TALLINN UNIVERSITY OF TECHNOLOGY
School of Information Technologies

Marju Tomson YVEM192584

COMPARISON OF HEALTH APPS QUALITY REQUIREMENTS EVALUATION IN ESTONIAN TALTECH HEALTH APP EVALUATOR WITH THE CEN ISO/TS 82304-2:2021 "HEALTH AND WELLNESS APPS - QUALITY AND RELIABILITY"

Master's thesis

Supervisor: Priit Kruus
MSc

Co-supervisor: Carmen Mäe
MSc

Tallinn 2022

TALLINNA TEHNIKAÜLIKOOL
Infotehnoloogia teaduskond

Marju Tomson YVEM192584

TALTECH'I TERVISERAKENDUSTE HINDAJA JA CEN ISO/TS 82304-2:2021 "TERVISE- JA HEAOLURAKENDUSED – KVALITEET JA USALDUSVÄÄRSUS" KVALITEEDINÕUETE HINDAMISE VÕRDLUS

Magistritöö

Juhendaja: Priit Kruus
MSc

Co-supervisor: Carmen Mäe
MSc

Tallinn 2022

Author’s declaration of originality

I hereby certify that I am the sole author of this thesis. All the used materials, references to the literature and the work of others have been referred to. This thesis has not been presented for examination anywhere else.

Author: Marju Tomson

09.05.2022

Abstract

Background: In recent years a variety international regulatory initiative on health apps evaluation have evolved and new frameworks and standards have been developed. Evaluation frameworks are expected to provide assurance on health apps quality and reliability, and build trust, which is essential for the further adoption of new solutions that have potential to empower citizens in their health and wellness care, and thereby support healthcare systems sustainability.

Aim: The aim of the thesis is to compare health apps evaluation quality requirements in the Estonian TalTech Health App Evaluator with the CEN ISO/TS 82304-2:2021 "Health and wellness apps - Quality and reliability" technical specification, to understand similarities, differences and analyse what these may entail for the TalTech and policymakers.

Methodology: Case study research methodology and cross-case analysis and synthesis methods are used for investigating quality requirements in three focus areas - health apps evaluation principles, evaluation criteria, and evaluation outcomes.

Results: Although both health apps evaluation frameworks have similar structures and purposes, the study highlights differences in approaches towards quality requirements evaluation and provides recommendations for change.

Conclusion: The results of this study provide input for international harmonization of the TalTech Health App Evaluator and development of the Estonian digital health technologies evaluation framework.

This thesis is written in English and is 54 pages long, including 5 chapters, 3 figures and 17 tables.

Annotatsioon

TalTech’i terviserakenduste hindaja ja CEN ISO/TS 82304-2:2021 hindamispõhimõtete võrdlus

Taust: Viimased aastad on kaasa toonud palju uusi terviserakenduste hindamist käsitlevaid rahvusvahelisi arenguid ja regulatiivseid algatusi, ning välja on töötatud uusi hindamisraamistikke ja standardeid erinevate lähenemistega hindamise põhimõtetele, kriteeriumidele ja tulemustele. Rakenduste hindamine tõstab terviserakenduste kvaliteeti ja usaldusväärsust, suurendab usaldust uute innovaatiliste terviserakenduste edasiseks kasutuselevõtuks, loob kasutajatele suuremaid võimalusi oma tervise- ja heaolu eest hoolitsemiseks ning toetab seeläbi tervishoiusüsteemide jätkusuutlikkust.

Eesmärk: Lõputöö eesmärk on võrrelda Eesti TalTech’i terviserakenduste hindaja ja CEN ISO/TS 82304-2:2021 “Tervise- ja heaolurakendused – kvaliteet ja usaldusväärsus” terviserakenduste hindamise kvaliteedinõudeid, et analüüsida neis olevaid sarnasusi ja erinevusi.

Metoodika: Juhtumianalüüsi *(case study)* uurimise metoodikat ning ristjuhtumite analüüsi ja sünteesi *(cross-case analysis and synthesis)* meetodeid kasutatakse kvaliteedinõuete uurimiseks kolmes fookusvaldkonnas – terviserakenduste hindamispõhimõtted, hindamiskriteeriumid ja hindamistulemused.

Tulemused: Kuigi mõlemal terviserakenduste hindamisraamistikul on sarnane struktuur ja eesmärgid, toob uuring esile erinevused kvaliteedinõuete hindamise lähenemisviisides ja annab soovitusi muudatuste tegemiseks.

Kokkuvõte: Käesoleva uuringu tulemused annavad sisendi TalTech’i terviserakenduse hindaja rahvusvaheliseks ühtlustamiseks ja Eesti digitaalse tervisetehnoloogiate hindamisraamistiku väljatöötamiseks.

Lõputöö on kirjutatud inglise keeles ning sisaldab teksti 54 leheküljel, 5 peatükki, 3 joonist, 17 tabelit.

