Yekutieli method[206]. Data were analysed using the R programming language (R Foundation for Statistical Computing, Vienna, Austria; Version: 4.2.1, 2022-06-23) in RStudio (RStudio Inc., Boston, MA, USA; Version: 2022.07.1). See the R script in Appendix A.

7.5 Results

7.5.1 Demographic and Other Study Coverage Data

In total, for 19,961 patients users had logged in at least once during the nine-week study period. Logins came from all age groups, the largest group logged in for patients in their thirties (18.1%), followed by sixties (17.4%) and fifties (16.5%). Relatively more logins were for women (61.8%) and few users logged in with proxy accounts (0.2%). Table 7.3 shows the user characteristics for each step. Appendix D contains the complete results dataset.

7.5.2 Conversion Rates

Table 7.4 shows the overall conversion and the number of actions carried out for each step. The problem list of 6,530 patients was viewed (32.7% of the patients for whom users had logged in), 2,660 (13.3%) had viewed DT encoded diagnoses on their problem list, and 2,363 (11.8%) had viewed info buttons on their problem list that they could have clicked. Therefore, for 88.8% (2,363 / 2,660) of patients of whom DT encoded diagnoses on their problem list were viewed, an info button was available to view a clarification. When info buttons were available, a median of 1 (IQR: 1 – 2; maximum: 10) info button was on their problem list. The median problem clarification recall was 0.33 (IQR: 0.00 – 0.67). The median diagnosis clarification recall was 1.00 (IQR: 0.75 – 1.00). One or more info buttons were clicked on for diagnoses of 1,291 patients, which is 54.6% of the patients for whom info buttons were viewed and 6.5% of the patients for whom was logged in. On average, users clicked twice (IQR: 1 – 3; maximum: 31) on one info button (IQR: 1 – 1; maximum: 8). The median click precision per patient was 0.50 (IQR: 0.00 – 1.00). Of the patients that clicked on an info button, 108 (8.4%) provided a rating (0.5% of the patients that had logged in).

Table 7.3 User characteristics for users that had logged in, viewed their problem list, viewed info buttons on their problem list, clicked on info buttons on their problem list and provided feedback, with the number of users and percentages of the total number of users that carried out that action.

Statistic	Value	Logged in		Viewed problem list		Viewed info buttons on list		Clicked on info buttons		Provided feedback	
		n	(%)	n	(%)	n	(%)	n	(%)	n	(%)
Age	00 - 09	373	(1.9)	130	(2.0)	54	(2.3)	32	(2.5)	2	(2)
	10 - 19	492	(2.5)	166	(2.5)	58	(2.5)	33	(2.6)	0	(0)
	20 - 29	2,280	(11.4)	764	(11.7)	181	(7.7)	95	(7.4)	6	(6)
	30 - 39	3,612	(18.1)	1,085	(16.6)	265	(11.2)	138	(10.7)	7	(6)
	40 - 49	2,807	(14.1)	929	(14.2)	333	(14.1)	191	(14.8)	17	(15)
	50 - 59	3,284	(16.5)	1,110	(17.0)	442	(18.7)	259	(20.1)	18	(16)
	60 - 69	3,478	(17.4)	1,193	(18.3)	487	(20.6)	264	(20.5)	28	(25)
	70 - 79	2,861	(14.3)	923	(14.1)	417	(17.7)	218	(16.9)	25	(23)
	80 - 89	711	(3.6)	211	(3.2)	114	(4.8)	56	(4.3)	5	(5)
	90 -	63	(0.3)	19	(0.3)	12	(0.5)	5	(0.4)	0	(0)
Age subgroup	00	17	(0.1)	7	(0.1)	2	(0.1)	1	(0.1)	0	(0)
	01 - 09	356	(1.8)	123	(1.9)	52	(2.2)	31	(2.4)	2	(2)
	10 - 11	54	(0.3)	14	(0.2)	3	(0.1)	1	(0.1)	0	(0)
	12 - 15	131	(0.7)	51	(0.8)	24	(1.0)	15	(1.2)	0	(0)
	16 - 17	94	(0.5)	33	(0.5)	13	(0.6)	6	(0.5)	0	(0)
	18 - 19	213	(1.1)	68	(1.1)	18	(0.8)	11	(0.9)	0	(0)
Gender	Male	7,629	(38.2)	2,375	(36.4)	856	(36.2)	471	(36.5)	46	(43)
	Female	12,332	(61.8)	4,155	(63.6)	1,507	(63.8)	820	(63.5)	62	(57)
Account	Proxy	42	(0.2)	12	(0.2)	7	(0.3)	2	(0.2)	0	(0)
	Patient	19,923	(99.8)	6,519	(99.8)	2,357	(99.8)	1,289	(99.9)	108	(100)
Total		**19,961**	**(100.0)**	**6,530**	**(100.0)**	**2,363**	**(100.0)**	**1,291**	**(100.0)**	**108**	**(100)**

Table 7.4 Percentages of logins, views, clicks and ratings, with the number of patients n for whom were logged in, the percentage of patients for whom the problem list was viewed, for whom a problem was viewed on the problem list, for whom diagnoses were viewed, for whom info buttons were viewed, for whom an info button was clicked on and for whom a clarification was rated, of the total number of patients for whom users logged in. Additionally shows the total number of actions, with the quartiles min, 25%, median, 75% and max, and the average number of actions. The number of actions is distinguished from the number of problems or diagnoses that were viewed (i.e. the number of times ratings were provided for which number of diagnoses).

Statistic	Patients		Actions							
	n	%	Level	Total	Min	25%	Median	75%	Max	Average
Logged in	19,961	100.0	Logins	69,112	1	1	2	3	260	3.5
Viewed problem list	6,530	32.7	Views	17,414	1	1	2	3	63	2.7
Viewed problems	3,961	19.8	Views	34,539	1	2	4	9	474	8.7
			Problems	11,145	1	1	2	4	20	2.8
Viewed diagnoses	2,660	13.3	Views	16,012	1	2	3	6	197	6.0
			Diagnoses	4,977	1	1	1	2	16	1.9
Viewed info buttons	2,363	11.8	Views	13,235	1	1	3	6	165	5.6
			Diagnoses	4,069	1	1	1	2	10	1.7
Clicked on info buttons	1,291	6.5	Clicks	2,979	1	1	2	3	31	2.3
			Diagnoses	1,770	1	1	1	1	8	1.4
Rated clarifications	108	0.5	Ratings	133	1	1	1	1	4	1.2
			Diagnoses	127	1	1	1	1	4	1.2

7.5.3 Clarification Quality Ratings

108 users provided feedback on 127 diagnoses (103 unique diagnoses with 95 unique clarifications). Users rated the clarifications with a median of 6 (IQR: 4 – 7), see Figure 7.4. Clarifications with synonyms and definitions were rated higher than clarifications with generalizations to supertypes (median: 6, compared to median: 5.5; p = .0379), see Figure 7.5. Users provided a comment on 66 of the 127 diagnoses (56%). The most common comments were that they found the clarification clear (n = 25; 38%) or incomplete (n = 10; 15%), provided input for improvement (n = 10; 15%), found the clarification unclear (n = 5; 8%), or disagreed with the diagnosis rather than the clarification (n = 4; 6%). Additionally, some users (n ≤ 3; ≤ 5%) commented they recognized the clarification based on their own experience, that they found the clarification was right or useful, asked for a solution for their health problem, disagreed with the treatment, clarification, and/or diagnosis (sometimes not a clear distinction), or mentioned the existence of alternative sources of clarifications.

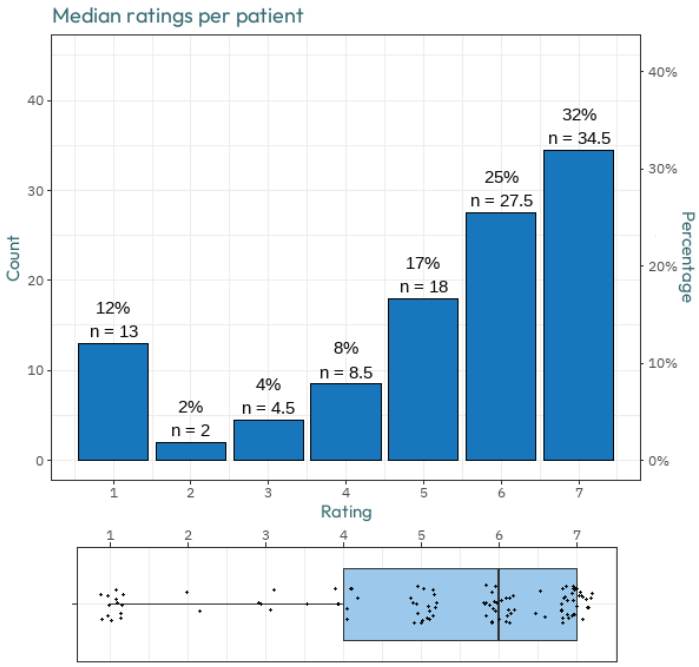


Figure 7.4 Bar plot and box plot of the median ratings per patient. The bar plot shows the median ratings per patient for each rating from (1) very bad to (7) very good and the percentage, and the number of patients n. The box plot below the bar plot shows the median (median: 6), interquartile range (IQR: 4 – 7) and jittered scatter of the ratings.

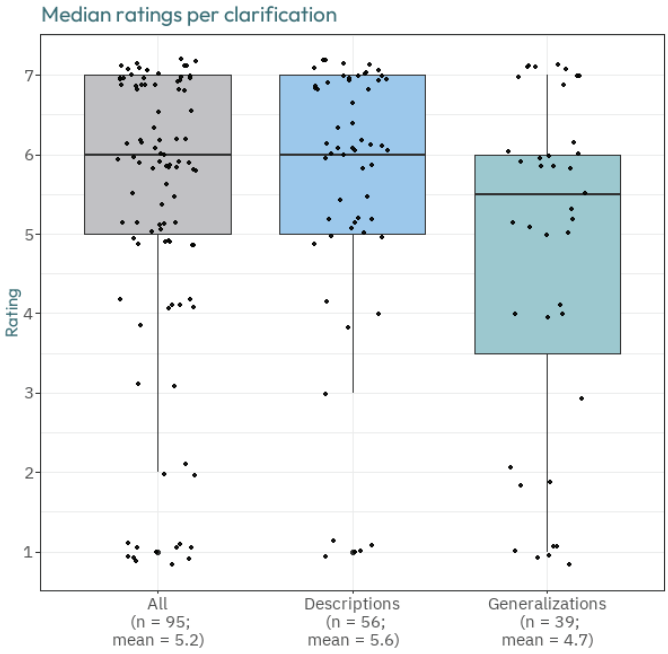


Figure 7.5 Boxplots of ratings from (1) very bad to (7) very good for all clarifications ("All", left), clarifications with patient-friendly synonyms and definitions ("Descriptions", centre) and clarifications with generalization to supertypes with patient-friendly synonyms and definitions ("Generalizations", right).

7.5.4 Differences Between Subgroups

After correcting for the false discovery rate (see Table 7.5 in Appendix C), differences in the proportion of users that viewed the problem list were found significant for gender (p = .0042) and latest diagnosis year (p = .0042). The odds of viewing the problem list were lower for male compared to female patients (OR: 0.89; CI: 0.84 – 0.95) and higher for patients having the latest diagnosis in the year 2022 (when the study was carried out) compared to those having no diagnosis (OR: 1.35; CI: 1.20 – 1.53). Differences in the proportion of users that clicked on an info button were significant for the latest diagnosis year (p = .0003) and medical specialty (p = .0046). The odds of clicking on an info button were higher for patients having the latest diagnosis in 2022 (OR: 3.08; CI: 2.30 – 4.15) and 2021 (OR: 1.33; CI: 1.02 – 1.74) compared to 2020. Compared to orthopaedics, the odds of clicking were lower for ear, nose and throat surgery (OR: 0.78; CI: 0.61 – 0.99), dermatology (OR: 0.56; CI: 0.42 – 0.74), surgery (OR: 0.75; CI: 0.56 – 0.99), ophthalmology (OR: 0.64; 0.48 – 0.85), urology (OR: 0.52; CI: 0.34 – 0.78), plastic surgery (OR: 0.51; CI: 0.28 – 0.93) and gynaecology (OR:

0.51; CI: 0.26 – 0.98). See Appendix C for the proportions and odds ratios of the subgroup variables.

7.5.5 Unexpected Observations

During monitoring, we noticed two events that were not expected. One user (rating: 5) wrote '*I have this pain already for [x] years, why can they not do anything about it, life keeps getting more unbearable*'. The hospital verified whether the patient required follow-up, but there already was a follow-up scheduled. Therefore, it was decided that further action was not necessary. In a second case (rating: 1), a user commented he or she did not have the diagnosis and that this was confirmed by the clinician.

7.6 Discussion

This study provided insight into patient portal user information needs by measuring and evaluating the actual coverage and use of a clarification functionality for the problem list. The coverage of diagnoses by clarifications was high, with almost ninety percent of patients having clarifications for one or more diagnoses on their problem list. More than half of the users that could use the info buttons clicked on them during the study period and on average they clicked on half of the info buttons available in their problem list. Overall, clarifications were rated as having good quality. Clarifications by synonyms and definitions of supertypes were rated relatively lower than clarifications with synonyms and definitions of the diagnoses themselves. The odds that the problem list was viewed were relatively higher for patients of the female gender and with a more recent diagnosis. The odds that info buttons were clicked to view clarifications were relatively higher for patients with a recent diagnosis and relatively lower (compared to orthopaedics) for diagnoses from the specialties ear, nose and throat surgery, dermatology, surgery, ophthalmology, urology, plastic surgery and gynaecology.

Similar studies have not carried out an evaluation study in clinical practice, but relied on online surveys[173], laboratory situations[173,195,207], or only carried out expert evaluation[191,208,209]. Additionally, previous studies did not use personal medical data and were focused on notes, rather than encoded diagnoses. Therefore, the current study is novel in that we prospectively evaluated clarifications in a real-time patient portal with patients' personal medical data, showing that end-users use and appreciate clarification functionality. Patients have been reported to find errors in their notes and to consider some medical record content to be judgmental and offensive[5,210]. It appears that the clarification helps users to

verify whether the diagnosis is correct, as our second example in section 7.5.5 illustrates. Some authors[5,6] argue that medical jargon should be replaced by language that treats patients less belittling, passive, childish and blameable. The evaluated solution in the present study, however, does not require clinicians to change the way they register their data. It combines the strength of more professional phrasing, as the content was already encoded with terminology systems, with clarification by the functionality.

To our knowledge, this is the first study that evaluates clarifications in a patient portal. Reusing existing log and EHR data provides a more representative picture of users and their behaviour than making patients or laymen fill out surveys and using fabricated non-personal data[173,195,207], as we were able to include a wide variety of users in the convenience samples of each conversion step. The brief quality ratings were minimally invasive for end users. Some users disagreed with the diagnosis and one with their treatment, and accordingly rated the clarification as very bad. Conversely, a user commented that the clarification was a good addition to the drawing a clinician made and rated the clarification as very good. Where users did not comment, we could not verify whether they based the rating on the clarification only or also on the diagnosis or experience with their clinician. This might affect the ratings and the ratings thus reflect a mix of the quality of the clarification, the data quality and the experience with the clinician. Without the permission of the users, we could not obtain individual patient data to run a multivariate model. Therefore, this research was limited to aggregate data and associations could hence not be corrected for confounders. The aggregate data provided insight into different user groups. However, the few differences in conversion we found between users and diagnoses were based on sample sizes that lowered along the conversion steps. Differences might have resulted coincidentally due to multiple testing and confounding. We tried to minimize the false discovery rate and might have unnecessarily discarded associations such as age and problem list viewing (e.g. the problem list appear to be viewed significantly more often for patients in their thirties compared to patients in their sixties). However, we still were able to provide some insight for further studies with a rich descriptive dataset.

This study shows that generalization is a useful technique to generate clarifications from the perspective of actual patient portal users. For terminology developers, the approach has the potential to make more maintainable terms and definitions, that can be reused among several medical concepts. In further research, tailoring clarifications to end-users, especially on a more accessible language difficulty

level, and developing clarifications for particular diagnosis classes should be investigated, improving the clarifications and functionality. The coverage of the current system can be increased by updating the terminology versions, developing clarifications for other types of medical data and applying other clarification methods, such as using relationships other than is-a relationships in SNOMED CT, such as the finding site (e.g. pancreas) and associated morphology (e.g. inflammation) to clarify concepts (e.g. deriving 'inflammation of the pancreas' from 'pancreatitis'). The associations found indicate that there are differences in usage between groups, which might reflect that they have different information needs. The unexpected observations imply that asking for feedback about diagnoses should also involve follow-up, as patients sometimes do not understand or agree with the diagnosis. The hospital decided to continue showing the clarifications after the study period, but without asking for free-text feedback, because there was no solution yet for continuing follow-up and free-text anonymization to share the feedback for clarification quality improvement. Healthcare institutions should determine how to deal with these issues before implementing such functionality, as user input can help improve medical record accuracy and clarification quality.

7.7 Conclusion

The coverage of diagnoses by clarifications based on an algorithm that generalizes diagnoses to concepts with patient-friendly terms and definitions was high and the majority of users used the clarification functionality. Overall, users considered it good clarifications, but they also identified opportunities for improving the clarity and completeness of some clarifications. Future research should address the improvement of the clarification coverage and quality, and further investigate differences between subgroups to assess specific user group needs and prioritize areas of improvement.

7.8 Clinical Relevance Statement

While medical data had traditionally been registered for clinical purposes and clinicians only, patients – who often have not had any medical training – currently access their health records. This study presents a generic solution to make medical data, in particular diagnoses, more understandable for patients, without creating an additional administrative burden for clinicians, because clarifications are provided to data that already are routinely registered in health records. The functionality is used and appreciated by patient portal users.

7.9 Acknowledgements

Thanks to Mark de Keizer, Thabitha Scheffers-Buitendijk, Savine Martens, Daniël Huliselan, Theo van Mens, and Leonieke van Mens for their contributions to this study.

7.10 Protection of Human and Animal Subjects

The study was performed in compliance with the World Medical Association Declaration of Helsinki on Ethical Principles for Medical Research Involving Human Subjects. A waiver from the Medical Research Ethics Committee of Amsterdam UMC, location AMC was obtained on June 3, 2021, and filed under reference number W21_259 # 21.285. It confirmed that the Medical Research Involving Human Subjects Act (in Dutch: "WMO") does not apply to the study and that therefore an official approval of this study by the ethics committee was not required under Dutch law. Approval from the Data Protection Officer and the Scientific Bureau of Franciscus was obtained on March 10, 2022 (reference 2021-109). The final research protocol was registered at the ISRCTN[203]. No consent for publication was required because no individual person's data was published.

7.11 Appendices

7.11.1 Appendix A

Appendix A Script.R will be available in the online version of this paper.

7.11.2 Appendix B

Appendix B Modifications

To minimize the errors and clarifications considered to be unacceptable that we had identified in our validation study we made 219 modifications in the patient-friendly extension, thus resulting in a local extension that included the modifications. We excluded 13 concepts and 22 descriptions from the extension, and excluded 20 diagnoses from the Diagnosethesaurus, accepted 2 descriptions, preferred 23 descriptions, added 20 synonyms and 73 text definitions, and replaced 8 synonyms and 38 text definitions. Additionally, we decided to provide the same general clarification for all metastases (based on the September 2021 edition), to prevent problems with metastases clarifications. The coverage of clarifications by supertypes had dropped from 86.79% to 83.13% because of the modifications (e.g. excluding the clarification 'a type of disease')

and dropped further to 65.99%, because of newly introduced concepts in the Diagnosethesaurus.

7.11.3 Appendix C

Appendix C Tables

7.11.3.1 Differences in conversion per user characteristic and diagnosis

Table 7.5 P-values of the Fisher Exact test and adjusted p-values after Benjamini Yekutieli correction for the false discovery rate, testing whether the differences in proportions of users that viewed their problem list and that clicked on an info button are different

Level	Variable	Viewed differences		Clicked differences	
		p-value	adjusted p	p-value	adjusted p
Patient	Age group	.0063	.0567	.4734	1
	Age subgroup	.5644	1	.8607	1
	Gender	.0002	.0042	.7965	1
	User type	.6249	1	.2561	1
	Latest diagnosis year	.0002	.0042	< .0001	.0003
	Number of diagnoses	.0094	.0728	.0458	.2753
Diagnosis	Diagnosis concept	.5228	1	.0048	.0523
	Medical specialty	.0440	.2753	.0003	.0046

7.11.3.2 Conversion per user characteristic

Table 7.6 Logins and problem list views per user characteristic with number of users that logged in l, the percentage of total number of patients for whom users logged in (19,961), the number of users that viewed their problem list v, with the percentage of users that logged in that viewed their problem list, the odds ratio of viewing the problem list after logging in, with 95% confidence intervals, p-value of the Fisher Exact test, and * indicating whether the odds the direction of the odds (+ or -) or whether the category was the reference category (indicated with "ref").

Characteristic	Category	Logged in		Viewed problem list						
		l	% l	v	% v	OR	(95% CI)	P-value	*	ref
Age	00 - 09	373	1.9	130	34.9	1.25	(0.99, 1.57)	.0587		
	10 - 19	492	2.5	166	33.7	1.19	(0.96, 1.45)	.0951		
	20 - 29	2,280	11.4	764	33.5	1.17	(1.05, 1.32)	.0057	+	
	30 - 39	3,612	18.1	1,085	30.1	-	- -	-		ref
	40 - 49	2,807	14.1	929	33.1	1.15	(1.03, 1.28)	.0092	+	
	50 - 59	3,284	16.5	1,110	33.8	1.19	(1.07, 1.32)	.0008	+	
	60 - 69	3,478	17.4	1,193	34.3	1.22	(1.10, 1.35)	.0001	+	
	70 - 79	2,861	14.3	923	32.3	1.11	(1.00, 1.23)	.0583		
	80 - 89	711	3.6	211	29.7	0.98	(0.82, 1.18)	.8580		
	≥ 90	63	0.3	19	30.2	1.01	(0.55, 1.77)	1		
Age subgroups	00	17	0.1	7	41.2	1.32	(0.42, 3.97)	.6072		
	01 - 09	356	1.8	123	34.6	-	- -	-		ref
	10 - 11	54	0.3	14	25.9	0.66	(0.32, 1.30)	.2782		
	12 - 15	131	0.7	51	38.9	1.21	(0.78, 1.86)	.3944		
	16 - 17	94	0.5	33	35.1	1.02	(0.61, 1.69)	.9038		
	18 - 19	213	1.1	68	31.9	0.89	(0.61, 1.29)	.5822		

Table 7.6 (continued)

Characteristic	Category	Logged in		Viewed problem list						
		l	% l	v	% v	OR	(95% CI)	P-value	*	ref
Gender	Male	7,629	38.2	2,375	31.1	0.89	(0.84, 0.95)	.0002	-	
	Female	12,332	61.8	4,155	33.7	-	- -	-		ref
Account	Proxy	42	0.2	12	28.6	0.82	(0.38, 1.66)	.6249		
	Patient	19,923	99.8	6,519	32.7	-	- -	-		ref
Diagnoses	≥ 2	3,288	16.5	1,124	34.2	1.11	(1.02, 1.20)	.0139	+	
	1	4,638	23.2	1,567	33.8	1.09	(1.01, 1.17)	.0210	+	
	0	12,032	60.3	3,839	31.9	-	- -	-		ref
Latest diagnosis year	2022	1,275	6.4	495	38.8	1.35	(1.20, 1.53)	< .0001	+	
	2021	1,457	7.3	490	33.6	1.08	(0.96, 1.22)	.1814		
	2020	1,599	8.0	517	32.3	1.02	(0.91, 1.14)	.7107		
	2019	1,421	7.1	477	33.6	1.08	(0.96, 1.22)	.1961		
	2018	984	4.9	328	33.3	1.07	(0.93, 1.23)	.3745		
	≤ 2017	1,193	6.0	384	32.2	1.01	(0.89, 1.15)	.8452		
	No diagnosis	12,032	60.3	3,839	31.9	-	- -	-		ref
Total		**19,961**	**100.0**	**6,530**	**32.7**					

Table 7.7 Number of users with certain user characteristics that viewed info buttons and clicked on info buttons, with percentages and 95% confidence intervals (CI). The percentages and CIs concern the number of users that viewed info buttons compared to the number that viewed the problem list, the number of users that clicked on an info button compared to the number of users that viewed an info button and the number of users that provided a rating (answer) compared to the number of users that clicked on an info button. The columns with the asterisks * show whether the CI is significantly differs from (whether it is lower or higher than) the CI of the total number of info button views, clicks or answers.

Characteristic	Value	Viewed info button		Clicked on info button						
		v	% v	c	% c	OR	(95% CI)	P-value	*	ref
Age	00 - 09	54	41.54	32	59.26	1.23	(0.67, 2.29)	.5649		
	10 - 19	58	34.94	33	56.90	1.11	(0.62, 2.02)	.7807		
	20 - 29	181	23.69	95	52.49	0.93	(0.65, 1.33)	.7272		
	30 - 39	265	24.42	138	52.08	0.92	(0.67, 1.25)	.5927		
	40 - 49	333	35.84	191	57.36	1.14	(0.85, 1.52)	.3910		
	50 - 59	442	39.82	259	58.60	1.20	(0.91, 1.56)	.1858		
	60 - 69	487	40.82	264	54.21	-	- -	-		ref
	70 - 79	417	45.18	218	52.28	0.93	(0.71, 1.21)	.5928		
	80 - 89	114	54.03	56	49.12	0.82	(0.53, 1.25)	.3489		
	≥ 90	12	63.16	5	41.67	0.60	(0.15, 2.25)	.5597		
Age subgroups	00	2	28.57	1	50.00	0.68	(0.01, 55.66)	1		
	01 - 09	52	42.28	31	59.62	-	- -	-		ref
	10 - 11	3	21.43	1	33.33	0.35	(0.01, 7.03)	.5652		
	12 - 15	24	47.06	15	62.50	1.13	(0.38, 3.50)	1		
	16 - 17	13	39.39	6	46.15	0.59	(0.14, 2.36)	.5327		
	18 - 19	18	26.47	11	61.11	1.06	(0.31, 3.80)	1		
Gender	Male	856	36.04	471	55.02	1.02	(0.86, 1.22)	.7965		
	Female	1,507	36.27	820	54.41	-	- -	-		ref

Table 7.7 (continued)

Characteristic	Value	Viewed info button		Clicked on info button						
		v	% v	c	% c	OR	(95% CI)	P-value	*	ref
Account	Proxy	7	58.33	2	28.57	0.33	(0.03, 2.03	.2561		
	Patient	2,357	36.16	1,289	54.69	-	- -	-		ref
Diagnoses	≥ 2	1,094	97.33	622	56.86	1.18	(1.00, 1.40)	.0467	+	
	1	1,269	80.98	669	52.72	-	- -	-		ref
Latest diagnosis year	2022	426	86.06	323	75.82	3.08	(2.30, 4.15)	< .0001	+	
	2021	449	91.63	258	57.46	1.33	(1.02, 1.74)	.0348	+	
	2020	476	92.07	240	50.42	-	- -	-		ref
	2019	420	88.05	205	48.81	0.94	(0.71, 1.23)	.6399		
	2018	276	84.15	120	43.48	0.76	(0.55, 1.03)	.0695		
	≤ 2017	316	82.29	145	45.89	0.83	(0.62, 1.12)	.2179		
Total		**2,363**	**36.19**	**1,291**	**54.63**					

7.11.3.3 Conversion per medical specialty

Table 7.8 Logins and problem list views of patients with one or more diagnoses per medical specialty with number of users that logged in or viewed their problem list with percentages and 95% confidence intervals (CI). The percentage of logins concerns the percentage of the sum of the total number patients with one or more diagnoses per specialty (12,100), while the percentage of views concerns the percentage of users that viewed their problem list of the total number of users with that characteristic that logged in. The column with an asterisk * shows whether the CI is significantly lower or higher than the CI of the total number of views.

Specialty	Logged in		Viewed problem list						
	l	% l	v	% v	OR	(95% CI)	P-value	*	ref
Orthopaedics	2,033	100.0	701	34.5	-	- -	-		ref
Ear, nose and throat surgery	1,636	16.8	570	34.8	1.02	(0.88, 1.17)	.8342		
Ophthalmology	1,359	13.5	433	31.9	0.89	(0.77, 1.03)	.1188		
Surgery	1,086	11.2	358	33.0	0.93	(0.80, 1.10)	.4048		
Dermatology	1,037	9.0	370	35.7	1.05	(0.90, 1.24)	.5219		
Neurology	777	8.6	260	33.5	0.96	(0.80, 1.14)	.6250		
Internal medicine	541	6.4	161	29.8	0.81	(0.65, 0.99)	.0404	-	
Rheumatology	511	4.5	172	33.7	0.96	(0.78, 1.19)	.7546		
Cardiology	501	4.2	161	32.1	0.90	(0.73, 1.11)	.3434		
Urology	486	4.1	147	30.2	0.82	(0.66, 1.03)	.0781		
Plastic surgery	363	4.0	107	29.5	0.79	(0.62, 1.02)	.0705		
Pulmonary medicine	356	3.0	133	37.4	1.13	(0.89, 1.44)	.3057		
Gastrointestinal and liver diseases	271	2.9	99	36.5	1.09	(0.83, 1.43)	.4980		
Paediatrics	265	2.2	94	35.5	1.04	(0.79, 1.37)	.7837		
Gynaecology	169	2.2	59	34.9	1.02	(0.72, 1.43)	.9330		
Revalidation	160	1.4	66	41.2	1.33	(0.95, 1.87)	.0857		

Table 7.8 (continued)

Specialty	Logged in		Viewed problem list						
	l	% l	v	% v	OR	(95% CI)	P-value	*	ref
Psychiatry	153	1.3	53	34.6	1.01	(0.70, 1.44)	1		
Pain medicine	146	1.3	61	41.8	1.36	(0.95, 1.94)	.0875		
Anaesthesiology	132	1.2	37	28.0	0.74	(0.49, 1.11)	.1549		
Geriatrics	109	1.1	31	28.4	0.76	(0.48, 1.17)	.2143		
Intensive care	5	0.9	3	60.0	2.85	(0.33, 34.18)	.3477		
Neurosurgery	2	0.0	1	50.0	1.90	(0.02, 149.09)	1		
Nephrology	1	0.0	0	0.0					
Obstetrics	1	0.0	0	0.0					

Table 7.9 Number, percentage and confidence interval of patient account users that viewed info buttons, clicked on info buttons and answered the questionnaire per medical specialty of the diagnoses registered. The percentages and confidence intervals (CIs) concern the number of users that viewed info buttons compared to the number that viewed the problem list, the number of users that clicked on an info button compared to the number of users that viewed an info button and the number of users that provided a rating (answer) compared to the number of users that clicked on an info button. The columns with the asterisks * show whether the CI is significantly differs from (whether it is lower or higher than) the CI of the total number of info button views, clicks or answers.

Specialty	Viewed info button		Clicked on info button						
	v	% v	c	% c	OR	(95% CI)	P-value	*	ref
Orthopaedics	637	90.9	343	53.8	-	- -	-		ref
Ear, nose and throat surgery	527	92.5	251	47.6	0.78	(0.61, 0.99)	.0392	-	
Dermatology	318	85.9	125	39.3	0.56	(0.42, 0.74)	< .0001	-	
Surgery	305	85.2	142	46.6	0.75	(0.56, 0.99)	.0369	-	
Ophthalmology	300	69.3	128	42.7	0.64	(0.48, 0.85)	.0016	-	
Neurology	220	84.6	101	45.9	0.73	(0.53, 1.00)	.0503		
Rheumatology	156	90.7	73	46.8	0.75	(0.52, 1.09)	.1284		
Internal Medicine	139	86.3	70	50.4	0.87	(0.59, 1.28)	.5115		
Urology	127	86.4	48	37.8	0.52	(0.34, 0.78)	.0013	-	
Cardiology	117	72.7	66	56.4	1.11	(0.73, 1.69)	.6156		
Pulmonary medicine	115	86.5	53	46.1	0.73	(0.48, 1.11)	.1293		
Gastrointestinal and liver diseases	97	98.0	47	48.5	0.81	(0.51, 1.26)	.3280		
Paediatrics	73	77.7	32	43.8	0.67	(0.40, 1.12)	.1092		
Revalidation	60	90.9	36	60.0	1.29	(0.73, 2.31)	.4166		
Plastic surgery	56	52.3	21	37.5	0.51	(0.28, 0.93)	.0249	-	
Gynaecology	48	81.4	18	37.5	0.51	(0.26, 0.98)	.0353	-	

Specialty	Viewed info button			Clicked on info button					
	v	% v	c	% c	OR	(95% CI)	P-value	*	ref
Pain medicine	43	70.5	25	58.1	1.19	(0.61, 2.37)	.6373		
Anaesthesiology	31	83.8	11	35.5	0.47	(0.20, 1.05)	.0639		
Geriatrics	18	58.1	13	72.2	2.23	(0.73, 8.07)	.1523		
Psychiatry	6	11.3	3	50.0	0.86	(0.11, 6.45)	1		
Intensive care	3	100.0	2	66.7	1.71	(0.09, 101.42)	1		
Neurosurgery	1	100.0	1	100.0					
Nephrology	0	0.0	0	0					
Obstetrics	0	0.0	0	0					

7.11.3.4 Conversion per diagnosis

Table 7.10 Top 20 of patient account users with a certain diagnosis that logged in, with the SNOMED CT concept Id, diagnosis description (SNOMED CT preferred term) and clarification, the number of diagnoses for which patients logged in, percentage of the sum of the number of patients that logged per diagnosis (13,939) and 95% confidence intervals.

Id	Diagnosis	Clarification	Logged in		Viewed problem list						
			l	%	v	%	OR	(95% CI)	P-value	*	ref
239873007	Osteoarthritis of knee	Osteoarthritis: This is a chronic joint disease in which alteration of the cartilage in the joints leads to pain and stiffness.	214	1.5	61	28.5	-	- -	-		ref
442520000	Inflammation of rotator cuff tendon	A condition of tendon and shoulder.	185	1.3	71	38.4	1.56	(1.00, 2.43)	.0427	+	
267103008	Feeling of lump in throat (globus feeling in Dutch)	Another word for “globus feeling” is lump in the throat: Feeling that there is a slimy plug in the throat that cannot be swallowed.	139	1.0	58	41.7	1.79	(1.12, 2.88)	.0114	+	
21522001	Abdominal pain	Nagging, cramping or stabbing pain in the abdominal area.	133	0.9	58	43.6	1.94	(1.20, 3.13)	.0052	+	
239872002	Osteoarthritis of hip	Osteoarthritis: This is a chronic joint disease in which alteration of the cartilage in the joints leads to pain and stiffness.	133	0.9	45	33.8	1.28	(0.78, 2.10)	.3378		
57406009	Carpal tunnel syndrome	Disease in the wrist that results from putting too much pressure on a major nerve that runs through the wrist canal.	121	0.9	42	34.7	1.33	(0.80, 2.21)	.2677		
39450006	Senile cataract	Another word for “senile cataract” is elderly cataract: Lens clouding, caused by aging.	119	0.8	33	27.7	0.96	(0.56, 1.63)	.8997		
16491000146102	Injury of medial meniscus structure	A disease of cartilage and joint.	104	0.7	44	42.3	1.84	(1.09, 3.08)	.0160	+	

Table 7.10 (continued)

Id	Diagnosis	Clarification	Logged in		Viewed problem list						
			l	%	v	%	OR	(95% CI)	P-value	*	ref
72866009	Varicose veins of lower extremity	A type of varicose veins.	91	0.6	35	38.5	1.57	(0.90, 2.70)	.1057		
25071000146105	Osteoarthritis of medial compartment of knee	Osteoarthritis: This is a chronic joint disease in which alteration of the cartilage in the joints leads to pain and stiffness.	84	0.6	28	33.3	1.25	(0.70, 2.22)	.4819		
254837009	Malignant neoplasm of breast	Another word for “malignant neoplasm of breast” is breast cancer: A malignant tumour in the breast.	80	0.6	23	28.7	1.01	(0.55, 1.84)	1		
78275009	Obstructive sleep apnoea syndrome	Another word for “obstructive sleep apnoea syndrome” is OSA or OSAS: Temporary stop of breathing during sleep caused by the airway becoming blocked by the tongue, the soft part of the palate or the wall of the pharynx.	79	0.6	27	34.2	1.3	(0.72, 2.33)	.3894		
239865003	Osteoarthritis of acromioclavicular joint	A disease of the shoulder. Osteoarthritis: This is a chronic joint disease in which alteration of the cartilage in the joints leads to pain and stiffness.	78	0.6	26	33.3	1.25	(0.69, 2.26)	.4702		
195967001	Asthma	Attacks of breathlessness and shortness of breath, caused by narrowing of the smaller airways.	74	0.5	31	41.9	1.8	(1.00, 3.24)	.0426	+	
39898005	Sleep disorder	Collective term for disorders in which sleep is disturbed or in which abnormal behaviour occurs during sleep.	74	0.5	21	28.4	0.99	(0.52, 1.84)	1		

Table 7.10 (continued)

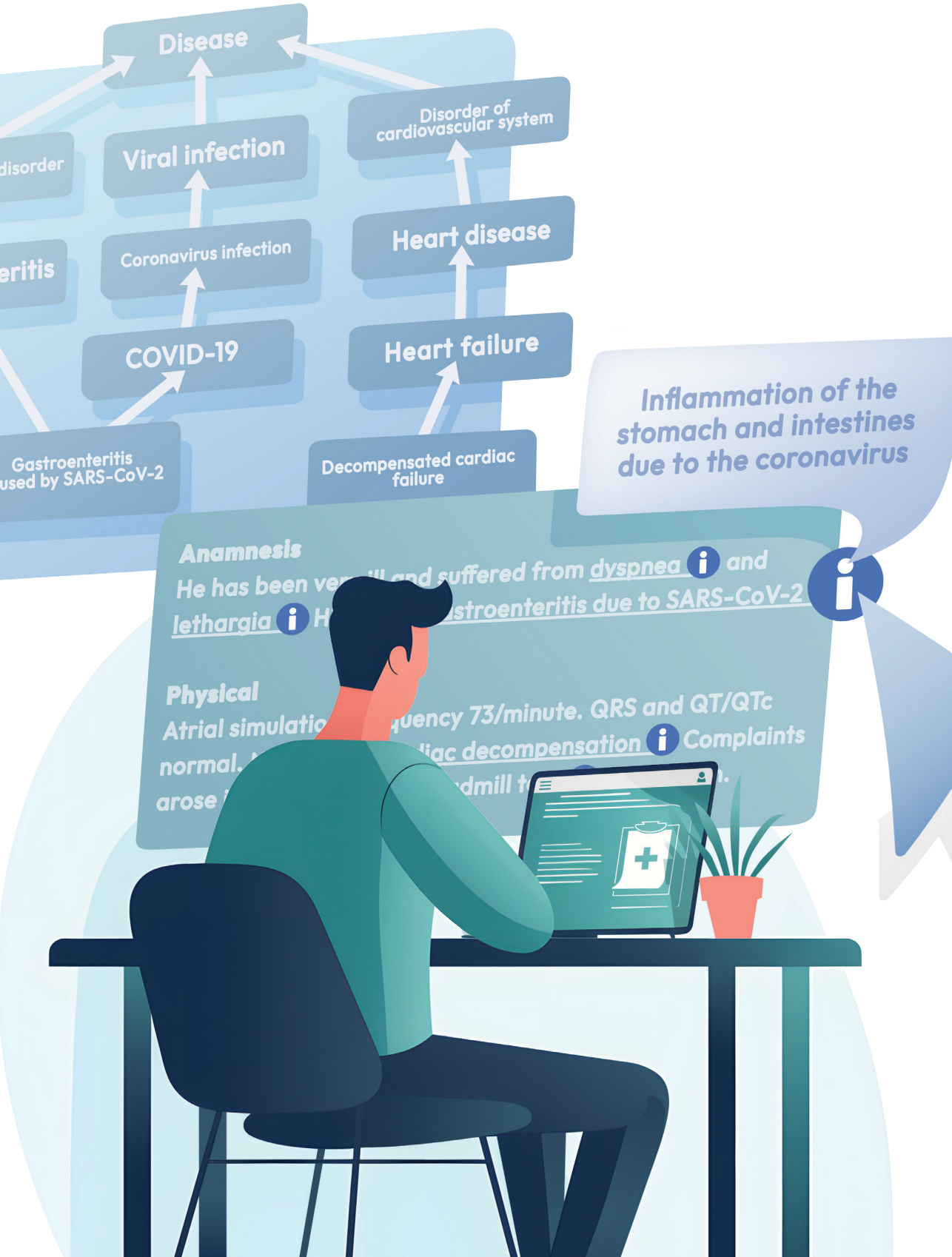
Id	Diagnosis	Clarification	Logged in		Viewed problem list						
			l	%	v	%	OR	(95% CI)	P-value	*	ref
202842005	Partial thickness rotator cuff tear	A disease of the shoulder.	72	0.5	18	25.0	0.84	(0.43, 1.59)	.6485		
49526009	Presbycusis	A type of hearing loss.	72	0.5	19	26.4	0.9	(0.46, 1.70)	.7639		
254701007	Basal cell carcinoma of skin	Another word for “basal cell carcinoma of skin” is basal cell cancer: Form of skin cancer, usually in areas of the skin exposed to the sun; rarely metastasizes.	71	0.5	18	25.4	0.85	(0.43, 1.62)	.6491		
26241000146104	Primary osteoarthritis of knee	Osteoarthritis: This is a chronic joint disease in which alteration of the cartilage in the joints leads to pain and stiffness.	71	0.5	22	31.0	1.13	(0.60, 2.09)	.7633		
430725003	Patellofemoral stress syndrome	Irritation that can occur between the kneecap and the knee joint because the kneecap slides over the joint when bending and stretching; gives pain at the kneecap, especially when squatting, climbing stairs or cycling; mainly occurs in teenagers and young adults.	69	0.5	22	31.9	1.17	(0.62, 2.18)	.6487		
60862001	Tinnitus	Another word for “tinnitus” is ringing in the ears: Hearing sounds (such as humming, whistling and wind noise) that do not come from outside.	69	0.5	23		1.25	(0.67, 2.32)	.4520		

Table 7.11 Patient account users with a certain diagnosis that viewed a diagnosis with an info button, clicked on it and answered the questionnaire, with the numbers. The columns with the asterisks * show whether the CI is significantly differs from (whether it is lower or higher than) the CI of the total number of info button views, clicks or answers.