List of abbreviations and terms

CE	Compliance with European community regulatory requirements
DHT	Digital health technologies
DiGA	Digital health application
DiGAV	Digital Health Applications Regulation
DTAC	Digital Technology Assessment Criteria
DVG	Digital Health Care Act
EC	European Commission
EHIF	Estonian Health Insurance Fund
EU	European Union
EUnetHTA	European Network for Health Technology Assessment
HTA	Health technology assessment
GDPR	General Data Protection Regulation
ISO	International Organization for Standardization
MDR	Medical Device Regulation
mHealth Hub	European Innovation and Knowledge mHealth Hub
NHS	The National Health Service
NICE	The National Institute for Health and Care Excellence
PII	Personally identifiable information
REA	Rapid relative effectiveness assessment
TalTech	Tallinn University of Technology
UK	United Kingdom
WHO	World Health Organization

Table of contents

List of figures

List of tables

1 Introduction

Digital health is defined as “the use of digital, mobile and wireless technologies to support the achievement of health objectives [1]” by the World Health Organization (WHO). The decisions to combine digital health and current care practices are derived from the need for efficient and accessible healthcare, and an economically sustainable healthcare system [2]. ‘Digital health provides possibility for patient-centred, convenient, engaging personal care that is quickly accessible and 24 hours a day available before, during and after treatment [3]. Digital health technologies create possibilities for patients’ empowerment, reduce the number of required healthcare appointments, and enable cost-efficient scalability of solutions [4]. Mobile health technologies have considerable potential to positively impact the sustainability of health systems by creating new, more cost-effective healthcare access, communication, and delivery pathways [5].

An increasing number of mobile health apps are helping to improve the availability and affordability of healthcare, but there have been only a limited number of controls, regulatory procedures, standards, and guidelines supporting the growth, limiting risks, and ensuring quality [6]. With health apps it could be possible to improve the quality of care and achieve cost-savings, but there are several risks derived from the low quality and safety of many of these apps [7]. The complexity of choosing between good- and low-quality health apps has steadily increased, with greater risks for both patients and healthcare professionals worldwide [8].

To ensure effective and accessible health care, policy makers need clarity on how health innovations that are increasingly integrated into health systems are assessed before they are implemented [2]. Over the last years new digital health solutions evaluation initiatives to support and guide different stakeholders have emerged worldwide [9]. Health and wellness apps evaluation and standardization can benefit users, healthcare organizations and policy makers [10]. Digital health solutions evaluation that is objective, transparent and based on standards adds clarity and confidence to all

stakeholders starting from patients and healthcare professionals to developers and policy makers [11]. The use of a standardized evaluation approach increases comparability and availability of assured high-quality apps to choose from for patients, harmonizes international market access for manufacturers, and in longer term supports sustainability of healthcare systems [12].

1.1 Aim and research question

Problem

The increase in in the number of health apps [13], [14] has brought along the risks for potential harm on users safety, and complexities of differentiating between good- and low-quality solutions [7], [8]. This thesis is focusing on the need for the international comparison and harmonization of the Tallinn University of Technology (TalTech) Health App Evaluator to support the development of an Estonian digital health technologies evaluation framework that can provide assurance on health apps quality and reliability for stakeholders and build trust, which is essential for the further adoption of new solutions that have potential to empower citizens in their health and wellness care, and thereby support healthcare systems sustainability [15], [12], [5]. The study may have in a longer time horizon an impact on wider public and healthcare decision-making while recommendations are adopted by the policy makers.

Aim

The aim on the thesis is to compare the Estonian TalTech Health App Evaluator [16] and the CEN ISO/TS 82304-2:2021 “Health and wellness apps - Quality and reliability” [17] health apps evaluation quality requirements to understand similarities, differences and analyse what these may entail for the TalTech and policymakers by using case study and cross-case analysis and synthesis methods.

Research question

How the TalTech Health App Evaluator quality requirements could be enriched based on the comparison of similarities and differences with an international health apps evaluation harmonization initiative, the CEN ISO/TS 82304-2:2021 “Health and wellness apps - Quality and reliability” in three focus areas: 1) evaluation principles, 2) evaluation criteria, 3) evaluation outcome.

1.2 Background

1.2.1 Evaluation in theory

In this chapter the evaluation theory background overview is introduced. According to the benchmark evaluation guide, “Evaluation Theory, Models, and Applications” by Daniel L. Stufflebeam’s and Chris L. S. Coryn’s [18], evaluation has been defined over the years as an assessment of specific objectives accomplishment, testing based on set references or carrying out controlled experiments that generate quality information for decision-making [18, p. 6]. However, evaluation that is based on objectives that are inappropriate, irrelevant, or not oriented towards the needs of planned stakeholders, can be limiting and not beneficial for consumer products and services assessment [18, p. 7]. In addition to objectives, processes are also relevant and should be considered for successful evaluation outcomes [18, p. 7]. All the suitable, relevant, necessary, and useful methods are recommended to be used for carrying out evaluation to reach valuable results [18, p. 8]. Besides outcome achievement assessment, evaluation can also be used for systematically and objectively determining value, efficiency, effectiveness, impact, and sustainability; and for gathering credible and useful information for decision-making [19]. Evaluation credibility is established by independent, autonomous assessment that is carried out without the influence of related counterparties [19].

Evaluation should be based on a defensible and relevant set of reference values like safety, usability, costs, legality [18, p. 8]. Safety issues are the core focus of many evaluations and apply to all fields, products, and services to minimize possible risks posed on consumers [18, p. 9]. The assessed value can be defined as merit while representing internal values and quality; and as worth or external value related to context, needs, and costs [18, p. 8]. Merit expresses excellence at object level showing how well it performs in respect of established respective standards or compared with similar objects [18, p. 9]. However, for being worthy it is necessary to demonstrate besides required level of quality as well relevance in given context related to target group needs and considering both involved monetary and nonmonetary costs [18, p. 9].

Criterion-referenced assessment against published standards provides the clearest findings, while absence of published standards may lead to results that lack

transparency, reliability, or validity [18, p. 19]. Professional standards agreed by experts and reflecting stakeholders input establish sound principles and common criteria for conducting credible, accountable, valuable, and fair evaluations [18, p. 69]. Evaluator has a key role in ensuring evaluation's impact and relevance by effectively communicating with stakeholders as well by supporting the use of findings [18, p. 9]. The evaluation results and feedback should be reported and presented timely, effectively, and accurately to relevant audiences [18, p. 15].

Formative evaluations are used for providing information and feedback during the development phase to ensure and improve quality and are carried out prospectively and proactively [18, p. 21]. Summative evaluations establish retrospectively accountability on already finished products and services and inform consumers regarding the quality and safety, also other related counterparties can obtain information about the assessed phenomena [18, p. 22]. Internal formative and summative self-evaluations conducted by organisations provide continuous support for analysing and improving production and supplying data for external independent evaluations needed for establishing public accountability [18, p. 27]. Variety of accrediting organisations perform regular accreditation evaluations that are based on clearly established evaluation criteria and self-assessment guidelines and make the results publicly available [18, p. 27].

Formal evaluations are distinguished from informal quick and intuitive judgements by used methodologies that involve "complex areas of epistemology, rules of evidence, information sciences, research design, measurement, statistics, communication, and some others [18, p. 29]". An evaluation theory consists of a "coherent set of conceptual, hypothetical, pragmatic, and ethical principles" that create a framework for guiding the study and practice [18, p. 50]. Like theory informs practice, the feedback from practice has a role to validate and strengthen theory [18, p. 47]. However, using only a narrow evaluation theory that provides a systematic framework for assessing objectives and intended outcome achievement has led to disregarding context and process for decades [18, p. 47].

The evaluation theory has been advanced and dialogue on its meaning and use in real-world settings has been supplemented by creative theorists like Robert Stake by recommending using predetermined responsive evaluations; Michael Scriven suggesting use of a wider goal-free approach; Lee Cronbach advising to focus on generalizable

contingency-based assessment; and Egon Guba advocating use of experimental design instead of a naturalistic approach [18, p. 48]. However, the evaluating theory tends not to be followed by many practising evaluators possibly due to insufficient reasoning for practical use, practitioners' deficient training and proficiency, or missing convincing evidence on properly applied theory yielding sound evaluation outcomes [18, p. 49]. Instead, pragmatic principles that have grown out from vast experiences and have proven to work well have been directing ways for evaluation practice [18, p. 51]. For further advancement of effective evaluation practice thorough and validated evaluation theories are needed [18, p. 64].

1.2.2 Evaluation in practice

In the following chapter the theoretical evaluation concept is supplemented with examples from selected practical digital health technologies quality evaluation approaches, and international developments in health applications quality requirements are introduced.

The increased interest towards mobile health apps has created a need for evaluation that increases trust of users and improves quality of apps [20]. Quality has been regarded from different perspectives as "excellence, value, conformance to specifications, meeting or exceeding expectations" [21]. Quality evaluation is an activity that focuses on assessment and improvement of developed solutions value or compliance with standards and is related to proper use of evaluation criteria, followed by objective analysis and presentation of results [19]. Evaluation reliability is achieved by using a consistent and dependable approach for collecting and interpreting evaluation data that yields repeatable results under similar conditions [19].

Digital health technologies (DHT) are defined as "apps, programmes and software used in the health and care system. They may be standalone or combined with other products such as medical devices or diagnostic tests [4]". Mobile health apps used for monitoring and prevention are one of the categories of digital health together with electronic health, telehealth, and health data solutions [22]. The increase in the use of apps in the health and wellness field, that supports lifestyle and nutrition choices, is associated with the search for a more holistic view towards health [23]. These apps are sold directly to users through app stores, often without any official evaluation and supporting evidence [17]. An international consensus is reached on consistency in

terminology to use the term “app” in app-related scientific articles while referring to mobile medical applications [24].