Id	Diagnosis	Viewed info button		Clicked on info button						
		v	% v	c	% c	OR	95% CI	P-value	*	ref
239873007	Osteoarthritis of knee	214	1.5	61	28.5	-	- -	-		ref
442520000	Inflammation of rotator cuff tendon	185	1.3	71	38.4	1.56	(1.00, 2.43)	.0427	+	
267103008	Feeling of lump in throat	139	1.0	58	41.7	1.79	(1.12, 2.88)	.0114	+	
21522001	Abdominal pain	133	0.9	58	43.6	1.94	(1.20, 3.13)	.0052	+	
239872002	Osteoarthritis of hip	133	0.9	45	33.8	1.28	(0.78, 2.10)	.3378		
57406009	Carpal tunnel syndrome	121	0.9	42	34.7	1.33	(0.80, 2.21)	.2677		
39450006	Senile cataract	119	0.8	33	27.7	0.96	(0.56, 1.63)	.8997		
16491000146102	Injury of medial meniscus structure	104	0.7	44	42.3	1.84	(1.09, 3.08)	.0160	+	
72866009	Varicose veins of lower extremity	91	0.6	35	38.5	1.57	(0.90, 2.70)	.1057		
25071000146105	Osteoarthritis of medial compartment of knee	84	0.6	28	33.3	1.25	(0.70, 2.22)	.4819		
254837009	Malignant neoplasm of breast	80	0.6	23	28.7	1.01	(0.55, 1.84)	1		
78275009	Obstructive sleep apnoea syndrome	79	0.6	27	34.2	1.3	(0.72, 2.33)	.3894		
239865003	Osteoarthritis of acromioclavicular joint	78	0.6	26	33.3	1.25	(0.69, 2.26)	.4702		
195967001	Asthma	74	0.5	31	41.9	1.8	(1.00, 3.24)	.0426	+	
39898005	Sleep disorder	74	0.5	21	28.4	0.99	(0.52, 1.84)	1		
202842005	Partial thickness rotator cuff tear	72	0.5	18	25.0	0.84	(0.43, 1.59)	.6485		
49526009	Presbycusis	72	0.5	19	26.4	0.9	(0.46, 1.70)	.7639		
254701007	Basal cell carcinoma of skin	71	0.5	18	25.4	0.85	(0.43, 1.62)	.6491		
26241000146104	Primary osteoarthritis of knee	71	0.5	22	31.0	1.13	(0.60, 2.09)	.7633		
430725003	Patellofemoral stress syndrome	69	0.5	22	31.9	1.17	(0.62, 2.18)	.6487		
60862001	Tinnitus	69	0.5	23		1.25	(0.67, 2.32)	.4520		

7.11.4 Appendix D

Appendix D Dataset.xlsx will be available in the online version of this paper.


Disease
Viral infection
Disorder of cardiovascular system
Coronavirus infection
Heart disease
COVID-19
Heart failure
Gastroenteritis used by SARS-CoV-2
Decompensated cardiac failure
Inflammation of the stomach and intestines due to the coronavirus
Anamnesis
dyspnea
and
lethargia
Physical
Complaints

8

General discussion

This thesis aimed to provide an overview of the determinants and outcomes of patients accessing their EHRs and to develop and evaluate applications using medical terminology systems to clarify medical data to patients.

8.1 Principal findings

Using the clinical adoption framework (CAF), we synthesized the results from systematic reviews on patient access to electronic health records (EHRs), through patient portals and personal health records, which was described in **Chapter 2**. Patient portals and PHRs are types of health information systems (HISs). The reviews reported on outcomes of patient access to EHRs regarding care quality, productivity and healthcare access. The determinants of HIS quality and HIS use were widely reported, as well as associations with people, organization and implementation dimensions of the CAF. There were fewer reviews addressing healthcare standards, health data infrastructures, legislation, policy, governance, funding, incentives, and social, economic and political trends. In **Chapter 3**, we further elaborated on the CAF to include patients as end-users of HISs in the evaluation of HIS adoption and to develop more explicit guidance documentation for the application of the CAF.

In **Chapter 4**, we evaluated lexical clarification functionality in a patient portal that provided consumer-oriented definitions from the Thesaurus Zorg en Welzijn (TZW) to medical terms in patients' clinical notes. We invited patients to identify terms that they did not understand or for which they wanted to view a clarification. Fifteen patients with adequate health literacy marked about ten terms as difficult per note. The functionality, however, only provided clarifications for less than one of those terms per note, and for many other terms that were not considered difficult by the patients. Despite this low precision and recall of the functionality, most participants found the functionality easy to use and useful.

In **Chapter 5** we proposed an algorithm to generate clarifications for diagnoses encoded with the Diagnosethesaurus. The clarifications were generated by generalizing them to concepts with patient-friendly terms from the SNOMED CT Netherlands patient-friendly reference set. We found that this increases the coverage of diagnoses by clarifications from 1.2% to 71% (a 59 times increase). We assessed the medical validity of the clarifications and identified problems in **Chapter 6**. We found that the room for improvement was mostly related to the patient-friendly reference set, terminology mappings and terminology modelling, but also the algorithm. We modified the patient-friendly reference set and the

clarifications to largely resolve the identified problems and implemented the clarifications in a hospital patient portal, which we reported in **Chapter 7**. By including generalizations in addition to clarifications with synonyms and definitions, we achieved a significantly larger diagnosis clarification precision, although generalizations were rated slightly, but significantly lower compared to clarifications with synonyms and definitions. The majority of patients viewed the clarifications when they could and they considered most of the clarifications to be of good quality. Figure 8.1 summarizes the research carried out on clarifying medical data into the CAF.

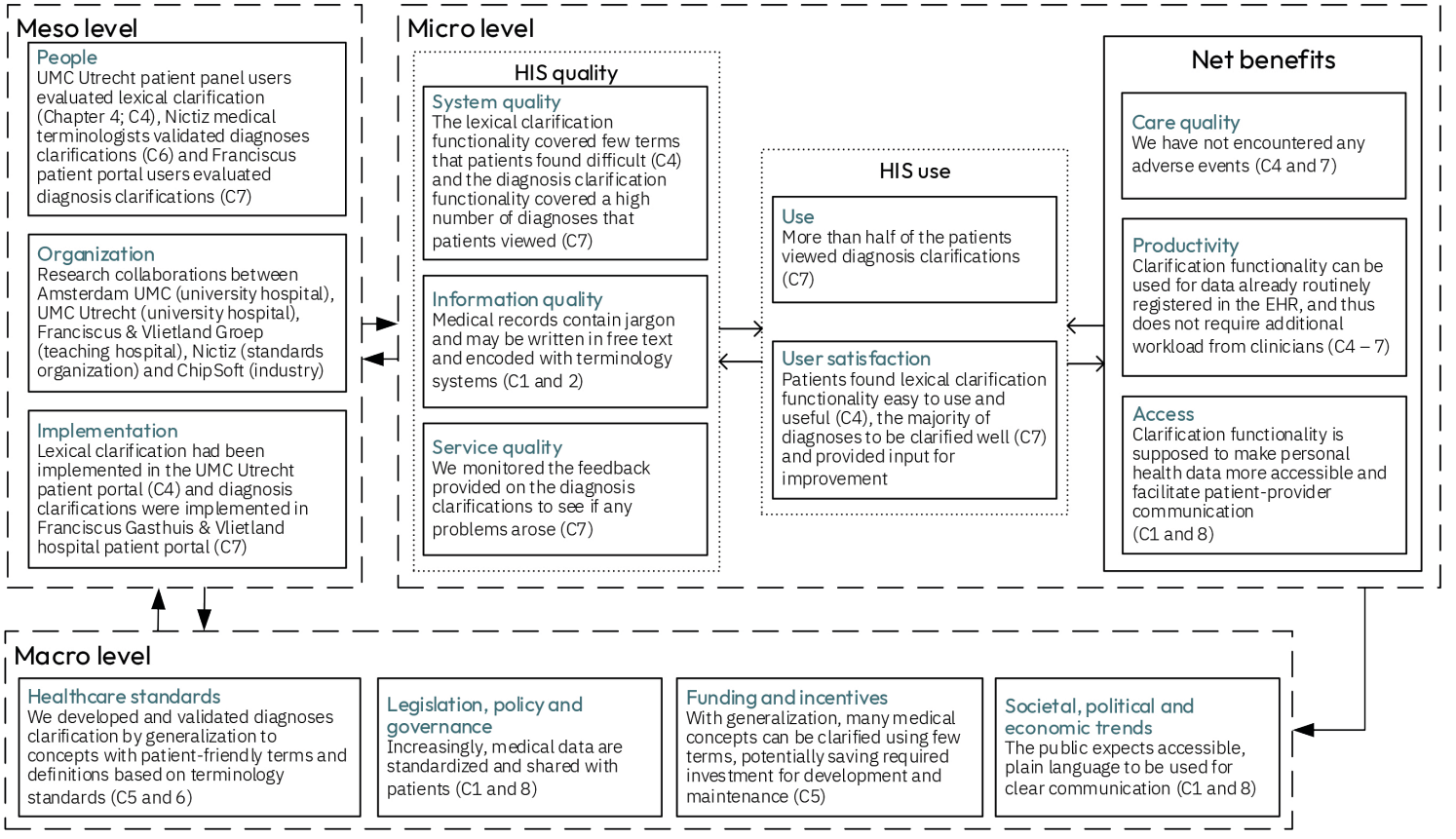


Figure 8.1 Principal findings from this thesis about clarification of medical data in electronic health records (EHRs) ordered in the Clinical Adoption Framework. HIS: Health Information System.

8.2 Strengths and weaknesses

We used the CAF to summarize the literature and provide an overview of the determinants and outcomes of the adoption of patient portals and PHRs in Chapter 2. A strength of using the CAF was that it is a general, integrative framework and that it enabled us to relate all the results from the literature to it. However, it was difficult to apply the categorization unambiguously and to apply it to patient access specifically, rather than HIS adoption in general. Therefore an additional strength was that we reported these issues and were able to suggest improvements to the CAF definitions and documentation. Admittedly, being a general, all-encompassing framework that might risk losing concrete value, our CAF-based literature synthesis enables finding reviews on specific topics, such as medical terminology. While it was not the primary aim of our systematic review, several included reviews reported the issue of difficult medical terminology and suggested providing definitions or explanations.[56,106-108,114,117] We have not assessed the strength of the association between determinants and outcomes or assessed the risk of bias in primary studies. A more recent meta-review[55] did assess bias in primary studies. It confirmed low to moderate evidence on most patient portal usage associations between determinants and outcomes. However, high evidence was found for ease of use and usefulness, secure messaging, prescription refills, medication information, and easy-to-understand information in lay language.[55] We addressed the latter issue in our research about patient-friendly terminology, Chapters 4 – 7.

A general issue with text simplification or, more particularly, lexical clarification, is that the exact concepts have to be identified in the text to be able to correctly clarify them. This can be challenging, because of typos, synonyms, homonyms, partial matching and alternative formulations. Conversely, an advantage of our focus on clarifying diagnoses is that we did not have to deal with this challenge of encoding free text, because we were able to reuse the diagnoses that were already encoded. Providing a patient-friendly synonym or a definition was only possible for a few terms due to the lack of available patient-friendly or consumer-oriented terminology. Therefore we introduced the algorithm that uses the SNOMED CT hierarchy to generalize diagnoses to supertype concepts with patient-friendly terms and definitions. Hence the lack of coverage of the thesaurus used in the lexical clarification feature can be partly addressed by generating clarifications using the algorithm we developed to clarify diagnoses. Lexical clarification compared to text simplification has the advantage of keeping the original text and data, while enabling the user to choose to receive clarification for

certain terms rather than rewriting and simplifying a whole text. An advantage of our knowledge-representation-based approach is that it can be easily explained how a certain result was derived, which is becoming increasingly more important with the requirements for explainable artificial intelligence, especially in critical domains such as healthcare.[211,212]

A strong point of our evaluation studies was that we evaluated functionalities with actual patients and their actual EHRs. For the diagnosis clarifications, we even went through the full cycle from the development and validation to the implementation and evaluation in clinical practice. This was possible thanks to a fruitful collaboration between academia (Amsterdam UMC, University of Amsterdam and UMC Utrecht, University of Utrecht), a healthcare standards organization (Nictiz), industry (ChipSoft) and clinical practice (UMC Utrecht and Francicus Gasthuis & Vlietland).

8.3 Significance and implications

In our research on patient-friendly terminology, we found that the quality of clarifications depended largely on the data quality. Clarifying standardized, encoded data is more straightforward than clarifying free text. In the review on reviews (Chapter 2), we found that policy, healthcare standards and regional healthcare infrastructures to support patient access to EHRs were addressed by only a few reviews. Recent developments in Europe and the USA, however, have shown increased attention to patient access to EHRs. Policymakers have tried to accelerate the adoption of EHRs, patient portals and PHRs, and the standardization of data and data exchange. In the USA the Meaningful Use program was introduced for this reason.[105,213] In the Netherlands the VIPP (an abbreviation for "Versnellingprogramma Informatieuitwisseling Patiënt en Professional" in Dutch or "acceleration program information exchange patient and professional" in English) incentivised many hospitals to adopt certain standards for data exchange between healthcare institutions and open up their EHRs for patients through patient portals and personal health records (PHRs, sometimes also referred to as patient-held records).[214] In 2022 a new law (Wegiz, an abbreviation for "Wet elektronische gegevensuitwisseling in de zorg" in Dutch or "Electronic Data Exchange in Healthcare Act" in English) was passed by the House of Representatives (but still needs to be approved by the Senate). This law will further enable standards-based data exchange between healthcare providers, and between healthcare providers and patients.[215] It also anticipates the level of the European Union (EU), where the European Health Data Space regulation

is being developed, to enable reuse for secondary usage purposes.[216-218] In the research community, the FAIR-data guidelines are widely acknowledged among research data managers. The FAIR principles state that research data should be findable, accessible, interoperable and reusable (FAIR) for humans and machines.[219,220] These developments in policy and trends in research practice will hopefully additionally enable more accessible and reusable health data for patients.

Despite the limited quality and coverage of the lexical clarification functionality evaluated in Chapter 4, the participants appreciated the functionality even for those few terms that were easily and usefully clarified. Since its implementation in the hospital-wide patient portal several years ago at UMC Utrecht no problems have been reported to our knowledge, such as anxious patients contacting their healthcare provider about the clarifications from the functionality. Similarly, we did not know in advance what to expect from the patient's interaction with diagnosis clarifications derived from SNOMED CTs hierarchy (Chapter 7) and we closely monitored any problems. Clinicians were afraid that this might increase the number of questions from patients or that errors might occur because not all clarifications had been validated. However, no problems have been reported, neither have questions been asked about the clarifications particularly since the functionality was implemented. This shows that the implementation of these functionalities is feasible.

Healthcare standards or terminology development organizations should consider adopting the generalization method presented in this thesis. Translating each medical concept manually is labour-intensive and might result in inconsistencies. We believe our method will enable them to make patient-friendly terminology better maintainable by being able to reuse parts of clarifications across medical concepts. Validation of the resulting clarifications can be carried out to find errors in terminology mappings and modelling, as we found in the validation study, and thus can be used to perform terminology auditing and quality improvement. Furthermore, while sometimes definitions from TZW and SNOMED CT were considered to be quite difficult to read, patients with adequate health literacy still found them useful. This shows that tailoring depends on the literacy level of end-users. A more general, plain language clarification may be sufficient for some users in some contexts, while others may prefer more details and specific

language use. In the September 2022 release[f], a B1-level reference set was added to the SNOMED CT Netherlands edition. This will enable tailoring to end-user preferences.

8.4 Further research

Research on patient access to EHRs can apply the adapted CAF to identify areas to address in HIS implementation and evaluation, structure their results and make results more comparable. Additionally, the developed guidance documentation may help in the application of the categories.

To further improve the text-matching algorithm of the lexical clarification functionality, natural language processing techniques, such as named entity linking and medical concept normalization should be applied.[221,222] The other way around, machine learning approaches can be improved by incorporating knowledge represented in ontologies such as SNOMED CT. As they are generally based on large text corpora, terms and concepts that are not common might be problematic for these models; while this knowledge might be readily available in domain-specific ontologies.[223] Clinical notes contain data that might be copied from other places and lose their original encodings. The EHR copy-paste functionality should be improved to prevent this and maintain the provenance of the data.

We used several methods to generate clarifications using SNOMED CT: providing synonyms, definitions and generalizations. In further research, defining relationships additional to hierarchical relationships in SNOMED CT should be considered to generate clarifications, similar to earlier efforts by Zeng et al[179], who used relationships from the UMLS. The defining relationships in SNOMED CT represent knowledge about those medical concepts. By deriving clarifications from the SNOMED CT ontology, parts of the text can be reused among more concepts. For example, Staphylococcus epidermidis ventriculitis can be generalized to bacterial infection and brain disorder using our generalization algorithm. However, this way some information is lost from its definition, such as the particular bacteria that caused the infection and that it involves an inflammation of the brain ventricles. From the defining relationships, it can be derived that it concerns a disorder with inflamed (associated morphology

f Note that the SNOMED CT Patient-friendly extension was merged with the SNOMED CT Netherlands edition and thus the Patient-friendly extension referred to in this thesis no longer exists, because the Patient-friendly reference set is now part of the Netherlands edition itself.

relationship) brain ventricles (finding site) caused by an infection (pathological process) with a bacteria called Staphylococcus epidermidis (causative agent).[9] See Figure 8.2. These relationships fully define the semantics of the concept. Hence, some of the problems with generalization, such as the loss of detail, can be overcome by utilizing additional defining relationships. Furthermore, natural language generation techniques should be employed to improve the current modular rule-based approach.[224]

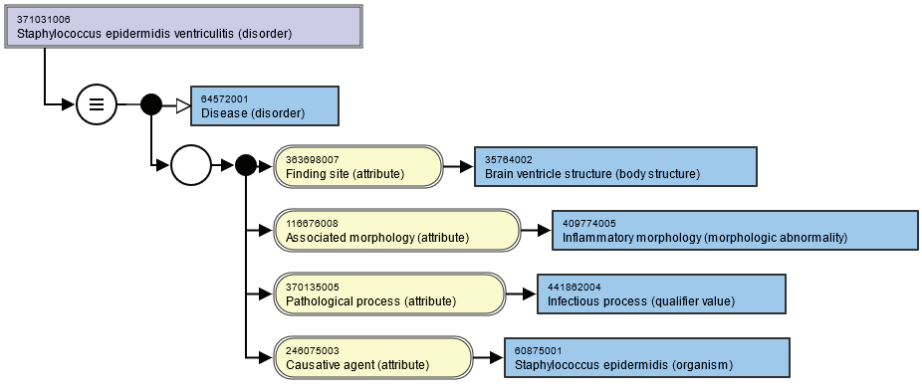


Figure 8.2 Diagram of stated defining relationships of 371031006 |Staphylococcus epidermidis ventriculitis|, from the SNOMED CT International browser[9]

Similarly, illustrations of anatomy and patient education leaflets can automatically be provided to patients or suggested to be used by clinicians if they are annotated with SNOMED CT concepts. For example, for all types of malignant neoplasms of the colon, a general patient education leaflet might be provided. An additional opportunity to clarify encoded data is to translate them to other languages using existing terminology content available, for those users that do not have English (or Dutch) as their primary language. For languages other than English, and for minority languages, fewer materials are available for language processing.[183] The approach we presented can potentially be a more efficient method of modelling and creating patient-friendly terminology, reducing effort and cost. This might lower the threshold to create clarifications for minority languages in the Netherlands such as local languages Frisian and Low Saxon, or foreign languages such as Turkish, Arabic, Chinese and Polish.[225] For the United States, Spanish, Chinese, Tagalog, Vietnamese and Arabic could be targeted.[226] Table 8.1 lists the methods to clarify medical diagnoses and provides examples of each.