Some of the apps are also classified as **medical devices** [17]. For medical devices and related software separate requirements are applicable that are outlined in the European Union (EU) Medical Device Regulation (MDR) [25]. Medical devices are defined as “any instrument, apparatus, appliance, software, implant, reagent, material or other article intended by the manufacturer to be used, alone or in combination, for human beings for one or more of the following specific medical purposes: 1. diagnosis, prevention, monitoring, prediction, prognosis, treatment or alleviation of disease, 2. diagnosis, monitoring, treatment, alleviation of, or compensation for, an injury or disability, 3. investigation, replacement or modification of the anatomy or of a physiological or pathological process or state [25]”. Medical device evaluation covers clinical safety assessment [25].

The global mobile healthcare **market** is expected to grow from USD 23 billion in 2016 to USD 190 billion in 2025, increasing the relevance of accessibility, transparency, privacy, and security of healthcare-related services, while ensuring efficiency and cost savings [14]. While in 2008 the Apple App Store was launched with 500 apps altogether [26], in 2021 overall 350 000 consumer health apps were available globally, and more than 250 new healthcare apps were released on average per day [13]. Freely available health apps are used by 75% of smartphone users, the most common are apps for home exercising, body and fitness data like heart rate and steps recording, information providing, motivating, and advising, as well for reminding medicines taking [27]. At the same time the number of digital therapeutics and care products, that also can be health apps, and which are recognized as medical devices used for treating, preventing, and managing of health conditions, and may be prescribed and reimbursed, was only 250 in 2021 [13]. By a study from 2019, 80% of health apps were not medical devices [28].

The growth in use of mobile technologies has also led to an increase in the different **risk** types that are related to the health app use and functionalities, the context of use, and target users’ profiles [20]. The development of health apps without support of usable and effective standards has carried along heightened risks like “stress, dissatisfaction, delay in effective treatment, loss of privacy, poor lifestyle choices and deterioration in health” for users and possible negative impacts on reputation, care quality and demand,

and lost opportunities to developers and healthcare professionals [7]. While the risk factors that are inherent to an app like intended function, content accuracy, complexity or feedback mechanisms can be lowered by an appropriate regulation; the external contextual risk factors that are related to app proper use and users may be targeted by user groups awareness raising [29]. An app may perform either one or several functions like informing, calculating, instructing, communicating, and the risks related to the use of the app must be communicated to the users via appropriate indicators and in relevant channels [20]. Reliability and accuracy of information provided by mobile health apps that are used in the healthcare area is of utmost importance as it has further impact on health-related decisions [29].

An evaluation framework of digital health services benefits different **stakeholders** by enabling informed evidence-based decision-making, provides information to healthcare professionals about available tools and benefits, and to manufacturers about the app's approval requirements [30]. The new developed solutions will benefit service users - either people, their caretakers or wider healthcare system [4]. Evaluation serves different purposes for stakeholders and each of these groups has its own sector-specific objectives that are reflected in different evaluation domains and processes, which complicates harmonization of assessment systems [20]. The evaluation framework can create different values for stakeholders – enable better access and easier decision-making for patients and citizens; support healthcare professionals on clinical decision-making and patient empowerment; clarify benefits and relative value for informing decisions related to policies, funding, and reimbursement; and provide good practice guidelines for developers of health apps [9].

The **evidence** on mobile health efficacy and effectiveness has been limited despite global growth in mobile phone use and resulting popularity of mobile health interventions with potential for efficient high-quality care delivery [31]. The number of academic studies on health apps' clinical impact has also been limited, and behaviour change techniques have not been considered neither in apps' development nor in evaluation methodologies so far [8]. However, the amount of scientific literature on mobile health utility has increased recently [20], similarly the clinical evidence on digital health apps effectiveness has been growing and maturing, and the amount of published studies has increased altogether by 2000 since 2007, with three-quarters of them published over the last five years [13].

Health technologies assessment

Due to diverse needs of stakeholders and variable characteristics of technologies, different evaluation approaches and criteria have been developed and used for the health technologies assessment [32]. Health Technology Assessment (HTA) has conventionally been used for evaluating pharmaceutical technologies' influence on health and care [2]. HTA has been in use already since the 1980s to support the development of safe and effective patient-focused health policies by compiling the medical, social, economic, and ethical aspects of the use of health technologies [33]. HTA has been a good source of evidence on safety, efficacy, and cost-effectiveness of interventions [34]. The aim of the HTA is to support new health technologies usage by providing information to decision makers [30].

The European network for HTA (EUnetHTA), supported by the European Commission, has since 2006 developed a full HTA model which informs decisions on technology and can be used in local and international contexts [35]. The full HTA model was created for the evaluation of "pharmaceuticals, medical and surgical interventions, diagnostic technologies, and screening [35]". EUnetHTA HTA Core Model® for comprehensive full HTA consists of two appraisal modules - a rapid relative effectiveness assessment (REA) that covers four domains: "(1) the health problem and current use of technology; (2) description and technical characteristics of the new technology; (3) safety assessment; (4) clinical effectiveness [36]"; and a national appraisal module that adds five more categories of "(5) economic evaluation, typically cost-effectiveness or cost-utility analysis; (6) ethical analysis; (7) organizational aspects; (8) patient and social aspects; (9) legal aspects [36]". From nine domains four have focus on clinical and five on non-clinical topics [37]. The compiled evidence and information on health technologies can be applied across organizations and internationally; and shared and reused via repository, thus avoiding duplication of work [35]. HTA focus is on determining and comparing added value and relative effectiveness of new or already available health technology based on a scientific evidence-based process carried out by designated authorities to promote these innovations that have the best result for people and society [37]. HTA results are used for supporting budgetary, pricing or reimbursement related decisions [37].

However, carrying out parallel assessments in multiple EU member states creates a burden of duplicate processes and varying outcomes resulting from local diversified

requirements [37]. In December 2021 the EU HTA Regulation was adopted by the European Parliament and the Council governing the clinical evaluation of health technologies in order to harmonize evaluation principles, ensure uniform evaluation across the EU and lower administrative burden for developers of medical devices and medicinal products, starting from the devices belonging into the highest risk classes by MDR [37], [25].

HTA frameworks have been developed for the evaluation of pharmaceuticals, medical devices, and services; their suitability and sufficiency for DHTs purpose is questionable due to different benefit and risk profiles [38]. Even of the same type DHTs need to be assessed due to their inherent variabilities instead of intervention levels on product levels [30]. An extensive systematic review on digital technologies evaluation frameworks identified and developed a list of DHTs specific topics that are relevant to be considered additionally when conducting an HTA, resulting in six of EUNetHTA Core model domains to be complemented by eight additional topics mostly in the fields of safety and clinical effectiveness [38]. The need for more profound technology-specific questions was identified in all nine EUNetHTA Core Model domains implying that the current HTA approach is insufficient for evaluating DHTs [38].

On the one hand, the existing evaluation frameworks only partially cover the domains of HTA, but also the classic HTA model also needs to be complemented by additional technology-specific aspects related to data privacy and protection, connectivity, compatibility, and software updates for enabling a thorough evaluation of digital health applications [39]. Several evaluation domains like data security, data protection and accessibility, that are important for digital tools, are not included into the traditional HTA assessment [30]. Absence of critically relevant DHT-specific content in HTA research in the areas of technical reliability, stability, cybersafety and cybersecurity, as well as in patient satisfaction has been identified [40]. Over the last years the amount of DHT-specific evaluation frameworks and guidance has grown and regulations like EU Medical Device Regulation (MDR) [25] and the EU General Data Protection Regulation (GDPR) that establishes rules governing “the protection of natural persons with regard to the processing of personal data and rules relating to the free movement of personal data [41]”, have improved clarity of what is required and relevant for HTA [38].

National approaches to DHT evaluation

There is a large heterogeneity and diversity of approaches towards health apps evaluation principles, criteria, and outcomes in Europe [9]. The focus on the existing European frameworks is primarily on supporting end-users - people and healthcare professionals by providing confidence and trust for smoother apps adoption [9]. The assessment approach involves different combinations of self-assessment, owner, and expert assessments [9]. Majority of the current national and regional frameworks operate on a voluntary basis by providing secondary benefits, while a few mandatory ones have higher potential for supporting integration into the healthcare systems [9]. However, the scalability may be limited on frameworks that use for assessment and publication of results in a repository only in the local language [9].

Evaluation criteria can support a variety of stakeholders – developers, patients, healthcare professionals and others starting from informing health apps development to quality measurement [6]. Quality evaluation of health apps is a complicated undertaking, a remarkable variety of assessment criteria is present in different methodologies, and several relevant aspects are not included in popular methodologies [8]. The large diversity in health apps evaluation criteria is a consequence of variable assessment approaches, numerous definitions, and lack of consensus regarding related concepts [32]. Also, not all the higher risk class digital health technologies evaluation criteria are applicable to solutions belonging to lower risk classes [38]. Accurate, decisive, and reliable evaluation criteria are needed to support the current health apps compliance assessment with existing best practices and regulations, however due to the ever-changing nature of these technologies the criteria will not be perfect and complete, also it should not add further complications to an already complex apps marketplace [32].

The evaluation outcomes have been diversified with a variety of additional approaches like scorings, labelling, certification, and online repositories over the last five years [20]. The presentation of the final evaluation outcome varies across current frameworks, including quantitative final scores and qualitative visualized quality marks or labels, and profound end-results reports [9]. Frameworks serve as well as guidelines for developers and for other stakeholders' assessments [9]. The creation of new regulations and reimbursement solutions for approved apps has been growing worldwide, increasing the health apps relevance on impacting personal health [13].

Historically the first leading European mobile health adopters with the most favourable market conditions have been the United Kingdom (UK), Germany, and Scandinavian countries [42]. The following paragraphs introduce initiatives from the UK, Germany, Belgium, and Finland, that represent the most advanced European active frameworks on digital health technologies evaluation. These country-specific frameworks have been created mostly as governmental initiatives, representing non-standardized heterogenous approaches towards digital health technologies evaluations.