Table 8.1 Methods to clarify medical diagnoses, with example diagnoses and clarifications

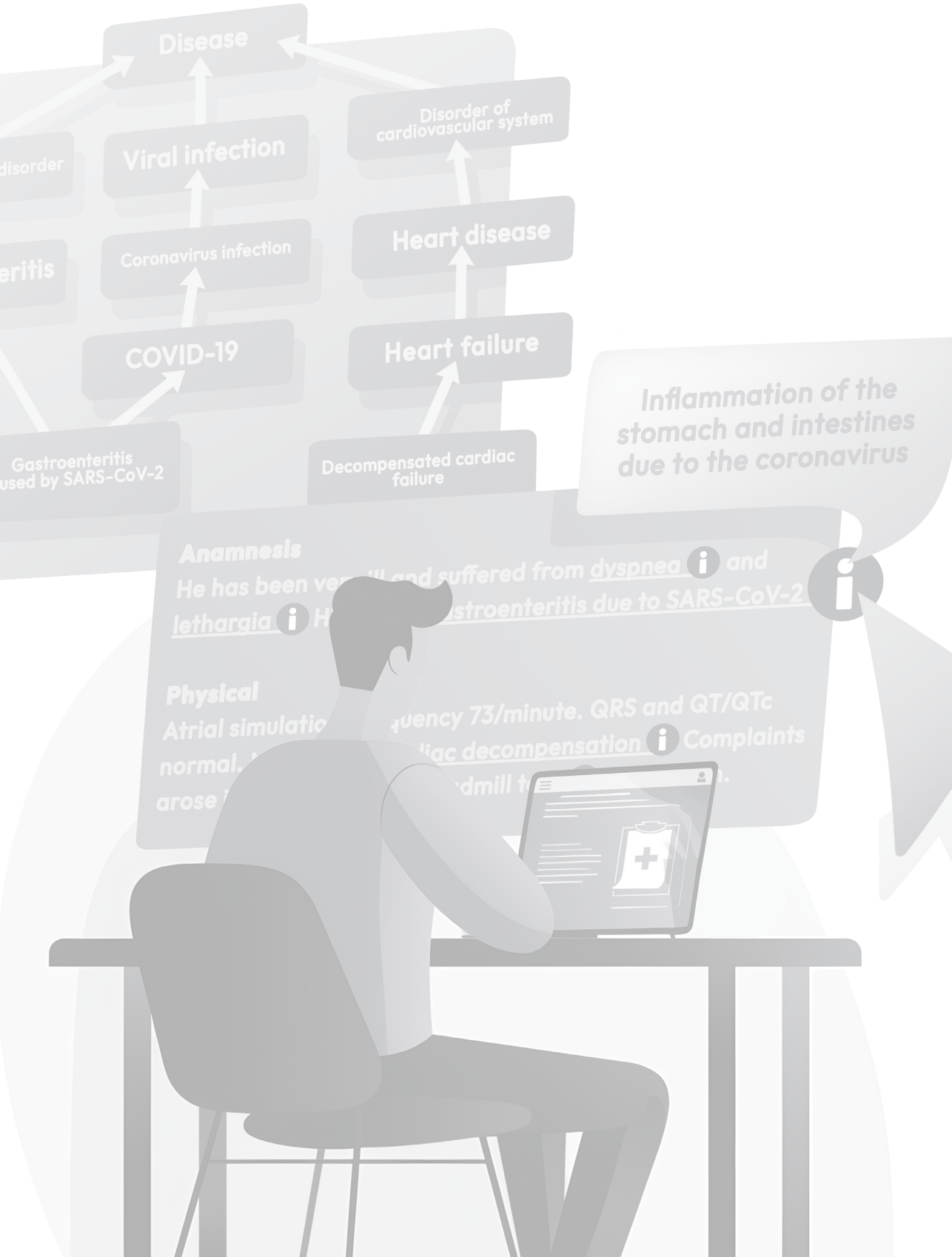
Method	Example diagnosis	Example clarification
Synonym	Dorsalgia	Another word for dorsalgia is back pain.[78]
Definition	Aortic valve stenosis	Narrowing of the aortic valve, making it difficult for blood to flow through the heart.[78]
Translation	Cerebrovascular accident	Beroerte (Dutch; "stroke" in English)[78]
Generalization	Hereditary factor IX deficiency disease	A blood clotting disorder. It is inborn and hereditary.[78]
Derivation	Staphylococcus epidermidis ventriculitis	An inflammation of the brain ventricles due to an infection with the bacteria Staphylococcus epidermidis.[78]
Illustration	Pancreatitis	Illustration of the pancreas, e.g. from Wikimedia.[227]
Education	Malignant neoplasm of colon	Link to patient information website medlineplus.gov[228], to specific patient information leaflets used locally by the healthcare provider, or to multimedia such as videos or interactive patient education applications.

8.5 Conclusions

This thesis contributes to the reuse of EHRs by patients and clarifying medical data to patients and their significant others. We created an overview of determinants and outcomes of patient access to EHRs, facilitated by patient portals and personal health records. The associations between determinants and outcomes of patient access to EHRs through patient portals and PHRs have been widely reported in systematic reviews, but the evidence was weak for most reviews.

We evaluated a lexical clarification tool that provides clarifications to medical terms in clinical notes in a patient portal. Although the consumer-oriented thesaurus from the tool provided limited coverage and quality, patient portal users still considered it easy to use and useful. We suggest the coverage could be increased by using the generalization method proposed in this thesis to generate clarifications for more medical concepts. Additionally, there are some challenges in lexical clarification related to natural language processing. To further improve the text-matching algorithm, named entity linking and medical concept normalization should be applied.

We developed an algorithm to generate diagnosis clarifications by using patient-friendly terms, definitions and generalizations. By generalization, we were able to increase the coverage of diagnoses by clarifications significantly. We identified problems in the clarifications and improved the patient-friendly terms used, to implement the clarifications in a hospital patient portal. The majority of patient portal users were found to use the clarifications when they could and rated most of the clarifications as having good quality. This shows that knowledge represented in SNOMED CT can be utilized to generate and reuse clarifications across medical concepts. Further work will be dedicated to extending the method by using non-hierarchical relationships, the adoption of this technology by terminology standards developers, quality improvement and the implementation in clinical practice.


Disease
Viral infection
Disorder of cardiovascular system
Coronavirus infection
Heart disease
COVID-19
Heart failure
Gastroenteritis
Decompensated cardiac failure
Inflammation of the stomach and intestines due to the coronavirus
Anamnesis
Physical

A

Appendices

References

1. Lerner EB, Jehle DV, Janicke DM, Moscati RM. Medical communication: do our patients understand? *Am J Emerg Med.* 2000;18(7):764-766.
2. Graham S, Brookey J. Do patients understand? *Perm J.* 2008;12(3):67-69.
3. Gotlieb R, Praska C, Hendrickson MA, et al. Accuracy in Patient Understanding of Common Medical Phrases. *JAMA Netw Open.* 2022;5(11):e2242972.
4. Keselman A, Smith CA. A classification of errors in lay comprehension of medical documents. *J Biomed Inform.* 2012;45(6):1151-1163.
5. Fernández L, Fossa A, Dong Z, et al. Words Matter: What Do Patients Find Judgmental or Offensive in Outpatient Notes? *Journal of general internal medicine.* 2021;36(9):2571-2578.
6. Cox C, Fritz Z. Presenting complaint: use of language that disempowers patients. *BMJ (Clinical research ed).* 2022;377:e066720.
7. Swiech P. Medical jargon can be confusing, misinterpreted. https://web.archive.org/web/20180511105726/http://www.pantagraph.com/lifestyles/health-med-fit/medical-jargon-can-be-confusing-misinterpreted/article_8e6f2742-6773-11df-afd7-001cc4c002e0.html. Published 2010. Accessed May 11, 2018.
8. Pada S, Lishi P, Ng KS, et al. The impact of a 'milking the COW' campaign in a regional hospital in Singapore. *Antimicrob Resist Infect Control.* 2021;10(1):81.
9. SNOMED International. SNOMED International SNOMED CT Browser. https://browser.ihtsdotools.org/. Accessed December 19, 2022.
10. Ayers Career College. When Medical Terminology is Appropriate to Use. https://web.archive.org/web/20220407214203/https://ayers.edu/blog/medical-terminology-appropriate-use-with-patients/. Published 2019. Accessed April 7, 2022.
11. van Rosse F, de Bruijne M, Suurmond J, Essink-Bot ML, Wagner C. Language barriers and patient safety risks in hospital care. A mixed methods study. *International journal of nursing studies.* 2016;54:45-53.
12. Matlow AG, Wishen A, Read SE, Raboud JM. A study of provider-caregiver communication in paediatric ambulatory care. *Paediatr Child Health.* 2006;11(4):217-221.
13. Hull M. Medical language proficiency: A discussion of interprofessional language competencies and potential for patient risk. *International journal of nursing studies.* 2016;54:158-172.
14. Karim SK, Amin OSM. Stroke in Ancient Mesopotamia. *Med Arch.* 2018;72(6):449-452.
15. Al-Awqati Q. How to write a case report: lessons from 1600 B.C. *Kidney International.* 2006;69(12):2113-2114.
16. Dalianis H. The History of the Patient Record and the Paper Record. In: *Clinical Text Mining: Secondary Use of Electronic Patient Records.* Cham: Springer International Publishing; 2018:5-12.
17. Lorkowski J, Pokorski M. Medical Records: A Historical Narrative. *Biomedicines.* 2022;10(10).
18. Cheng TO. Hippocrates and cardiology. *American Heart Journal.* 2001;141(2):173-183.
19. Johnson KB, Neuss MJ, Detmer DE. Electronic health records and clinician burnout: A story of three eras. *Journal of the American Medical Informatics Association : JAMIA.* 2021;28(5):967-973.

20. Rodríguez-Vera FJ, Marín Y, Sánchez A, Borrachero C, Pujol E. Illegible handwriting in medical records. *J R Soc Med.* 2002;95(11):545-546.
21. Bhojwani D, Javed H, Amala C, Hashmi S. 330 An End to Doctor's Handwriting in the Digital Age: Evaluating the Legibility of Operation Notes in Otolaryngology. *British Journal of Surgery.* 2022;109(Supplement_1).
22. Dalianis H. Characteristics of Patient Records and Clinical Corpora. In: *Clinical Text Mining: Secondary Use of Electronic Patient Records.* Cham: Springer International Publishing; 2018:21-34.
23. Arts DG, De Keizer NF, Scheffer GJ. Defining and improving data quality in medical registries: a literature review, case study, and generic framework. *Journal of the American Medical Informatics Association : JAMIA.* 2002;9(6):600-611.
24. de Keizer NF, Abu-Hanna A, Zwetsloot-Schonk JH. Understanding terminological systems. I: Terminology and typology. *Methods of information in medicine.* 2000;39(1):16-21.
25. International Statistical Classification of Diseases and Related Health Problems (ICD). Wold Health Organization. https://web.archive.org/web/20221228150420/https://www.who.int/standards/classifications/classification-of-diseases. Published 2022. Accessed December 28, 2022.
26. Alonso V, Santos JV, Pinto M, et al. Health records as the basis of clinical coding: Is the quality adequate? A qualitative study of medical coders' perceptions. *Health Information Management Journal.* 2020;49(1):28-37.
27. International Classification of Primary Care. The WONCA ICPC Foundation. https://web.archive.org/web/20221229145703/https://www.icpc-3.info/. Published 2022. Accessed December 29, 2022.
28. Busse R, Geissler A, Aaviksoo A, et al. Diagnosis related groups in Europe: moving towards transparency, efficiency, and quality in hospitals? *BMJ : British Medical Journal.* 2013;346:f3197.
29. Fetter RB. Diagnosis related groups: understanding hospital performance. *Interfaces.* 1991;21(1):6-26.
30. Oostenbrink JB, Rutten FF. Cost assessment and price setting of inpatient care in The Netherlands. the DBC case-mix system. *Health care management science.* 2006;9(3):287-294.
31. Handleiding dbc-systematiek. NZA. https://puc.overheid.nl/nza/doc/PUC_721877_22/1/. Published 2022. Accessed January 9, 2023.
32. Zimmerman JE, Kramer AA, McNair DS, Malila FM. Acute Physiology and Chronic Health Evaluation (APACHE) IV: hospital mortality assessment for today's critically ill patients. *Critical care medicine.* 2006;34(5):1297-1310.
33. Nationale Intensive Care Evaluatie. Wat wij registreren (in Dutch; 'What we register' in English). https://web.archive.org/web/20221005211422/https://www.stichting-nice.nl/watweregistreren.jsp. Published 2022. Accessed October 5, 2022.
34. Medicine NLo. MeSH. National Intitutes of Health. https://www.ncbi.nlm.nih.gov/mesh. Published 2022. Accessed December 28, 2022.
35. Kroth PJ, Morioka-Douglas N, Veres S, et al. Association of Electronic Health Record Design and Use Factors With Clinician Stress and Burnout. *JAMA Network Open.* 2019;2(8):e199609-e199609.
36. Joukes E, Abu-Hanna A, Cornet R, de Keizer NF. Time Spent on Dedicated Patient Care and Documentation Tasks Before and After the Introduction of a Structured and Standardized Electronic Health Record. *Applied clinical informatics.* 2018;9(1):46-53.

37. Bakhshi-Raiez F, de Keizer NF, Cornet R, Dorrepaal M, Dongelmans D, Jaspers MWM. A usability evaluation of a SNOMED CT based compositional interface terminology for intensive care. *International journal of medical informatics.* 2012;81(5):351-362.
38. Zheng K, Hanauer DA, Padman R, et al. Handling anticipated exceptions in clinical care: investigating clinician use of 'exit strategies' in an electronic health records system. *Journal of the American Medical Informatics Association : JAMIA.* 2011;18(6):883-889.
39. Klappe ES, van Putten FJP, de Keizer NF, Cornet R. Contextual property detection in Dutch diagnosis descriptions for uncertainty, laterality and temporality. *BMC medical informatics and decision making.* 2021;21(1):120.
40. DHD. T-REX. https://trex.dhd.nl/. Published 2022. Accessed December 28, 2022.
41. NHG. ICPC-1 code-zoeker. https://viewers.nhg.org/icpcviewer/. Published 2022. Accessed December 28, 2022.
42. The WONCA ICPC Foundation. ICPC-3 browser. https://browser.icpc-3.info/. Published 2022. Accessed December 28, 2022.
43. Organization WH. ICD-10 Version:2010. https://icd.who.int/browse10/2010/en#/S06.2. Published 2022. Accessed December 28, 2022.
44. Joukes E, Cornet R, de Keizer N, de Bruijne M. Collect Once, Use Many Times: End-Users Don't Practice What They Preach. *Stud Health Technol Inform.* 2016;228:252-256.
45. Oplossingen voor registratievraagstukken DHD. http://web.archive.org/web/20230104121230/https://www.dhd.nl/producten-diensten/registratie-data/oplossingen-voor-registratievraagstukken. Published 2023. Accessed January 4, 2023.
46. SNOMED CT Starter Guide. International Health Terminology Standards Development Organisation. https://confluence.ihtsdotools.org/display/DOCSTART. Published 2022. Accessed January 4, 2023.
47. SNOMED CT Search and Data Entry Guide. International Health Terminology Standards Development Organisation. https://confluence.ihtsdotools.org/display/DOCSEARCH. Published 2017. Accessed January 4, 2023.
48. Data Analytics with SNOMED CT. International Health Terminology Standards Development Organisation. https://confluence.ihtsdotools.org/display/DOCANLYT. Published 2021. Accessed January 4, 2023.
49. Decision Support with SNOMED CT. International Health Terminology Standards Development Organisation. https://confluence.ihtsdotools.org/display/DOCCDS. Published 2021. Accessed January 4, 2023.
50. van Mens HJT. *Terminology system-based data encoding for intensive care: Deriving the APACHE-IV reasons for ICU admission classification through SNOMED CT and optimizing the user interface for diagnostic data entry*. Amsterdam: Faculty of Medicine, Department of Medical Informatics, Academic Medical Center, University of Amsterdam; 2016.
51. Bakhshi-Raiez F, Cornet R, de Keizer NF. Cross-mapping APACHE IV "reasons for intensive care admission" classification to SNOMED CT. *Stud Health Technol Inform.* 2008;136:779-784.
52. Artikel 446 Burgerlijk Wetboek Boek 7. In:2016.
53. General Data Protection Regulation (GDPR). GDPR.EU. https://gdpr.eu/tag/gdpr/. Published 2023. Accessed January 2, 2023.
54. Joseph L, Lavis A, Greenfield S, et al. Systematic review on the use of patient-held health records in low-income and middle-income countries. *BMJ Open.* 2021;11(9):e046965.

55. Antonio MG, Petrovskaya O, Lau F. The state of evidence in patient portals: umbrella review. *Journal of medical Internet research.* 2020;22(11):e23851.
56. Irizarry T, DeVito Dabbs A, Curran CR. Patient Portals and Patient Engagement: A State of the Science Review. *Journal of medical Internet research.* 2015;17(6):e148.
57. Vermeir P, Degroote S, Vandijck D, et al. The patient perspective on the effects of medical record accessibility: a systematic review. *Acta clinica Belgica.* 2017;72(3):186-194.
58. Mold F, De Lusignan S, Sheikh A, et al. Patients' online access to their electronic health records and linked online services: A systematic review in primary care. *British Journal of General Practice.* 2015;65(632):e141-e151.
59. van der Vaart R, van Deursen L, Standaar L, et al. E-healthmonitor 2021 Stand van zaken digitale zorg. Rijksinstituut voor Volksgezondheid en Milieu. https://web.archive.org/web/20230104155051/https://www.rivm.nl/documenten/e-healthmonitor-2021-stand-van-zaken-digitale-zorg. Published 2022. Accessed January 4, 2023.
60. Archer N, Fevrier-Thomas U, Lokker C, McKibbon KA, Straus SE. Personal health records: a scoping review. *Journal of the American Medical Informatics Association : JAMIA.* 2011;18(4):515-522.
61. Niazkhani Z, Toni E, Cheshmekaboodi M, Georgiou A, Pirnejad H. Barriers to patient, provider, and caregiver adoption and use of electronic personal health records in chronic care: a systematic review. *BMC medical informatics and decision making.* 2020;20(1):153.
62. Harahap NC, Handayani PW, Hidayanto AN. Functionalities and Issues in the Implementation of Personal Health Records: Systematic Review. *Journal of medical Internet research.* 2021;23(7):e26236.
63. McMullan M. Patients using the Internet to obtain health information: how this affects the patient–health professional relationship. *Patient education and counseling.* 2006;63(1-2):24-28.
64. Tan SS-L, Goonawardene N. Internet Health Information Seeking and the Patient-Physician Relationship: A Systematic Review. *Journal of medical Internet research.* 2017;19(1):e9.
65. Nutbeam D. Health literacy as a public health goal: a challenge for contemporary health education and communication strategies into the 21st century. *Health promotion international.* 2000;15(3):259-267.
66. Baccolini V, Rosso A, Di Paolo C, et al. What is the Prevalence of Low Health Literacy in European Union Member States? A Systematic Review and Meta-analysis. *Journal of general internal medicine.* 2021;36(3):753-761.
67. Berkman ND, Sheridan SL, Donahue KE, Halpern DJ, Crotty K. Low health literacy and health outcomes: an updated systematic review. *Annals of internal medicine.* 2011;155(2):97-107.
68. Otte-Trojel T, de Bont A, Rundall TG, van de Klundert J. What do we know about developing patient portals? A systematic literature review. *Journal of the American Medical Informatics Association.* 2016;23(e1):e162-e168.
69. Coughlin SS, Prochaska JJ, Williams LB, et al. Patient web portals, disease management, and primary prevention. *Risk Management & Healthcare Policy.* 2017;10:33-40.
70. Janssen SL, Venema-Taat N, Medlock S. Anticipated Benefits and Concerns of Sharing Hospital Outpatient Visit Notes With Patients (Open Notes) in Dutch Hospitals: Mixed Methods Study. *Journal of medical Internet research.* 2021;23(8):e27764.