In **the UK** since 2021 the **Digital Technology Assessment Criteria (DTAC)** for health and social care by the National Health Service (NHS) Digital has been in use for supporting digital health technologies baseline assessment and procurement by healthcare organizations before these new solutions are accepted into the into the NHS and social care systems [43]. The DTAC evaluation establishes requirements for developers of digital technologies facing patients and care providers, apps, systems, web portals etc; and provides assurance to users - people and healthcare professionals in "clinical safety, data protection, technical security, interoperability and usability and accessibility" domains [43]. The assessment is voluntary, during which questions are initially answered by developers and approved by the NHS experts and approved assessors [9]. The results of the assessment are published in the NHS Apps Library reaching thereby to a larger community of users, giving developed solutions wider exposure, and supporting uptake by providing trust and assurance on clinical safety and security [43]. In the UK also a **digital health technologies evidence standards framework** was developed by the National Institute for Health and Care Excellence (**NICE**) and The NHS England in 2018-2019, and reviewed in 2021 (NICE, 2021). The NICE framework focuses on digital technologies by their functions and risks and describes corresponding levels of evidence needed to prove their effectiveness and economic value for the UK health and care system, being less relevant to DHT's like apps that are sold directly to users (NICE, 2021).

In **Germany**, the Digital Health Care Act (DVG) entered into force in 2019, and the Digital Health Applications Regulation (DiGAV) in 2020, which establish digital health applications (**DiGA**) evaluation pathway, enable prescription, and set the requirements for reimbursement by health insurance companies [44]. The DiGA is a CE-compliant medical device from class I or IIa according to the MDR [25] that has successfully passed the evaluation; it is an approved reimbursable digital health application, that is

used either by the patient alone or together with a healthcare professional; and which is listed in the DiGA directory [45]; creating transparency, trust, and informing decisions [44]. The DiGA evaluation process entails 122-item criteria about the app quality, functionality, security, data protection, interoperability, and user friendliness, as well provision of clinical evidence latest within 12-24 months demonstrating positive care effects [44].

In 2018 the **mHealth Belgium** platform was created for assessment of mobile apps that are classified as medical devices [46]. Since 2021 a new three-level pyramid-shaped process is in place that at the first base level requires from a manufacturer proof on device's CE-marking and the GDPR compliance; on the next level criteria on interoperability and connectivity to eHealth platform's basic services are added; followed on the third level by the demonstration of added social-economic value and financing after approval of funding request [47]. The approved devices are added to the mandatory health insurance package, and insurance companies reimburse the device to the end user; information about the apps and their level is published on the mHealth Belgium library webpage [46]. This process is optional for getting the quality label and gaining visibility on the portal, but mandatory for being included into the reimbursement by national authorities [9].

In **Finland** in 2019 a framework for **digi-HTA** process was created covering the fields of mHealth, AI and robotics, and targeting HTA experts and decision-makers by supporting the assessment of the product or service's suitability for healthcare use and introduction into the Finnish healthcare system [30]. For developers digi-HTA is a free tool that can be used for self-assessment and product development [48]. Traditional HTA domains covering security, effectiveness, usability, accessibility, economical, organizational, and technical aspects were supplemented by interoperability, artificial intelligence, and robotics; only ethical, social, and legal topics were left out from the framework to create a fast evaluation possibility for quickly changing health technologies [30]. Data security and data protection domains are assessed separately by outsourced data protection experts [30]. Certified applications are published in a national repository webpage [49].

Standardization

As was highlighted in previous paragraphs, a large variety and great heterogeneity is present in current approaches towards digital health technologies evaluation. The frameworks have different methods, specificity, and outcomes; and standardization could benefit the comparability and quality across the assessments [2]. While considering the implementation of a new assessment framework, public and private entities would benefit from the experiences and lessons learned from existing approaches that could add Europe-wide scalability [9].

Standards can be considered as a compilation of the best methods by experts in the respective fields towards a variety of actions like development, management, delivery or supply of processes, products, services, or materials [50]. In the EU most of the standards and guidelines are issued by the European Committee for Standardization (CEN) in cooperation with the International Organization for Standardization (ISO) and 34 European countries standardization bodies [42]. Internationally, the ISO takes care of standards harmonization in 165 countries and only in the health and medicine field over 1600 standards have been issued by them [10]. Harmonized European standards have the same implementation obligations as national standards, and these are withdrawing conflicting local standards [51]. It is voluntary to use European harmonized standards and technical specifications, however the standards could be referred to by laws and regulations and can support, or even may be made compulsory by the EU legislation and policies [51].

The ISO standards are established by consensus and are created in cooperation with manufacturers, consumers and regulatory stakeholders, the standardization benefits all of them by improving safety and quality of products and by lowering price and transaction costs [52]. Standards support building of European goods and services internal market and removal of trade barriers; helping to ensure safety, reliability and interoperability of products, services, or processes; and can be used to demonstrate compliance with the EU legislation [52]. Conformity assessment procedures and certification audits are carried out by independent bodies from the ISO, who provide external assurance about meeting established product requirements [53]. However, the ISO compliance can be established as well by adherence to the ISO standards internally by manufacturers self-certification, and both – the ISO compliance and certification are voluntary [54].