71. Schreuder C, Pluut B. Handboek online inzage: Hoe regelen we dat?! https://web.archive.org/web/20230113142152/https://nictiz.nl/app/uploads/2022/08/Handboek_online_inzage_definitief.pdf. Published 2018. Accessed January 13, 2023.
72. Nictiz. Thesaurus Zorg & Welzijn. https://web.archive.org/web/20221212121517/https://www.thesauruszorgenwelzijn.nl/. Published 2022. Accessed December 28, 2022.
73. Shiri AA, Revie C. Thesauri on the Web: Current developments and trends. In: MCB University press; 2000.
74. Zeng QT. Consumer Health Vocabulary Initiative. https://web.archive.org/web/20221129171544/https://biomedinfo.smhs.gwu.edu/consumer-health-vocabulary-initiative. Published 2022. Accessed November 29, 2022.
75. Zorgkaart Nederland. Patiëntenfederatie. https://www.zorgkaartnederland.nl/. Published 2023. Accessed January 18, 2023.
76. Kanker.nl - We zijn er voor je. Stichting Kanker.nl. https://www.kanker.nl/. Published 2023. Accessed 13, 2023.
77. Taalunie. Taalunie. https://web.archive.org/web/20230113141122/https://taalunie.org/informatie/112/taalunie-union-for-the-dutch-language. Published 2023. Accessed January 13, 2023.
78. Nictiz. SNOMED CT. https://terminologie.nictiz.nl/art-decor/snomed-ct. Published 2022. Accessed December 28, 2022.
79. Henderson C, Laugharne R. Patient held clinical information for people with psychotic illnesses. *The Cochrane database of systematic reviews.* 2000(2):Cd001711.
80. Ross SE, Lin CT. The effects of promoting patient access to medical records: a review. *Journal of the American Medical Informatics Association : JAMIA.* 2003;10(2):129-138.
81. Ammenwerth E, Schnell-Inderst P, Hoerbst A. The impact of electronic patient portals on patient care: a systematic review of controlled trials. *Journal of medical Internet research.* 2012;14(6):e162.
82. McMillan B, Eastham R, Brown B, Fitton R, Dickinson D. Primary Care Patient Records in the United Kingdom: Past, Present, and Future Research Priorities. *Journal of medical Internet research.* 2018;20(12):e11293.
83. Prey JE, Polubriaginof F, Kuperman GJ, Tiase V, Collins SA, Vawdrey DK. International perspectives on sharing clinical data with patients. *International journal of medical informatics.* 2016;86:135-141.
84. Shah SGS, Fitton R, Hannan A, Fisher B, Young T, Barnett J. Accessing personal medical records online: A means to what ends? *International journal of medical informatics.* 2015;84(2):111-118.
85. Kharrazi H, Chisholm R, VanNasdale D, Thompson B. Mobile personal health records: An evaluation of features and functionality. *International journal of medical informatics.* 2012;81(9):579-593.
86. HealthIT.gov. Blue Button. https://www.healthit.gov/topic/health-it-initiatives/blue-button. Published 2016. Accessed May 1, 2019.
87. Hogan TP, Nazi KM, Luger TM, et al. Technology-assisted patient access to clinical information: an evaluation framework for blue button. *JMIR research protocols.* 2014;3(1):e18-e18.
88. Blue Button+ Implementation Guide. https://web.archive.org/web/20160331011031/http://bluebuttonplus.org/. Published 2016. Accessed May 1, 2019.

89. Chronaki C, Estelrich A, Cangioli G, et al. Interoperability standards enabling cross-border patient summary exchange. *Stud Health Technol Inform.* 2014;205:256-260.
90. Nazi KM, Turvey CL, Klein DM, Hogan TP, Woods SS. VA OpenNotes: exploring the experiences of early patient adopters with access to clinical notes. *Journal of the American Medical Informatics Association : JAMIA.* 2015;22(2):380-389.
91. Dobscha SK, Denneson LM, Jacobson LE, Williams HB, Cromer R, Woods S. VA mental health clinician experiences and attitudes toward OpenNotes. *General hospital psychiatry.* 2016;38:89-93.
92. Schoevers MA, van den Muijsenbergh ME, Lagro-Janssen AL. Patient-held records for undocumented immigrants: a blind spot. A systematic review of patient-held records. *Ethnicity & health.* 2009;14(5):497-508.
93. Beard L, Schein R, Morra D, Wilson K, Keelan J. The challenges in making electronic health records accessible to patients. *Journal of the American Medical Informatics Association : JAMIA.* 2012;19(1):116-120.
94. Saparova D. Motivating, influencing, and persuading patients through personal health records: a scoping review. *Perspectives in health information management / AHIMA, American Health Information Management Association.* 2012;9:1f.
95. Goldzweig CL, Towfigh AA, Paige NM, et al. VA Evidence-based Synthesis Program Reports. In: *Systematic Review: Secure Messaging Between Providers and Patients, and Patients' Access to Their Own Medical Record: Evidence on Health Outcomes, Satisfaction, Efficiency and Attitudes.* Washington (DC): Department of Veterans Affairs (US); 2012.
96. Farrelly S, Brown GE, Flach C, Barley E, Laugharne R, Henderson C. User-held personalised information for routine care of people with severe mental illness. *The Cochrane database of systematic reviews.* 2013(10):CD001711.
97. Goldzweig CL, Orshansky G, Paige NM, et al. Electronic patient portals: evidence on health outcomes, satisfaction, efficiency, and attitudes: a systematic review. *Annals of internal medicine.* 2013;159(10):677-687.
98. Otte-Trojel T, de Bont A, Aspria M, et al. Developing patient portals in a fragmented healthcare system. *International journal of medical informatics.* 2015;84(10):835-846.
99. Lau F, Price M, Keshavjee K. From benefits evaluation to clinical adoption: making sense of health information system success in Canada. *Healthc Q.* 2011;14(1):39-45.
100. Lau F, Price M, Boyd J, Partridge C, Bell H, Raworth R. Impact of electronic medical record on physician practice in office settings: a systematic review. *BMC medical informatics and decision making.* 2012;12:10.
101. van Mens HJT, Nienhuis R, Duijm RD, Cornet R, de Keizer NF. An overview of determinants and outcomes of providing patients access to and control over their own medical data: A systematic review of reviews. PROSPERO CRD42018084542. http://www.crd.york.ac.uk/PROSPERO/display_record.php?ID=CRD42018084542. Published 2018. Accessed December 28, 2022.
102. van Mens HJT, Nienhuis R, Duijm RD, Daams J, Cornet R, de Keizer NF. An overview of determinants and outcomes of providing patients access to and control over their own medical data: A systematic review of reviews protocol. https://www.crd.york.ac.uk/PROSPEROFILES/84542_PROTOCOL_20180411.pdf. Published 2018. Accessed May 7, 2018.

103. van Mens HJT, Nienhuis R, Daams J, Cornet R, de Keizer NF. An overview of determinants and outcomes of providing patients access to and control over their own medical data: Search strategy. PROSPERO. https://www.crd.york.ac.uk/PROSPEROFILES/84542_STRATEGY_20180409.pdf. Published 2018. Accessed May 7, 2018.
104. Shea BJ, Reeves BC, Wells G, et al. AMSTAR 2: a critical appraisal tool for systematic reviews that include randomised or non-randomised studies of healthcare interventions, or both. *BMJ (Clinical research ed).* 2017;358.
105. Kruse CS, Bolton K, Freriks G. The effect of patient portals on quality outcomes and its implications to meaningful use: a systematic review. *Journal of medical Internet research.* 2015;17(2):e44.
106. Bush RA, Connelly CD, Fuller M, Perez A. Implementation of the Integrated Electronic Patient Portal in the Pediatric Population: A Systematic Review. *Telemedicine Journal & E Health.* 2015;10:10.
107. Hemsley B, Rollo M, Georgiou A, Balandin S, Hill S. The health literacy demands of electronic personal health records (e-PHRs): An integrative review to inform future inclusive research. *Patient Education & Counseling.* 2018;101(1):2-15.
108. Kelly MM, Coller RJ, Hoonakker PL. Inpatient Portals for Hospitalized Patients and Caregivers: A Systematic Review. *Journal of hospital medicine.* 2017;20:20.
109. Kneale L, Demiris G. Lack of Diversity in Personal Health Record Evaluations with Older Adult Participants: A Systematic Review of Literature. *Journal of innovation in health informatics.* 2017;23(4):881.
110. Kruse CS, Argueta DA, Lopez L, Nair A. Patient and provider attitudes toward the use of patient portals for the management of chronic disease: a systematic review. *Journal of medical Internet research.* 2015;17(2):e40.
111. Powell KR. Patient-Perceived Facilitators of and Barriers to Electronic Portal Use: A Systematic Review. *CIN: Computers, Informatics, Nursing.* 2017;35(11):565-573.
112. Price M, Bellwood P, Kitson N, Davies I, Weber J, Lau F. Conditions potentially sensitive to a personal health record (PHR) intervention, a systematic review. *BMC Med Inf Decis Mak.* 2015;15:32.
113. Rathert C, Mittler JN, Banerjee S, McDaniel J. Patient-centered communication in the era of electronic health records: What does the evidence say? *Patient education and counseling.* 2017;100(1):50-64.
114. Sakaguchi-Tang DK, Bosold AL, Choi YK, Turner AM. Patient Portal Use and Experience Among Older Adults: Systematic Review. *JMIR medical informatics.* 2017;5(4):e38.
115. Sartain SA, Stressing S, Prieto J. Patients' views on the effectiveness of patient-held records: a systematic review and thematic synthesis of qualitative studies. *Health Expectations.* 2015;18(6):2666-2677.
116. Showell C. Barriers to the use of personal health records by patients: a structured review. *PeerJ.* 2017;5:e3268.
117. Thompson MJ, Reilly JD, Valdez RS. Work system barriers to patient, provider, and caregiver use of personal health records: A systematic review. *Applied Ergonomics.* 2016;54:218-242.
118. Turner K, Klaman SL, Shea CM. Personal health records for people living with HIV: a review. *AIDS Care - Psychological and Socio-Medical Aspects of AIDS/HIV.* 2016;28(9):1181-1187.
119. Al Dahdah M, Desgrees Du Lo UA, Meadel C. Mobile health and maternal care: A winning combination for healthcare in the developing world? *Health policy and technology.* 2015;4(3):225-231.

120. Arens-Volland AG, Spassova L, Bohn T. Promising approaches of computer-supported dietary assessment and management-Current research status and available applications. *International Journal of Medical Informatics.* 2015;84(12):997-1008.
121. Baysari MT, Westbrook JI. Mobile Applications for Patient-centered Care Coordination: A Review of Human Factors Methods Applied to their Design, Development, and Evaluation. *Yearb.* 2015;10(1):47-54.
122. Canneaux M, Kobilinsky N, Wolf JP, Golse B, Beauquier-Maccotta B. [Information, transmission, secrecy: What kind of discourse for children born by gamete donation?]. *Gynecologie, Obstetrique & Fertilite.* 2016;44(7):410-416.
123. Moon LA. Factors influencing health data sharing preferences of consumers: A critical review. *Health Policy and Technology.* 2017;6(2):169-187.
124. O'Connor S, Hanlon P, O'Donnell CA, Garcia S, Glanville J, Mair FS. Understanding factors affecting patient and public engagement and recruitment to digital health interventions: a systematic review of qualitative studies. *BMC Med Inf Decis Mak.* 2016;16(1):120.
125. Robinson JR, Huth H, Jackson GP. Review of information technology for surgical patient care. *Journal of Surgical Research.* 2016;203(1):121-139.
126. Turvey CL, Roberts LJ. Recent developments in the use of online resources and mobile technologies to support mental health care. *International Review of Psychiatry.* 2015;27(6):547-557.
127. Wu H, LaRue EM. Linking the health data system in the U.S.: Challenges to the benefits. *International Journal of Nursing Sciences.* 2017;4(4):410-417.
128. Yang Y, Asan O. Designing Patient-facing Health Information Technologies for the Outpatient Settings: A Literature Review. *J Innov Health Inform.* 2016;23(1):185.
129. Bachiri M, Idri A, Fernandez-Aleman JL, Toval A. Mobile personal health records for pregnancy monitoring functionalities: Analysis and potential. *Computer Methods and Programs in Biomedicine.* 2016;134:121-135.
130. Fant CT, Adelman DS. Sharing clinical notes with patients. *NURSE PRACT.* 2017;42(10):1-5.
131. Fraccaro P, Vigo M, Balatsoukas P, Buchan IE, Peek N, van der Veer SN. Patient Portal Adoption Rates: A Systematic Literature Review and Meta-Analysis. *Stud Health Technol Inform.* 2017;245:79-83.
132. Gefen R, Bruno MA, Abujudeh HH. Online Portals: Gateway to Patient-Centered Radiology. *AJR American Journal of Roentgenology.* 2017;209(5):987-991.
133. Heart T, Ben-Assuli O, Shabtai I. A review of PHR, EMR and EHR integration: A more personalized healthcare and public health policy. *Health Policy and Technology.* 2016.
134. Hemsley B, Georgiou A, Balandin S, et al. The Personally Controlled Electronic Health Record (PCEHR) for Adults with Severe Communication Impairments: Findings of Pilot Research. *Stud Health Technol Inform.* 2015;214:100-106.
135. Jilka SR, Callahan R, Sevdalis N, Mayer EK, Darzi A. "Nothing About Me Without Me": An Interpretative Review of Patient Accessible Electronic Health Records. *J Med Internet Res.* 2015;17(6):e161.
136. Kendrick EJ, Benson C. Patient Portals in Child and Adolescent Psychiatry. *Child and adolescent psychiatric clinics of North America.* 2017;26(1):43-54.
137. Kildea J, Hendren L, Hererra D, Joseph A, Maglieri R, Hijal T. Opal-The oncology portal and application. *Radiotherapy and Oncology.* 2017;123:S950-S951.
138. Lee CI, Langlotz CP, Elmore JG. Implications of Direct Patient Online Access to Radiology Reports Through Patient Web Portals. *Journal of the American College of Radiology.* 2016;13(12):1608-1614.

139. Martinez W, Threatt A, Rosenbloom ST, Wallston KA, Hickson GB, Elasy TA. Development and usability of a patient-facing diabetes dashboard. *Diabetes.* 2017;66:A221.
140. Mohammed AA, Bagalagel A, Noor A, Husband AK, Forrest SP, Slight SP. A systematic review of the experiences of health care staff on using electronic systems for the follow-up of patient test results. *Journal of General Internal Medicine.* 2017;32 (2 Supplement 1):S103.
141. Mohsen MO, Aziz HA. The Blue Button Project: Engaging Patients in Healthcare by a Click of a Button. *Perspect.* 2015;12:1d.
142. Prashad R. The Role of Personal Health Record Systems in Chronic Disease Management. *Stud Health Technol Inform.* 2017;234:275-279.
143. Shah SD, Liebovitz D. It Takes Two to Tango: Engaging Patients and Providers With Portals. *Pm & R.* 2017;9(5):S85-S97.
144. Showell C, Cummings E, Turner P. The Invisibility of Disadvantage: Why Do We Not Notice? *Stud Health Technol Inform.* 2017;235:388-392.
145. Sorensen T, Johansen MA. Developing and Implementing Patients' Full-Scale Electronic Access to Their Health Record. *Stud Health Technol Inform.* 2016;228:85-89.
146. Tietze M, Winters C, Hoelscher SH. Personal health records and patient portals. *McBride, Susan [Ed]; Tietze, Mari [Ed] (2016) Nursing informatics for the advanced practice nurse: Patient safety, quality, outcomes, and interprofessionalism (pp 369-388) xxii, 720 pp New York, NY, US: Springer Publishing Co; US.* 2016:369-388.
147. Tulu B, Trudel J, Strong DM, Johnson SA, Sundaresan D, Garber L. Patient Portals: An underused resource for improving patient engagement. *Chest.* 2016;149(1):272-277.
148. Wake DJ, He J, Czesak AM, Mughal F, Cunningham SG. MyDiabetesMyWay: An Evolving National Data Driven Diabetes Self-Management Platform. *J Diabetes Sci Technol.* 2016;10(5):1050-1058.
149. Wildenbos GA, Peute L, Jaspers M. Facilitators and Barriers of Electronic Health Record Patient Portal Adoption by Older Adults: A Literature Study. *Stud Health Technol Inform.* 2017;235:308-312.
150. Davis S, Roudsari A, Raworth R, Courtney KL, MacKay L. Shared decision-making using personal health record technology: a scoping review at the crossroads. *Journal of the American Medical Informatics Association.* 2017;24(4):857-866.
151. Fraccaro P, Vigo M, Balatsoukas P, Buchan IE, Peek N, van der Veer SN. The influence of patient portals on users' decision making is insufficiently investigated: A systematic methodological review. *International Journal of Medical Informatics.* 2018;111:100-111.
152. Hemsley B, Georgiou A, Carter R, et al. Use of the My Health Record by people with communication disability in Australia: A review to inform the design and direction of future research. *Health Inf Manage J.* 2016;45(3):107-115.
153. Idri A, Bachiri M, Fernandez-Aleman JL. A Framework for Evaluating the Software Product Quality of Pregnancy Monitoring Mobile Personal Health Records. *Journal of medical systems.* 2016;40(3):1-17.
154. Risling T, Martinez J, Young J, Thorp-Froslie N. Evaluating Patient Empowerment in Association With eHealth Technology: Scoping Review. *J Med Internet Res.* 2017;19(9):e329.
155. Roehrs A, da Costa CA, Righi RD, de Oliveira KS. Personal Health Records: A Systematic Literature Review. *J Med Internet Res.* 2017;19(1):e13.

156. Prey JE, Polubriaginof F, Kuperman GJ, Tiase V, Collins SA, Vawdrey DK. International perspectives on sharing clinical data with patients. *International journal of medical informatics.* 2016;86:135-141.
157. van Mens HJT, Duijm RD, Nienhuis R, de Keizer NF, Cornet R. Determinants and outcomes of patient access to medical records: Systematic review of systematic reviews. *International journal of medical informatics.* 2019;129:226-233.
158. Craven CK, Doebbeling B, Furniss D, Holden RJ, Lau F, Novak LL. Evidence-based Health Informatics Frameworks for Applied Use. *Stud Health Technol Inform.* 2016;222:77-89.
159. Lau F, Kuziemsky C. *Handbook of eHealth evaluation: an evidence-based approach.* Victoria: University of Victoria; 2016.
160. Ahmadi H, Darvishi M, Nilashi M, et al. Evaluating the critical factors for electronic medical record adoption using fuzzy approaches. *International Journal of Innovation and Scientific Research.* 2014;9(2):268-284.
161. Ahmadi H, Rad MS, Nilashi M, Ibrahim O, Almaee A. Ranking the Micro level critical factors of electronic medical records adoption using TOPSIS method. *Health Informatics.* 2013;4(2).
162. Bassi J, Lau F, Lesperance M. Perceived impact of electronic medical records in physician office practices: a review of survey-based research. *Interactive journal of medical research.* 2012;1(2):e3.
163. Gagnon MP, Payne-Gagnon J, Sicotte C, Langue-Dube JA, Motulsky A. Connecting primary care clinics and community pharmacies through a nationwide electronic prescribing network: A qualitative study. *Journal of innovation in health informatics.* 2015;22(3):359-367.
164. Lau F, Partridge C, Randhawa G, Bowen M. Applying the clinical adoption framework to evaluate the impact of an ambulatory electronic medical record. *Stud Health Technol Inform.* 2013;183:15-20.
165. Cohen J. A coefficient of agreement for nominal scales. *Educational and psychological measurement.* 1960;20(1):37-46.
166. Cohen J. Weighted kappa: Nominal scale agreement provision for scaled disagreement or partial credit. *Psychological bulletin.* 1968;70(4):213.
167. Walker J, Leveille S, Bell S, et al. OpenNotes After 7 Years: Patient Experiences With Ongoing Access to Their Clinicians' Outpatient Visit Notes. *Journal of medical Internet research.* 2019;21(5):e13876.
168. Delbanco T, Walker J, Bell SK, et al. Inviting Patients to Read Their Doctors' Notes: A Quasi-experimental Study and a Look Ahead. *Annals of internal medicine.* 2012;157(7):461-U436.
169. Esch T, Mejilla R, Anselmo M, Podtschaske B, Delbanco T, Walker J. Engaging patients through open notes: an evaluation using mixed methods. *BMJ Open.* 2016;6(1):e010034.
170. Wolff JL, Darer JD, Berger A, et al. Inviting patients and care partners to read doctors' notes: OpenNotes and shared access to electronic medical records. *Journal of the American Medical Informatics Association : JAMIA.* 2016.
171. Keselman A, Slaughter L, Smith CA, et al. Towards consumer-friendly PHRs: patients' experience with reviewing their health records. *Amia* 2007;Annual Symposium proceedings / AMIA Symposium. AMIA Symposium.:399-403.
172. Wernick M, Hale P, Anticich N, et al. A randomised crossover trial of minimising medical terminology in secondary care correspondence in patients with chronic health conditions: impact on understanding and patient reported outcomes. *Intern Med J.* 2016;46(5):596-601.

173. Lalor JP, Woolf B, Yu H. Improving Electronic Health Record Note Comprehension With NoteAid: Randomized Trial of Electronic Health Record Note Comprehension Interventions With Crowdsourced Workers. *Journal of medical Internet research.* 2019;21(1):e10793.
174. Choi YK, Kirchhoff K, Turner AM. Medical Text Simplification by Medical Trainees: A Feasibility Study. Paper presented at: 2016 IEEE International Conference on Healthcare Informatics (ICHI)2016.
175. Kandula S, Curtis D, Zeng-Treitler Q. A Semantic and Syntactic Text Simplification Tool for Health Content. *AMIA Annual Symposium Proceedings.* 2010;2010:366-370.
176. Teixeira M, Cook DA, Heale BSE, Del Fiol G. Optimization of infobutton design and Implementation: A systematic review. *Journal of Biomedical Informatics.* 2017;74(Supplement C):10-19.
177. Cook DA, Teixeira MT, Heale BSE, Cimino JJ, Del Fiol G. Context-sensitive decision support (infobuttons) in electronic health records: a systematic review. *Journal of the American Medical Informatics Association.* 2017;24(2):460-468.
178. Fransen M, Van Schaik T, Twickler T, Essink-Bot M. Applicability of internationally available health literacy measures in the Netherlands. *Journal of health communication.* 2011;16(sup3):134-149.
179. Zeng-Treitler Q, Goryachev S, Kim H, Keselman A, Rosendale D. Making texts in electronic health records comprehensible to consumers: a prototype translator. *AMIA Annual Symposium proceedings / AMIA Symposium AMIA Symposium.* 2007:846-850.
180. SNOMED CT Netherlands National Release Center. SNOMED CT Netherlands/ Nederlands Edition, March 31, 2017. Nictiz. https://mlds.ihtsdotools.org/. Published 2017. Accessed.
181. Diagnosethesaurus, Version 2.5, June 22, 2017. Dutch Hospital Data. https://www.dhd.nl/. Published 2017. Accessed2017.
182. Terminology Services Guide. SNOMED International. Published 2017. Accessed.
183. Cornet R, Van Eldik A, De Keizer N. Inventory of tools for Dutch clinical language processing. *Stud Health Technol Inform.* 2012;180:245-249.
184. van Mens HJT, de Keizer NF, Nienhuis R, Cornet R. Clarifying Diagnoses to Laymen by Employing the SNOMED CT Hierarchy. *Stud Health Technol Inform.* 2018;247:900-904.
185. Zeng QT, Tse T, Divita G, et al. Term Identification Methods for Consumer Health Vocabulary Development. *Journal of medical Internet research.* 2007;9(1):e4.
186. Keselman A, Tse T, Crowell J, Browne A, Ngo L, Zeng Q. Assessing Consumer Health Vocabulary Familiarity: An Exploratory Study. *Journal of medical Internet research.* 2007;9(1):e5.
187. Chen J, Jagannatha AN, Fodeh SJ, Yu H. Ranking Medical Terms to Support Expansion of Lay Language Resources for Patient Comprehension of Electronic Health Record Notes: Adapted Distant Supervision Approach. *JMIR medical informatics.* 2017;5(4):e42.
188. He Z, Chen Z, Oh S, Hou J, Bian J. Enriching consumer health vocabulary through mining a social Q&A site: A similarity-based approach. *Journal of biomedical informatics.* 2017;69:75-85.
189. Ibrahim M, Gauch S, Salman O, Alqahtani M. An automated method to enrich consumer health vocabularies using GloVe word embeddings and an auxiliary lexical resource. *PeerJ Computer science.* 2021;7:e668.

190. Zeng-Treitler Q, Goryachev S, Kim H, Keselman A, Rosendale D. Making texts in electronic health records comprehensible to consumers: a prototype translator. *Amia* 2007;Annual Symposium proceedings / AMIA Symposium. AMIA Symposium.:846-850.
191. Chen J, Druhl E, Polepalli Ramesh B, et al. A Natural Language Processing System That Links Medical Terms in Electronic Health Record Notes to Lay Definitions: System Development Using Physician Reviews. *Journal of medical Internet research.* 2018;20(1):e26.
192. Hallgren KA. Computing inter-rater reliability for observational data: an overview and tutorial. *Tutorials in quantitative methods for psychology.* 2012;8(1):23.
193. Cicchetti DV. Guidelines, criteria, and rules of thumb for evaluating normed and standardized assessment instruments in psychology. *Psychological assessment.* 1994;6(4):284.
194. Kloehn N, Leroy G, Kauchak D, et al. Improving Consumer Understanding of Medical Text: Development and Validation of a New SubSimplify Algorithm to Automatically Generate Term Explanations in English and Spanish. *Journal of medical Internet research.* 2018;20(8):e10779.
195. Ramadier L, Lafourcade M. Radiological Text Simplification Using a General Knowledge Base. 2018; Cham.
196. Liang SF, Stevens R, Scott D, Rector A. Automatic verbalisation of SNOMED classes using OntoVerbal. Paper presented at: Conference on Artificial Intelligence in Medicine in Europe2011.
197. Kanhov M, Feng X, Dalianis H. Natural Language Generation from SNOMED Specifications. Paper presented at: CLEF (Online Working Notes/Labs/Workshop)2012.
198. Lyudovyk O, Weng C. SNOMEDtxt: Natural Language Generation from SNOMED Ontology. *Stud Health Technol Inform.* 2019;264:1263-1267.
199. Nictiz. Taal en techniek ontmoeten elkaar bij brainstorm over patiëntvriendelijke termen (in Dutch). https://www.nictiz.nl/nieuws/taal-en-techniek-ontmoeten-elkaar-bij-brainstorm-over-patientvriendelijke-termen/. Published 2019. Accessed January, 24, 2022.
200. van Mens HJT, Martens SSM, Paiman EHM, et al. Diagnosis clarification by generalization to patient-friendly terms and definitions: Validation study. *J Biomed Inform.* 2022;129:104071.
201. SNOMED CT Netherlands National Release Center. SNOMED CT Patient-Friendly Extension Release. Nictiz. https://mlds.ihtsdotools.org/. Published 2022. Accessed September 21, 2022.
202. Franciscus Gasthuis & Vlietland. Jaarverslag 2021 (Annual Report). https://franciscus.maglr.com/franciscus-jaarverslag-2021/jaarverslag. Published 2021. Accessed September 21, 2022.
203. van Mens HJT, Hannen G, Nienhuis R, Bolt RJ, de Keizer NF, Cornet R. Better explanations for diagnoses in your medical record: use and ratings of an information button to clarify diagnoses in a hospital patient portal. *ISRCTN.* 2022.
204. Nachar N. The Mann-Whitney U: A test for assessing whether two independent samples come from the same distribution. *Tutorials in quantitative Methods for Psychology.* 2008;4(1):13-20.
205. Patefield WM. Algorithm AS 159: An Efficient Method of Generating Random R × C Tables with Given Row and Column Totals. *Journal of the Royal Statistical Society Series C (Applied Statistics).* 1981;30(1):91-97.

206. Benjamini Y, Yekutieli D. The Control of the False Discovery Rate in Multiple Testing under Dependency. *The Annals of Statistics.* 2001;29(4):1165-1188.
207. Lalor JP, Hu W, Tran M, Wu H, Mazor KM, Yu H. Evaluating the Effectiveness of NoteAid in a Community Hospital Setting: Randomized Trial of Electronic Health Record Note Comprehension Interventions With Patients. *Journal of medical Internet research.* 2021;23(5):e26354.
208. Alfano M, Lenzitti B, Bosco GL, Muriana C, Piazza T, Vizzini G. Design, Development and Validation of a System for Automatic Help to Medical Text Understanding. *Int J Med Inf.* 2020:104109.
209. Bala S, Keniston A, Burden M. Patient Perception of Plain-Language Medical Notes Generated Using Artificial Intelligence Software: Pilot Mixed-Methods Study. *JMIR formative research.* 2020.
210. Chimowitz H, Gerard M, Fossa A, Bourgeois F, Bell SK. Empowering Informal Caregivers with Health Information: OpenNotes as a Safety Strategy. *Jt Comm J Qual Patient Saf.* 2018;44(3):130-136.
211. Barredo Arrieta A, Díaz-Rodríguez N, Del Ser J, et al. Explainable Artificial Intelligence (XAI): Concepts, taxonomies, opportunities and challenges toward responsible AI. *Information Fusion.* 2020;58:82-115.
212. Payrovnaziri SN, Chen Z, Rengifo-Moreno P, et al. Explainable artificial intelligence models using real-world electronic health record data: a systematic scoping review. *Journal of the American Medical Informatics Association : JAMIA.* 2020;27(7):1173-1185.
213. Jones SS, Rudin RS, Perry T, Shekelle PG. Health information technology: an updated systematic review with a focus on meaningful use. *Annals of internal medicine.* 2014;160(1):48-54.
214. Versnellingsprogramma Informatie-uitwisseling Patiënt en Professional: de patiënt meer inzicht in zijn eigen zorg. Nederlandse Vereniging van Ziekenhuizen. https://web.archive.org/web/20221102220206/https://www.vipp-programma.nl/. Published 2023. Accessed January 4, 2023.
215. Gegevensuitwisseling in de zorg. Ministry of Health, Welfare and Sport. https://web.archive.org/web/20230104143249/https://www.gegevensuitwisselingindezorg.nl/gegevensuitwisseling. Published 2023. Accessed January 4, 2023.
216. European Health Data Space. European Commission. https://web.archive.org/web/20230104143729/https://health.ec.europa.eu/ehealth-digital-health-and-care/european-health-data-space_en. Published 2023. Accessed January 4, 2023.
217. Hussein R, Scherdel L, Nicolet F, Martin-Sanchez F. Towards the European Health Data Space (EHDS) ecosystem: A survey research on future health data scenarios. *International journal of medical informatics.* 2022;170:104949.
218. Horgan D, Hajduch M, Vrana M, et al. European Health Data Space-An Opportunity Now to Grasp the Future of Data-Driven Healthcare. *Healthcare (Basel).* 2022;10(9).
219. FAIR Principles. GO FAIR. https://web.archive.org/web/20230103214140/https://www.go-fair.org/fair-principles/. Published 2023. Accessed January 4, 2023.
220. Kersloot MG, van Damme P, Abu-Hanna A, Arts DL, Cornet R. FAIRification Efforts of Clinical Researchers: The Current State of Affairs. *Stud Health Technol Inform.* 2021;287:35-39.
221. Kalyan KS, Sangeetha S. BertMCN: Mapping colloquial phrases to standard medical concepts using BERT and highway network. *Artif Intell Med.* 2021;112:102008.

222. Kersloot MG, van Putten FJP, Abu-Hanna A, Cornet R, Arts DL. Natural language processing algorithms for mapping clinical text fragments onto ontology concepts: a systematic review and recommendations for future studies. *J Biomed Semantics.* 2020;11(1):14-14.
223. Kouris P, Alexandridis G, Stafylopatis A. Abstractive Text Summarization: Enhancing Sequence-to-Sequence Models Using Word Sense Disambiguation and Semantic Content Generalization. *Computational Linguistics.* 2021;47(4):813-859.
224. Pauws S, Gatt A, Krahmer E, Reiter E. Making effective use of healthcare data using data-to-text technology. In: *Data Science for Healthcare.* Springer; 2019:119-145.
225. Schmeets H, Cornips L. Talen en dialecten in Nederland: Wat spreken we thuis en wat schrijven we op sociale media? Centraal Bureau voor Statistiek. http://web.archive.org/web/20230103202446/https://www.cbs.nl/nl-nl/longread/statistische-trends/2021/talen-en-dialecten-in-nederland. Published 2021. Accessed January 3, 2023.
226. Dietrich S, Hernandez E. What Languages Do We Speak in the United States? Nearly 68 Million People Spoke a Language Other Than English at Home in 2019. United States Census Bureau. http://web.archive.org/web/20230103203947/https://www.census.gov/library/stories/2022/12/languages-we-speak-in-united-states.html. Published 2022. Accessed January 3, 2023.
227. Pancreas. Wikimedia. https://commons.wikimedia.org/wiki/Pancreas. Published 2021. Accessed January 2, 2023.
228. Colorectal Cancer. MedlinePlus. https://medlineplus.gov/colorectalcancer.html. Published 2023. Accessed January 2, 2023.
229. van Mens HJT, Duijm RD, Nienhuis R, de Keizer NF, Cornet R. Towards an Adoption Framework for Patient Access to Electronic Health Records: Systematic Literature Mapping Study. *JMIR medical informatics.* 2020;8(3):e15150.
230. van Mens HJT, van Eysden MM, Nienhuis R, van Delden JJM, de Keizer NF, Cornet R. Evaluation of lexical clarification by patients reading their clinical notes: a quasi-experimental interview study. *BMC medical informatics and decision making.* 2020;20(Suppl 10):278.
231. van Mens HJT, Hannen GEG, Nienhuis R, Bolt RJ, de Keizer NF, Cornet R. Evaluation of Patient-Friendly Diagnosis Clarifications in a Hospital Patient Portal (personal communication; not published yet).
232. Turner AM, Facelli JC, Jaspers M, et al. Solving Interoperability in Translational Health. Perspectives of Students from the International Partnership in Health Informatics Education (IPHIE) 2016 Master Class. *Applied clinical informatics.* 2017;8(2):651-659.
233. van Mens HJT. Utilizing the SNOMED CT hierarchy to generate patient-friendly clarifications: challenges and opportunities. In. *SNOMED CT Web Series*. YouTube2021.
234. Keur M. Betere communicatie arts-patiënt dankzij Eenheid van Taal. Qruxx tech. https://web.archive.org/web/20190225185221/https://tech.qruxx.com/betere-communicatie-arts-patient-dankzij-eenheid-van-taal/. Published 2019. Accessed February 25, 2019.
235. Zweers M. Software ondersteunt communicatie tussen arts en patiënt. Qruxx. https://web.archive.org/web/20200809232115/https://www.qruxx.com/software-ondersteunt-communicatie-tussen-arts-en-patient/. Published 2019. Accessed August 9, 2020.

236. Bukman B. Betere communicatie arts-patiënt dankzij Eenheid van Taal. Zorgvisie. https://web.archive.org/web/20190111192211/https://www.zorgvisie.nl/betere-communicatie-arts-patient-dankzij-eenheid-van-taal/. Published 2019. Accessed January 11, 2019.
237. Bruggen slaan tussen patiënt en arts. Customerfirst! https://web.archive.org/web/20191212200704/https://customerfirst.nl/nieuws/2019/01/bruggen-slaan-tussen-patient-en-arts/index.xml. Published 2019. Accessed December 12, 2019.
238. Blog: Verrijken van zorg en wetenschap. ChipSoft. https://web.archive.org/web/20210419133032/https://www.chipsoft.nl/nieuws/311/Blog-Verrijken-van-zorg-en-wetenschap. Published 2019. Accessed April 19, 2021.
239. Nictiz. Nictiz. https://web.archive.org/web/20190705102628/https://www.nictiz.nl/nieuws/taal-en-techniek-ontmoeten-elkaar-bij-brainstorm-over-patientvriendelijke-termen/. Published 2019. Accessed July 5, 2019.
240. Determinants and outcomes of patient access to medical records Amsterdam UMC, location AMC. https://web.archive.org/web/20190705100242/https://www.amc.nl/web/research-75/departments/medical-informatics-1/medical-informatics-1/-determinants-and-outcomes-of-patient-access-to-medical-records-.htm. Published 2019. Accessed July 5, 2019.
241. Promoveren en software ontwikkelen tegelijk. ChipSoft. https://web.archive.org/web/20210416072808/https://www.chipsoft.nl/nieuws/363/Promoveren-en-software-ontwikkelen-tegelijk. Published 2019. Accessed April 16, 2021.
242. Van Medische informatiekunde naar workshop over eenheid van taal. ChipSoft. https://web.archive.org/web/20210416085753/https://www.chipsoft.nl/nieuws/447/Van-Medische-informatiekunde-naar-workshop-over-eenheid-van-taal. Published 2020. Accessed April 16, 2021.
243. “Algoritme maakt medische terms begrijpelijk voor de patiënt”. ChipSoft. https://web.archive.org/web/20220523142733/https://www.chipsoft.nl/nieuws/707/Algoritme-maakt-medische-termen-begrijpelijk-voor-de-patient. Published 2022. Accessed May 23, 2022.
244. Nieuw algoritme zet diagnose om in begrijpelijke taal. ICT&health. https://web.archive.org/web/20220512090057/https://www.icthealth.nl/nieuws/algoritme-onderzoek/. Published 2022. Accessed May 12, 2022.

About the author

Hugo Johan Theodoor van Mens, son of Theo van Mens and Alda van Mens – Peper, was born on August 21st, 1986 in Naarden, the Netherlands. He grew up in de Hilversumse Meent with his parents and sister Leonieke. He graduated from scientific level (vwo) secondary education at St. Vituscollege in Bussum.

He started the interdisciplinary bachelor Beta Gamma at the University of Amsterdam and took several courses in philosophy of science. Later he moved to Amstelveen and switched to the transdisciplinary track Future Planet Studies of the Beta Gamma bachelor, about sustainable development. He founded the charity bookstore chapter Books 4 Life Amsterdam with volunteers from Oxfam Novib and visited Rwanda two summers in a private initiative to volunteer at a university and other organizations in Muhanga with students from the University of Amsterdam and Florida State University. After attempting to major in civil engineering in Delft within the bachelor Beta Gamma Future Planet Studies track in Amsterdam, he decided to major in Medical Informatics instead at the Academic Medical Center in Amsterdam (currently part of Amsterdam UMC), to move to Amsterdam, quit voluntary work and focus on his studies.

During his major, he worked as a research assistant at the Department of Medical Informatics on a web application to improve and evaluate alternative user interfaces for data encoding with terminology systems. After finishing his bachelor's thesis he was offered to start a PhD research program already during the beginning of his master's in Medical Informatics. However, after graduating with his master's he decided to work at the Department of Research and Development, ChipSoft as a software developer. There he worked on Care Portal, the web application layer of the health information system HiX. Together with Amsterdam UMC and ChipSoft he started a PhD research program about patient portals and semantic interoperability for two days per week, next to working three days as a software developer at Care Portal.

Hugo, at the time of writing this thesis, lives in Amsterdam with his wife Tania Estrella van Mens – Gutiérrez Calvo, and enjoys playing music, going to concerts and festivals, playing tennis, and learning about nature, culture, philosophy and sustainable development.