Several ISO standards provide guidance for different aspects of medical devices: the ISO 13485:2016 for quality management systems to ensure quality and safety, the ISO 14971:2019 for risks management during product development, manufacture, and use; the IEC 62366-1:2015 for usability engineering, verification, and validation [55]. For medical device software, as well as for medical devices that incorporate software, the IEC 62304:2006/AMD 1:2015 provides requirements for lifecycle processes – development, maintenance, problem resolution, and risk management [56]. The EN/IEC 82304-1:2016 focuses on the safety and security requirements for stand-alone health software-only products, it does not cover software that is intended to be a part of hardware [57]. The last two health software related standards on lifecycle processes and safety are applicable besides medical devices also to health apps [55].

The CEN ISO/TS 82304-2:2021

With special focus on health apps, the CEN ISO/TS 82304-2:2021 "Health and wellness apps - Quality and reliability" technical specification was released by the ISO in 2021, providing quality requirements and quality labels for communicating health apps quality and reliability [17]. The technical specification can be converted to a European standard in 2024, so far conflicting parallel national standards can be kept in force [17]. The document provides requirements on the product's health and societal impacts, ethics, user-friendliness, privacy, security, technical robustness, and interoperability [17]. This new standard may have a direct impact on the wider public, patients, and clinical practice through health apps quality labels [10]. This standard supports the self-certification of innovators and informs the various accreditation bodies that develop assessment processes [58], [59]. The principles that health apps standards create for manufacturers ensure that the solutions developed are trustworthy, reliable, and beneficial to users [60].

1.2.3 Estonia

Estonia is an innovative leader in the eHealth field [61]. Several local initiatives have been taken in relation to health technologies assessment (HTA). The University of Tartu has had a Health Technology Assessment Centre since 2012, which supports the Estonian Health Insurance Fund (EHIF) with cost-effectiveness and budget impact assessments of traditional health technologies [62]. The need to evaluate digital health solutions has increased in Estonia, but a nationwide system and cross-sectoral

agreements have not yet been established [16]. The EHIF's 2020-2023 development plan defines support for the development of innovative services, including the use of evidence-based and efficient digital health solutions in the Estonian health care system [63]. The EHIF has adopted the NICE Evidence standards framework for digital health technologies for the Estonian health system and people, which has been developed in the UK to support the wider use of evidence-based digital health solutions [64], [4]. The framework is relevant for the technologies used by the health and social care systems, but it is less focused on solutions that patients can directly download and use [4]. In Estonia the first digital health apps assessment initiative to support patients and healthcare professionals to make an informed choice on the mobile health and health behaviour apps, the Health App Evaluator, was launched by the TalTech in 2021 in Estonia [65], [16].

In 2021 a procurement “Assessment framework for digital healthcare technologies” was initiated by the EHIF for the development of a nationwide digital health technologies evaluation framework [15]. The framework would consist of a baseline assessment of all patient-centred health apps in the Estonian healthcare system, conducted either as a manufacturer self-assessment or through the TalTech Health App Evaluator to increase overall credibility; and of an additional assessment depending on the final purpose, for example the EHIF reimbursement [15]. The need to compare Estonian TalTech Health App Evaluator quality requirements to international evaluation criteria and principles was emphasized in the project conclusions [15]. This thesis is aiming to fill this gap, focuses on supporting the international harmonization of the TalTech Health App Evaluator and contributes to the development of an Estonian digital health technologies evaluation framework.

2 Methodology

This chapter describes case study research methodology and cross-case analysis and synthesis methods; and introduces the research process on data collection and analysis that are used for conducting the study. Approach explained by Robert K. Yin in his book “Case Study Research” is used as a guide for this case study [66]. Additionally, cross-case analysis and synthesis principles proposed by Matthew B. Miles and A. Michael Huberman are followed [67], [68].

The case study methodology is selected because it is one of the main methodologies used in researching the management of technology and innovation; and it supports studying of the dynamics underlying a complex context; and for narrowing down a broad investigated field into more easily researchable examples. It enables in-depth investigation of selected cases and can serve exploratory, descriptive, and explanatory purposes. Case study methodology supports investigation of a contemporary phenomenon that has a real-life context while the investigator has no control over events and searches answers to “how” and “why” questions [66].

This study provides an international comparison on the health apps evaluation and introduces new perspectives on the phenomenon. The case study approach was chosen as the most suitable methodology for examining the gap in existing knowledge and for investigating health apps evaluation quality requirements in three focus areas: evaluation principles, criteria, and outcomes.

2.1 Data collection

The study is conducted during September 2021 - May 2022. Background research is carried out to identify digital health apps evaluation related initiatives, regulations, and frameworks; to provide theoretical background on digital health apps evaluation evolvement; and to compile information for the case study.