Portfolio

PhD candidate	Hugo J.T. van Mens
Period	November 2016 – January 2023
Supervisors	Prof. dr. Nicolette F. de Keizer Dr. ir. Ronald Cornet

PhD Training

General courses

Description	Organization	Year	ECTS
Practical Biostatistics: e-learning (grade: 9 out of 10)	Amsterdam UMC Doctoral School (Doctoral School)	2017	1.4
Advanced Topics in Biostatistics	Doctoral School	2018	2.0
BROK registration: Basic course on regulations and organization for clinical investigators (grade: 94% out of 100%)	Examenbureau Medisch-Wetenschappelijk Onderzoeker (EMWO)	2018	1.0
Research Data Management	Doctoral School	2019	0.7
E-Science	Doctoral School	2019	0.7
BROK Re-registration	EMWO	2022	0.6
Total			**6.5**

Specific courses

Description	Organization	Year	ECTS
Clinical Epidemiology: Randomized Controlled Trials	Doctoral School	2017	0.6
Clinical Epidemiology: Observational Epidemiology (grade 9.5 out of 10)	Doctoral School	2017	0.6
Communication with Patients	Doctoral School	2018	0.3
Qualitative Health Research	Doctoral School	2019	1.9
Clinical Epidemiology: Evaluation of Medical Tests	Doctoral School	2020	0.9
Total			**4.3**

Seminars, workshops and masterclasses

Description	Location	Year	ECTS
International Partnership in Health Informatics Education (IPHIE) 2016 Masterclass	Salt Lake City, Utah, USA	2016	1.4
PhD candidate day (Promovendidag), Department of Medical Informatics	Breukelen and Amsterdam, The Netherlands	2017 2018 2019 2021	1.2
Amsterdam Public Health (APH) Annual Meeting	Amsterdam, The Netherlands	2017 2018	0.6
Working groups on patient-friendly terminology, Nictiz	Utrecht, The Netherlands	2019 2020	1.2
HIMMS SIIM Language Translation working group sessions	Online (Teams)	2021 2022	0.2
Total			**4.6**

Presentations

Description	Location	Year	ECTS
Interdisciplinary Alumni Network Amsterdam TED talks 'The medical record: ownership, content and use' (in Dutch)	Amsterdam, The Netherlands	2017	0.5
Medical Informatics Europe (MIE) 2018 conference 'Clarifying diagnoses to laymen by employing the SNOMED CT hierarchy'	Gothenburg, Sweden	2018	0.5
ChipSoft Care Portal User Day 'Patient-friendly health record without additional administrative burden?' (in Dutch)	Utrecht, The Netherlands	2018	0.5
ChipSoft I&S presentation 'Patient-friendly health record without additional administrative burden?' (in Dutch)	Amsterdam, The Netherlands	2018	0.5
ChipSoft R&D Seminar 'Graph vs. relational database: time for a new DBMS for HiX?' (in Dutch)	Bussum, The Netherlands	2018	0.5
ChipSoft Care Portal 'Statistics and A/B tests Care Portal' (in Dutch)	Amsterdam, The Netherlands	2019	0.5
Nictiz working group on patient-friendly terminology 'Dynamic patient-friendly descriptions based on SNOMED CT logic'	Utrecht, The Netherlands	2019	0.5
Nictiz working group on patient-friendly terminology 'Testing patient-friendly terms and definitions from various sources and terminology organizations'	Amsterdam, The Netherlands	2019	0.5
SNOMED Symposium 'Uniformity of language for patients' presentation together with Nictiz (in Dutch)	Zeist, The Netherlands	2020	0.5
ChipSoft CMIO Network pitch 'Statistics required to develop patient-friendly terminology' (in Dutch)	Online (Teams)	2020	0.5
SNOMED CT Research Webinar 'Utilizing the SNOMED CT hierarchy to generate patient-friendly clarifications: challenges and opportunities'	Online (Zoom and YouTube)	2021	0.5
HIMMS SIIM Language Translation working group 'Utilizing the SNOMED CT hierarchy to generate patient-friendly clarifications: challenges and opportunities'	Online (Teams)	2021	0.5
SNOMED CT Expo 'Patient-friendly language in a hospital patient portal: implementation and evaluation'	Lisbon, Portugal	2022	0.5
Total			**6.5**

Conferences

Description	Location	Year	ECTS
Medical Informatics Europe (MIE) 2018	Gothenburg, Sweden	2018	0.9
Medisch Informatica Congres (MIC) 2018	Antwerp, Belgium	2018	0.4
SNOMED Symposium	Zeist, The Netherlands	2020	0.3
2nd Dutch Meeting on Clinical NLP	Nijmegen, The Netherlands	2021	0.3
SNOMED CT Expo	Lisbon, Portugal	2022	0.6
Total			**2.5**

Other

Description	Location	Year	ECTS
Attending and contributing to research and staff meetings, Department of Medical Informatics, Amsterdam UMC	Amsterdam, The Netherlands	2017 - 2022	2.0
Attending and contributing to research group Reusable Health Data meetings, Department of Medical Informatics, Amsterdam UMC	Amsterdam, The Netherlands	2020 - 2022	1.0
Total			**3.0**

Teaching

Lecturing

Description	Location	Year	ECTS
BSc Medical informatics first year students alumni tour 'From medical informatics to PhD research and R&D' (in Dutch)	ChipSoft, Amsterdam, The Netherlands	2016 2018 2019	1.5
FHIR developer days talk for students from various international universities 'R&D at ChipSoft'	ChipSoft, Amsterdam, The Netherlands	2016 - 2018	1.5
IPHIE 2018 talk for international students 'Patient-friendly medical record without administrative burden?'	ChipSoft, Amsterdam, The Netherlands	2018	0.5
BSc Medical Informatics MiX 2.2 course knowledge clip 'Natural language generation' (in Dutch)	BSc Medical Informatics, online learning environment	2020	1.0
MSc Medical informatics students masterclass 'R&D: Patient Access to EHRs'	Online (Zoom)	2021	0.5
Total			**5.0**

Supervision

Description	Year	ECTS
Karen Goes, MSc Computational Linguistics. Part-time collection of text corpora for natural language processing and patient-friendly terminology.	2019	1.0
Mirte van Eysden, MA Journalism, Bachelor of Medicine. Five-month Master of Medicine scientific internship on the influence of lexical clarification on patients understanding their clinical notes.	2020	1.0
Li-anne Tjin, BSc Artificial Intelligence. Four-month internship on clustering to find popular medical terms.	2020	1.0
Daan de Regt. Five-month BSc Medical Informatics internship and bachelor thesis on collection of Dutch medical corpora, the development of an infrastructure.	2021	1.0
Pien van Putten, BSc Medical Informatics. Eight-month MSc Medical Informatics internship and master thesis on design of a tool to convert structured diagnoses of neoplasms and infectious diseases into patient-friendly explanations, focussed on patient portals.	2022	2.0
Kim Bladder. Two-month BSc Artificial Intelligence internship and bachelor thesis on determining the complexity of Dutch medical terms by using a familiarity metric.	2022	1.0
Total		**7.0**

Publications

Peer reviewed full papers in this thesis

Description	Year
Hugo J.T. van Mens, Nicolette F. de Keizer, Remko Nienhuis, Ronald Cornet. Clarifying Diagnoses to Laymen by Employing the SNOMED CT Hierarchy. *Studies in Health Technology and Informatics*, *247*, 900-904. DOI: 10.3233/978-1-61499-852-5-900.[184]	2018
Hugo J.T. van Mens, Ruben D. Duijm, Remko Nienhuis, Nicolette F. de Keizer, Ronald Cornet. Determinants and Outcomes of Patient Access to Medical Records: Systematic Review of Systematic Reviews. *International Journal of Medical Informatics*, *129*, 226-233. DOI: 10.1016/j.ijmedinf.2019.05.014.[157]	2019
Hugo J.T. van Mens, Ruben D. Duijm, Remko Nienhuis, Nicolette F. de Keizer, Ronald Cornet. Towards an Adoption Framework for Patient Access to Electronic Health Records: Systematic Literature Mapping Study. *JMIR Medical Informatics*, *8*(3), e15150. DOI: 10.2196/15150.[229]	2020
Hugo J.T. van Mens, Mirte M. van Eysden, Remko Nienhuis, Johannes J.M. van Delden, Nicolette F. de Keizer, Ronald Cornet. Evaluation of Lexical Clarification by Patients Reading Their Clinical Notes: a Quasi-Experimental Interview Study. *BMC Medical Informatics and Decision Making*, *20*(Suppl 10), 278. DOI: 10.1186/s12911-020-01286-9.[230]	2020
Hugo J.T. van Mens, Savine S.M. Martens, Elisabeth H.M. Paiman, Alexander C. Mertens, Remko Nienhuis, Nicolette F. de Keizer, Ronald Cornet. Diagnosis Clarification by Generalization to Patient-Friendly Terms and Definitions: Validation Study. *Journal of Biomedical Informatics*, *129*, 104071. DOI: 10.1016/j.jbi.2022.104071.[200]	2022
Hugo J.T. van Mens, Gaby E.G. Hannen, Remko Nienhuis, Roel J. Bolt, Nicolette F. de Keizer, Ronald Cornet. Evaluation of Patient-Friendly Diagnosis Clarifications in a Hospital Patient Portal. *Applied Clinical Informatics*. Published version: DOI: 10.1055/a-2067-5310.[231]	Accepted 2023

Other peer-reviewed full papers

Description	Year
Anne M. Turner, Julio C. Facelli, Monique W.M. Jaspers, Thomas Wetter, Daniel Pfeifer, Laël Cranmer Gatewood, Terry Adam, Yu-Chuan Li, Ming-Chin Lin, R. Scott Evans, Anna L. Beukenhorst, **Hugo J.T. van Mens**, Esmee Tensen, Christian Bock, Laura Fendrich, Peter Seitz, Julian Suleder, Ranyah Aldelkhyyel, Kent Bridgeman, Zhen Hu, Aaron Sattler, Shin-Yi Guo, Islam Md Mohaimenul Mohaimenul, Dina Nur Anggraini Ningrum, Hsin-Ru Tung, Jiantano Bian, Joseph M. Plasek, Casey Rommel, Juandalyn Burke, Harkirat Sohih. Solving Interoperability in Translational Health. Perspectives of Students from the International Partnership in Health Informatics Education (IPHIE) 2016 Master Class. *Applied Clinical Informatics*, *8*(2):651-659. DOI: 10.4338/ACI-2017-01-CR-0012.[232]	2016

Other publications

Description	Year
Hugo J.T. van Mens. Supervisors: Ronald Cornet, Nicolette F. de Keizer. Terminology system-based data encoding for intensive care: Deriving the APACHE-IV reasons for ICU admission classification through SNOMED CT and optimizing the user interface for diagnostic data entry. MSc Medical Informatics. (Master Thesis) DOI: 10.13140/RG.2.2.35624.67841.[50]	2016
Hugo J.T. van Mens, Remko Nienhuis, Ruben D. Duijm, Nicolette F. de Keizer, Ronald Cornet. An overview of determinants and outcomes of providing patients access to and control over their own medical data: A systematic review of reviews. *PROSPERO*, CRD42018084542. (Review Protocol Registration)[101]	2018
Hugo J.T. van Mens. Utilizing the SNOMED CT hierarchy to generate patient-friendly clarifications: challenges and opportunities. SNOMED CT Research Web Series. (Webinar)[233]	2021
Hugo J.T. van Mens, Gaby E.G. Hannen, Remko Nienhuis, Roel J. Bolt, Nicolette F. de Keizer, Ronald Cornet. Better explanations for diagnoses in your medical record: use and ratings of an information button to clarify diagnoses in a hospital patient portal. *ISRCTN*, ISRCTN59598141. (Protocol Registration) DOI: 10.1186/ISRCTN59598141.[203]	2022

Popular media

Description	Year
Better doctor-patient communication thanks to Unity of Language (in Dutch). Qruxx tech.[234]	2019
Software supports communication between doctor and patient (in Dutch). Qruxx.[235]	2019
Better doctor-patient communication thanks to Unity of Language (in Dutch). Zorgvisie. [236]	2019
Build bridges between patient and doctor (in Dutch). Customerfirst![237]	2019
Enriching care and science. ChipSoft.[238]	2019
Language and technology meet when brainstorming about patient-friendly terms (in Dutch). Nictiz.[239]	2019
Determinants and outcomes of patient access to medical records. Amsterdam UMC.[240]	2019
Doing a PhD and developing software at the same time (in Dutch). ChipSoft.[241]	2019
From medical informatics to workshop on unity of language (in Dutch). ChipSoft.[242]	2020
Algorithm makes medical terms understandable for patient (in Dutch). ChipSoft.[243]	2022
New algorithm converts diagnosis into understandable language (in Dutch). ICT&health.[244]	2022

Authors' contributions

Chapter 2

Hugo J.T. van Mens
Conceptualization, Methodology, Formal Analysis, Investigation, Data Curation, Writing – Original Draft, Visualization.

Ruben D. Duijm
Validation, Formal analysis, Investigation, Writing – Review & Editing.

Remko Nienhuis
Conceptualization, Methodology, Validation, Investigation, Writing – Review & Editing, Supervision.

Nicolette F. de Keizer
Conceptualization, Methodology, Validation, Investigation, Writing – Review & Editing, Supervision.

Ronald Cornet
Conceptualization, Methodology, Validation, Investigation, Writing – Review & Editing, Supervision.

Chapter 3

Hugo J.T. van Mens
Conceptualization, Methodology, Formal Analysis, Investigation, Data Curation, Writing – Original Draft, Visualization.

Ruben D. Duijm
Validation, Formal analysis, Investigation, Writing – Review & Editing.

Remko Nienhuis
Conceptualization, Methodology, Investigation, Validation, Writing – Review & Editing, Supervision.

Nicolette F. de Keizer
Conceptualization, Methodology, Investigation, Validation, Writing – Review & Editing, Supervision.

Ronald Cornet
Conceptualization, Methodology, Investigation, Validation, Writing – Review & Editing, Supervision.

Chapter 4

Hugo J.T. van Mens *
Conceptualization, Methodology, Formal Analysis, Investigation, Data Curation, Writing – Original Draft, Visualization, Project Administration.

Mirte M. van Eysden *
Conceptualization, Methodology, Investigation, Data Curation, Writing – Review & Editing, Project Administration.

Remko Nienhuis
Conceptualization, Methodology, Validation, Writing – Review & Editing, Supervision.

Johannes J.M. van Delden
Conceptualization, Methodology, Validation, Writing – Review & Editing, Supervision, Project Administration.

Nicolette F. de Keizer
Conceptualization, Methodology, Validation, Writing – Review & Editing, Supervision.

Ronald Cornet
Conceptualization, Methodology, Validation, Writing – Review & Editing, Supervision.

* contributed equally

Chapter 5

Hugo J.T. van Mens
Conceptualization, Methodology, Formal Analysis, Investigation, Data Curation, Writing – Original Draft, Visualization.

Nicolette F. de Keizer
Conceptualization, Methodology, Validation, Writing – Review & Editing, Supervision.

Remko Nienhuis
Conceptualization, Methodology, Validation, Writing – Review & Editing, Supervision.

Ronald Cornet
Conceptualization, Methodology, Validation, Writing – Review & Editing, Supervision.

Chapter 6

Hugo J.T. van Mens
Conceptualization, Methodology, Software, Validation, Formal analysis, Investigation, Resources, Data Curation, Writing - Original Draft, Visualization, Project administration.

Savine S.M. Martens
Investigation, Writing - Review & Editing.

Elisabeth H.M. Paiman
Investigation, Writing - Review & Editing.

Alexander C. Mertens
Software, Validation, Resources, Writing - Review & Editing.

Remko Nienhuis
Conceptualization, Methodology, Validation, Writing - Review & Editing, Supervision.

Nicolette F. de Keizer
Conceptualization, Methodology, Validation, Writing - Review & Editing, Supervision.

Ronald Cornet
Conceptualization, Methodology, Validation, Writing - Review & Editing, Supervision.

Chapter 7

Hugo J.T. van Mens
Conceptualization, Methodology, Software, Validation, Formal analysis, Investigation, Resources, Data Curation, Writing - Original Draft, Visualization, Project administration.

Gaby E.G. Hannen
Software, Validation, Resources, Writing - Review & Editing.

Remko Nienhuis
Conceptualization, Methodology, Validation, Writing - Review & Editing, Supervision.

Roel J. Bolt
Conceptualization, Methodology, Validation, Writing - Review & Editing.

Nicolette F. de Keizer
Conceptualization, Methodology, Validation, Writing - Review & Editing, Supervision.

Ronald Cornet
Conceptualization, Methodology, Validation, Writing - Review & Editing, Supervision.

Dankwoord (Acknowledgements)

Dit proefschrift is het resultaat van samenwerkingen met diverse mensen van verschillende organisaties, zonder wiens inzet en enthousiasme het er niet geweest was. Hierbij wil ik daarom graag iedereen bedanken die heeft bijgedragen aan de onderzoeksprojecten, softwareontwikkeling, implementatie en het promotietraject.

In de eerste plaats wil ik mijn promotoren Ronald en Nicolette bedanken voor de mogelijkheden die jullie mij hebben geboden, jullie geloof in mij en de prettige samenwerking. Het was een bijzonder leerzame tijd. Bij het volgen van vakken bij Medische informatiekunde tijdens mijn major hebben jullie mijn interesse aangewakkerd en al snel gevraagd om een webapplicatie te ontwikkelen en onderzoek te doen naar het op een gebruiksvriendelijke manier registreren van gegevens met medische terminologiesystemen. Ook toen ik na mijn master een andere richting in ben gegaan stonden jullie ervoor open om opnieuw een voor ons allen relevant onderzoekstraject op te zetten. Ronald bedankt voor je begeleiding en het overbrengen van je enthousiasme over medische terminologieën en gegevensuitwisseling. Nicolette bedankt voor het delen van je methodologische kennis en begeleiding. Het is bewonderenswaardig hoe jullie altijd klaarstaan voor jullie promovendi.

Remko, ik wil je bedanken voor hoe je eerst als teamleider, later als manager en later persoonlijk betrokken was bij het promotieonderzoek. Je was eigenlijk ook als een soort promotor. Bedankt voor de prettige samenwerking, je interesse in het opzetten van het wetenschappelijk onderzoek en het mogelijk maken van het promotietraject in de vorm van een samenwerking tussen ChipSoft en Amsterdam UMC.

De artikelen heb ik niet alleen geschreven, daarom wil ik ook mijn coauteurs die hierboven niet genoemd zijn bedanken voor hun bijdrage. Ruben, dank je voor je bijdrage aan de systematische review. Mirte en Hans, dank jullie voor jullie bijdragen aan het onderzoek bij UMC Utrecht. Savine, Malou en Sander, dank jullie voor jullie bijdragen aan de validatie van de diagnosetoelichtingen. Gaby en Roel, dank jullie voor jullie bijdragen aan de evaluatie van de toelichtingen in het patiëntenportaal van Franciscus Gasthuis & Vlietland. Karen, Li-anne, Daan, Pien en Kim dank jullie voor het volgen van een stage, dat was voor mij ook leerzaam en relevant voor het onderzoek dat in dit proefschrift is beschreven. De promotiecommissie wil ik bedanken voor het lezen en beoordelen van het

proefschrift en alvast voor de discussie tijdens de verdediging, waar ik naar uitkijk. Remko, Sander en Leonieke wil ik bedanken dat jullie mijn paranimfen zijn.

Al mijn collega's wil ik bedanken voor de gezelligheid, lunches, koffies, discussies, borrels, het meedenken en jullie bijdragen. De KiK'kers wil ik bedanken, in het bijzonder kamergenoten van J1B-109 en onderzoeksgroepsgenoten van Reusable Health Data. De ChipSoft-collega's wil ik bedanken, in het bijzonder van R&D Zorportaal en I&S Zorgportaal, R&D Techniek, UX en DCR. Velen van jullie hebben op de een of andere manier bijgedragen aan de resultaten van het onderzoek dat beschreven is in dit proefschrift. Met name Arthur wil ik bedanken voor het mogelijk maken om software te ontwikkelen die voor het onderzoek nodig was, daar meer mee te gaan doen en om me te betrekken bij de samenwerking met UMC Utrecht. Hans, Robert en George wil ik bedanken voor het steunen van het promotietraject, het formeel aangaan van de samenwerkingen met andere partijen en zo een bijdrage aan het wetenschappelijk onderzoek te leveren.

Lieve ouders, de bevlogenheid, nieuwsgierigheid en brede interesse heb ik voor een groot deel aan jullie te danken. Ik ben ontzettend dankbaar voor jullie onvoorwaardelijke steun en dat jullie mijn ouders zijn. Lieve Leonieke, mijn zusje, bedankt voor je vriendschap, gezelligheid en ondersteuning, het is fijn dat we het zo goed met elkaar kunnen vinden. En erg leuk dat Jenna en Noah deel uitmaken van onze familie.

Lieve schat, Tania, mariposita maravillosa. Bedankt voor je liefde, aanmoediging, interesse en geloof in mij als onderzoeker. Ik bewonder je gezelligheid, doorzettingsvermogen, inzicht, enthousiasme, schoonheid en vrolijkheid.

Summary

From jargon to clarity: On patient access to electronic health records and patient-friendly terminology

Personal health records (PHRs) and patient portals enable patients and their significant others to access their electronic health records (EHRs). This is supposed to help them learn more about their medical condition, prepare for consultations, remember what had been discussed and take care of themselves. Thus it is important to facilitate the adoption of patient portals and PHRs. However, several reviews on this topic addressed different aspects from different perspectives, which makes it difficult to get an overview of the determinants and outcomes patient access to EHRs, patient portals and PHRs. In particular, medical data are difficult to comprehend. Especially persons with limited health literacy experience barriers accessing their EHRs and using medical data and information to the benefit of their health. Most data in EHRs are hidden in free text, but increasingly data are encoded with standardized terminology systems, which makes it easier to reuse data for secondary purposes, such as obtaining statistics on morbidity and mortality. Consumer-oriented terminology bridges the gap between medical and consumer language, such as the Dutch Thesaurus Zorg en Welzijn (TZW, in English "Thesaurus Care and Well-being"). Additionally, the SNOMED CT Netherlands edition, a comprehensive standardized medical terminology system, includes patient-friendly terms and definitions. However, only a few medical concepts could be clarified with these terms and definitions. This thesis provides an overview of determinants and outcomes of patient access to EHRs and proposes and evaluates applications to clarify medical terms to patients using terminology systems.

In Chapter 2, we systematically reviewed systematic reviews on patient access to EHRs, patient portals and PHRs, which are types of health information systems (HISs). We synthesized the determinants and outcomes from these reviews into the Clinical Adoption Framework (CAF) and assessed their risk of bias with the AMSTAR2 checklist. The CAF is an integrative framework to evaluate HIS adoption. The reviews had several critical flaws according to the AMSTAR2 checklist, and therefore quality restrictions should be taken into account when interpreting their results. Reviews reported on the CAF dimensions of people, organization, implementation, HIS quality, HIS use, and benefits of HIS use. Few reviews found associations on the macro level of healthcare standards, legislation, policy, government, funding, incentives and social, political and economic trends, which indicates more research is needed on those levels.

There was little guidance documentation about applying the CAF and it was not specifically designed to consider patients as end-users of HISs. Therefore, in Chapter 3, we reported how we elaborated upon the CAF to include patients as users of HISs and defined inclusion and exclusion criteria for the CAF categories. We found that the interrater reliability was modest. The guidance documentation can be used in further research on patient access to EHRs.

In the patient portals of the university hospital UMC Utrecht, lexical clarification functionality was implemented to clarify medical terms in free-text clinical notes. The clarifications consisted of synonyms and definitions from the TZW. In Chapter 4, we evaluated this functionality to assess whether it met the information needs of the patients and how the functionality could be improved. We included fifteen patients. They had adequate health literacy. The participants marked and rated the ease of the terms they found difficult in a selection of clinical notes from their EHRs. Next, the clinical notes were read again but with the clarifications provided by the functionality. The participants rated the terms and clarifications on their ease and the clarifications on their usefulness. We found that the functionality had a median precision of 6.5% (interquartile range, IQR: 2.3 – 14.3%) and recall of 8.3% (IQR: 4.7 – 13.5%) per patient. This means that the functionality only clarified few terms that the patients found difficult and many other terms that they did not consider to be difficult. Despite this low precision and recall, the patients considered the functionality to be easy to use and useful.

In the Netherlands, diagnoses are registered with the coding system Diagnosethesaurus ("Diagnosis Thesaurus" in English). Because the Diagnosethesaurus is mapped to SNOMED CT, the patient-friendly terms from SNOMED CT could be used to clarify diagnoses to patients. However, the coverage of diagnoses by patient-friendly terms was only 1.2%. In Chapter 5, we proposed to generalize diagnoses to one or more concepts with patient-friendly terms, by using the SNOMED CT subtype-supertype hierarchy. This increased the coverage of diagnoses by clarifications to 71%. Therefore thousands of diagnoses could potentially be clarified using only a few hundred patient-friendly terms. In Chapter 6 we aimed to assess the medical validity and to identify problems with these diagnosis clarifications. Next to patient-friendly terms of more general, supertype concepts, we also included patient-friendly definitions in the clarifications. We took a representative random sample from the clarifications that covered all the supertype concepts with patient-friendly terms. Two raters with a medical and terminological background identified errors in 13% of the clarifications and considered 14% of the clarifications to be unacceptable to use in a hospital

patient portal. We found that the problems were related to patient-friendly synonyms and definitions, terminology mappings between the Diagnosethesaurus and SNOMED CT, SNOMED CT modelling and the algorithm. Subsequently, we adapted the patient-friendly reference set and the algorithm to address the issues that had been identified in the validation study. We implemented clarifications with synonyms, definitions and generalizations into the Franciscus Gasthuis & Vlietland hospital patient portal problem list. In Chapter 7, we evaluated the coverage, use and quality of the clarifications from the perspective of patient portal users. We found that clarifications were viewed by 55% of the users that viewed their problem list. The clarifications largely were considered to be of good quality, being rated with a median of 6 (interquartile range: 4 – 7) per patient from (1) very bad to (7) very good. See Figure 1 for an overview of the results.

In conclusion, this thesis contributes to the reuse of EHRs by patients and clarifying medical data. Patient access to EHRs is becoming increasingly more common and important, with new regulations being developed to standardize and facilitate data exchange such as the Wegiz (Wet op elektronische gegevensuitwisseling in de zorg" in Dutch, which means law on electronic data exchange in healthcare in English) in the Netherlands and the European Health Data Space in the European Union. Clarifications can be provided to data already routinely registered in EHRs and enable the user to choose to receive a clarification without clinicians having to change the way they record data in EHRs. We recommend terminology developers to apply the generalization method presented in this thesis to reuse patient-friendly terminology across medical concepts. In further research clinical language processing technologies for free-text encoding should be considered to improve lexical clarification performance. Additionally, relationships additional to the hierarchical relationships from SNOMED CT could be utilized to generate clarifications of medical concepts.

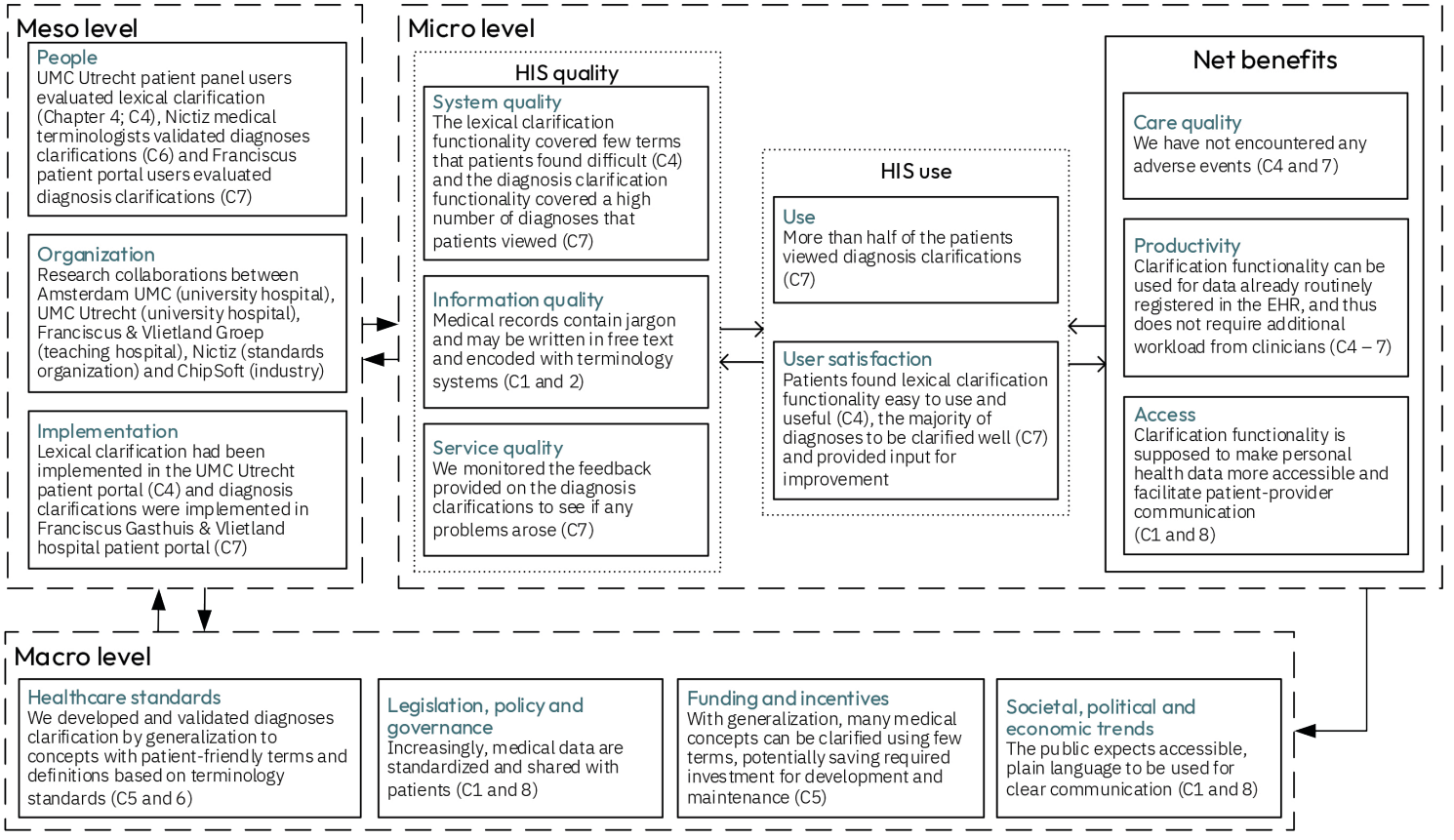


Figure 1 Principal findings from this thesis about clarification of medical data in electronic health records (EHRs) ordered in the Clinical Adoption Framework. HIS: Health Information System.

Nederlandse samenvatting (Dutch summary)

Van jargon naar duidelijkheid: Over toegang tot elektronische patiëntendossiers voor patiënten en patiëntvriendelijke terminologie

Persoonlijke gezondheidsomgevingen (PGO's) en patiëntenportalen stellen patiënten en hun naasten in staat om toegang te krijgen tot hun elektronische patiëntendossier (EPD). Dit zou hen moeten helpen meer te weten te komen over hun medische toestand, zich voor te bereiden op consulten, te onthouden wat er is besproken en voor zichzelf te zorgen. Het is dus belangrijk om de adoptie van patiëntenportalen en PGO's te faciliteren. Verschillende reviews over dit onderwerp behandelden echter verschillende aspecten vanuit verschillende perspectieven, waardoor het moeilijk is om een overzicht te krijgen van de determinanten en uitkomsten van de toegang van patiënten tot EPD's, patiëntenportalen en PGO's. In het bijzonder zijn medische gegevens moeilijk te begrijpen. Vooral personen met beperkte gezondheidsvaardigheden ervaren belemmeringen om toegang te krijgen tot hun EPD's en om medische gegevens en informatie te gebruiken ten behoeve van hun gezondheid. De meeste gegevens in EPD's zijn verborgen in vrije tekst, maar steeds meer gegevens worden gecodeerd met gestandaardiseerde terminologiesystemen, waardoor het makkelijker wordt om gegevens te hergebruiken voor secundaire doeleinden, zoals het verkrijgen van statistieken over morbiditeit en mortaliteit. Consumentgerichte terminologie overbrugt de kloof tussen medische en consumententaal, zoals de Nederlandse Thesaurus Zorg en Welzijn (TZW). Daarnaast bevat de SNOMED CT Nederlandse editie, een uitgebreid gestandaardiseerd medisch terminologiesysteem, patiëntvriendelijke termen en definities. Met deze termen en definities konden echter slechts enkele medische begrippen worden verduidelijkt. Dit proefschrift geeft een overzicht van determinanten en uitkomsten van de toegang van patiënten tot EPD's en stelt toepassingen voor en evalueert deze om medische termen toe te lichten voor patiënten met behulp van terminologiesystemen.

In Hoofdstuk 2 hebben we systematisch systematische reviews beoordeeld over de toegang van patiënten tot EPD's, patiëntenportalen en PGO's, die allemaal soorten zorginformatiesystemen (ZIS'en) zijn. We synthetiseerden de determinanten en uitkomsten van deze reviews in het Clinical Adoption Framework (CAF) en beoordeelden hun risico op bias met de AMSTAR2-checklist. Het CAF is een integratief raamwerk om de adoptie van ZIS'en te evalueren. De reviews hadden volgens de AMSTAR2-checklist verschillende kritieke tekortkomingen en daarom moet bij de interpretatie van de resultaten rekening worden gehouden met kwaliteitsbeperkingen. Reviews rapporteerden over de CAF-dimensies van

mensen, organisatie, implementatie, ZIS-kwaliteit, ZIS-gebruik en voordelen van ZIS-gebruik. Weinig reviews vonden associaties op het macroniveau van zorgstandaarden, wetgeving, beleid, overheid, financiering, prikkels en sociale, politieke en economische trends, wat aangeeft dat er meer onderzoek nodig is op die niveaus. Er was weinig begeleidende documentatie over het toepassen van het CAF en het was niet specifiek ontworpen om patiënten als eindgebruikers van ZIS'en te beschouwen. Daarom hebben we in hoofdstuk 3 gerapporteerd hoe we de CAF hebben uitgewerkt om patiënten op te nemen als gebruikers van ZIS'en en hebben we in- en exclusiecriteria voor de CAF-categorieën gedefinieerd. We vonden dat de interbeoordelaarsbetrouwbaarheid bescheiden was. De begeleidende documentatie kan worden gebruikt bij verder onderzoek naar de toegang van patiënten tot EPD's.

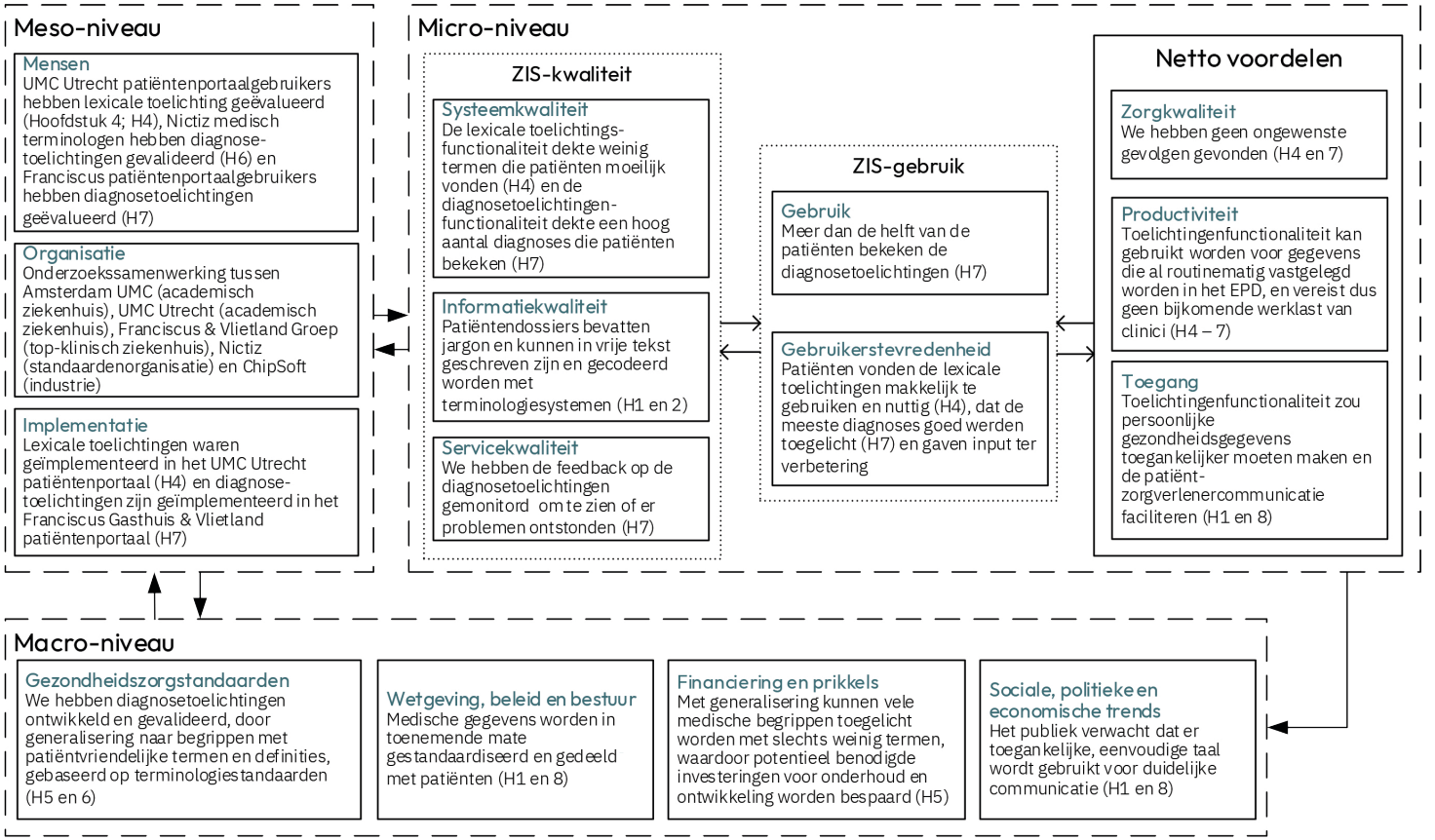
In het patiëntenportaal van het UMC Utrecht is de lexicale toelichtingsfunctionaliteit geïmplementeerd om medische termen in vrije tekst klinische aantekeningen toe te lichten. De toelichtingen bestonden uit synoniemen en definities uit de TZW. In hoofdstuk 4 hebben we deze functionaliteit geëvalueerd om te beoordelen of deze voorziet in de informatiebehoeftes van de patiënten en hoe de functionaliteit verbeterd kon worden. We includeerden vijftien patiënten. Ze hadden voldoende gezondheidsvaardigheden. De deelnemers markeerden en beoordeelden het gemak van de termen die ze moeilijk vonden in een selectie van klinische aantekeningen uit hun EPD's. Vervolgens werden de klinische aantekeningen opnieuw gelezen, maar dan met de toelichtingen die de functionaliteit bood. De deelnemers beoordeelden de termen en toelichtingen op de makkelijkheid en de toelichtingen op het nut. We stelden vast dat de functionaliteit een mediane precisie had van 6,5% (interkwartielafstand: 2,3 – 14,3%) en een sensitiviteit van 8,3% (interkwartielafstand: 4,7 – 13,5%) per patiënt. Dit betekent dat de functionaliteit slechts enkele termen toelichtte die de patiënten moeilijk vonden en veel andere termen die ze niet als moeilijk beschouwden. Ondanks deze lage precisie en sensitiviteit vonden de patiënten de functionaliteit gebruiksvriendelijk en nuttig.

In Nederland worden diagnoses geregistreerd met het coderingssysteem Diagnosethesaurus. Omdat de Diagnosethesaurus is gekoppeld aan SNOMED CT, kunnen de patiëntvriendelijke termen van SNOMED CT worden gebruikt om diagnoses voor patiënten toe te lichten. De dekking van diagnoses met patiëntvriendelijke termen was echter slechts 1,2%. In Hoofdstuk 5 stelden we voor om diagnoses te generaliseren naar een of meer begrippen met patiëntvriendelijke termen, door gebruik te maken van de SNOMED CT

subtype-supertype hiërarchie. Dit verhoogde de dekking van diagnoses met toelichtingen tot 71%. Er zouden dus in potentie duizenden diagnoses kunnen worden toegelicht met slechts een paar honderd patiëntvriendelijke termen. In hoofdstuk 6 wilden we de medische validiteit beoordelen en problemen met deze diagnosetoelichtingen identificeren. Naast patiëntvriendelijke termen van meer algemene, supertypebegrippen, hebben we in de toelichtingen ook patiëntvriendelijke definities opgenomen. We hebben een representatieve steekproef getrokken uit de toelichtingen die alle supertype-concepten met patiëntvriendelijke termen omvat. Twee beoordelaars met een medische en terminologische achtergrond identificeerden fouten in 13% van de toelichtingen en vonden 14% van de toelichtingen onaanvaardbaar voor gebruik in een patiëntenportaal van een ziekenhuis. We ontdekten dat de problemen verband hielden met patiëntvriendelijke synoniemen en definities, koppelingen tussen de Diagnosethesaurus en SNOMED CT, SNOMED CT-modellering en het algoritme. Vervolgens hebben we de patiëntvriendelijke referentieset en het algoritme aangepast om de problemen aan te pakken die in de validatiestudie waren geïdentificeerd. We hebben toelichtingen met synoniemen, definities en generalisaties geïmplementeerd in de probleemlijst Franciscus Gasthuis & Vlietland ziekenhuispatiëntenportaal. In hoofdstuk 7 evalueerden we de dekking, het gebruik en de kwaliteit van de toelichtingen vanuit het perspectief van patiëntenportaalgebruikers. We ontdekten dat toelichtingen werden bekeken door 55% van de gebruikers die hun probleemlijst bekeken. De toelichtingen werden grotendeels als van goede kwaliteit beschouwd en werden beoordeeld met een mediaan van 6 (interkwartielafstand: 4 – 7) per patiënt van (1) zeer slecht tot (7) zeer goed. Zie Figuur 1 voor een overzicht van de resultaten.

Concluderend draagt dit proefschrift bij aan het hergebruik van EPD's door patiënten en het toelichten van medische gegevens. Toegang van patiënten tot EPD's wordt steeds gangbaarder en belangrijker, met nieuwe regelgeving die wordt ontwikkeld om gegevensuitwisseling te standaardiseren en te faciliteren, zoals de Wegiz (Wet op elektronische gegevensuitwisseling in de zorg) in Nederland en de European Health Data Space in de Europese Unie. Er kunnen toelichtingen worden gegeven aan gegevens die al routinematig in EPD's worden geregistreerd, zodat de gebruiker zelf ervoor kan kiezen om een toelichting te ontvangen zonder dat clinici de manier hoeven te veranderen waarop ze gegevens in EPD's vastleggen. We raden terminologieontwikkelaars aan om de in dit proefschrift gepresenteerde generalisatiemethode toe te passen om patiëntvriendelijke terminologie te hergebruiken voor medische begrippen. In vervolgonderzoek zouden klinische taalverwerkingstechnologieën (natural

language processing) voor de codering van vrije tekst moeten worden overwogen om de lexicale toelichting te verbeteren. Daarnaast kunnen relaties aanvullend op de hiërarchische relaties van SNOMED CT worden gebruikt om medische begrippen toe te lichten.



Figuur 1 Voornaamste bevindingen uit dit proefschrift over toelichtingen van medische gegevens uit elektronische patiëntendossiers (EPD's) geordend in het Clinical Adoption Framework. ZIS: Zorginformatiesysteem

A

Patient access to electronic health records (EHRs) is facilitated by patient portals and personal health record systems. However, medical data are difficult to comprehend. In this thesis, we systematically reviewed systematic reviews on patient access to records and elaborated upon an adoption framework to include patients as users of EHRs. Furthermore, we showed that a tool to clarify medical terms in free-text EHR content to patients was appreciated by patients even when the coverage and quality of the clarifications was limited. Additionally, we developed and validated a novel approach to generate clarifications of medical diagnoses that increases the coverage of diagnoses by clarifications significantly. Finally, we implemented the clarifications into a hospital patient portal problem list and found that the clarifications were actually used and that most of the clarifications were considered to be of good quality by actual patient portal users. The presented work contributes to the (re)use of EHRs by patients and clarifying medical data to laymen.