The data collection is based on:

- academic literature, scientific papers, and formal studies; and
- publicly available information from organizations homepages, online reports, websites, official documentation, regulations, standards, news articles and official press releases.

The search is conducted from Google Scholar, Primo, PubMed, Startpage databases. The used search terms are digital health, evaluation framework, health technology assessment, apps evaluation, health app quality criteria, health apps standard. Also references of retrieved scientific articles are hand-searched. The information regarding the TalTech Health App Evaluator is translated to English by the author.

The study is based on published information and no human participants are involved; therefore, no privacy and confidentiality issues are identified, and no ethics approval is required for this study. However, the study may have in a longer time horizon an impact on wider public and healthcare decision-making while recommendations are adopted by the policy makers.

A case study protocol is created before engaging in data collection and analysis for selected cases. The protocol establishes levels of questions investigated of specific cases, patterns of outcomes across cases, of entire study, and of recommendations. The case study protocol serves as a data collection guide and during the process it is continuously adapted. According to Yin, a feedback loop should be used while a relevant discovery occurs, the study should be redesigned accordingly, and if needed data collection protocols should be changed respectively [66].

In this study, the case study protocol themes are shown in Table 1 on page 27 and results are presented in the forthcoming chapters. The questions are established, as proposed by Yin [66], at the levels of individual cases, on patterns of findings across cases, of entire study incorporating context from the background info, and of further recommendations, which are presented in the discussion chapter.

Table 1. Case study protocol themes.

Principles	**Criteria**	**Outcomes**
Purpose	Domains	Quality label
Value	Criteria	Scores
Creator	Questions	Report
Year	Topics	Published principles
Area		Repository
Stakeholders		Guideline
Technology		Optionality

Source: Author.

Cases selection process is based on the most similar case selection strategy with focus on comparable characteristics [69]. The cases selection starts with defining the cases universe, followed by identification of similar and different key variables across the cases, and finally selecting the cases [69]. This establishes a basis for the cases to be investigated and studies to be carried out [66].

The decision about the number of cases reflects the required amount of literal and theoretical replications [66]. Two or three replications are recommended while "theory is straightforward and the issue at hand does not demand an excessive degree of certainty [66]". Multiple case study replication logic differs and contrasts to sampling design logic, which "should not be used, the typical criteria regarding sample size also are irrelevant [66]". This study follows the simplest multiple-case study design, where chosen two cases are expected to provide literal direct replications.

In this study the most similar case selection approach, that emphasizes the most comparable characteristic, is used. Both chosen cases are designed for the same purpose to provide assurance on health apps quality. The first case, the Taltech Health App Evaluator, was created for assessing health and health behaviour apps used in Estonia in March 2021 [16]. The second case, the CEN ISO/TS 82304-2:2021 "Health and wellness apps - Quality and reliability" technical specification, was chosen as the most similar case to the Taltech Health App Evaluator. The CEN ISO/TS 82304-2:2021 is a new worldwide health apps evaluation initiative from July 2021 that has been developed based on already existing standards and frameworks on health apps evaluation [17]. The phenomenon studied is the Estonian TalTech Health App Evaluator conformity with the CEN ISO/TS 82304-2:2021 "Health and wellness apps - Quality and reliability".

Health apps evaluation hadn't had until recently a good, standardized approach. The CEN ISO/TS 82304-2:2021 is chosen from alternative available approaches due to its specialized focus on similar technologies evaluation and international harmonization potential. Traditional HTA approach that has been in use for over 40 years [33], lacks digital technologies focus and does not cover all the assessment criteria relevant to digital health technologies [39], [30]. At the same time also current HTA evaluation domains are covered incompletely in existing frameworks [39]. Scientific approaches are often not informed by quickly evolving real practice as the HTA related materials have not been necessarily published in the scientific literature. Countries' national frameworks have heterogeneous approaches towards evaluation principles, criteria, and outcomes; and their focus is mostly on reimbursement of medical devices [9].

Therefore, the CEN ISO/TS 82304-2:2021 was chosen to be the comparison case for the TalTech Health App Evaluator because of having specialized focus on similar technologies - health apps; being developed based on an international consensus having global harmonization potential; and being created to support health app assessment organizations among the other target groups [17]. In this study the data with focus on evaluation principles, criteria and outcomes on selected cases is collected simultaneously with parallel case studies.

2.2 Data analysis

After collecting documentary information on the selected cases, the TalTech Health App Evaluator, and the CEN ISO/TS 82304-2:2021 according to the created case study protocol, their evaluation principles, criteria, and outcomes are thoroughly analysed by using Excel Spreadsheets.

For evaluation principles and outcomes analysis a regular multiple case study approach is used. Due to the more profound needs related to the evaluation criteria analysis, for this part in the study additionally a cross-case study design mechanics and logic are used. In the cross-case method two interdependent strategies are used - cross-case analysis and variable-oriented analysis [67], [68]. The conducted steps during data collection and analysis are illustrated in the following Figure 1 on page 29.

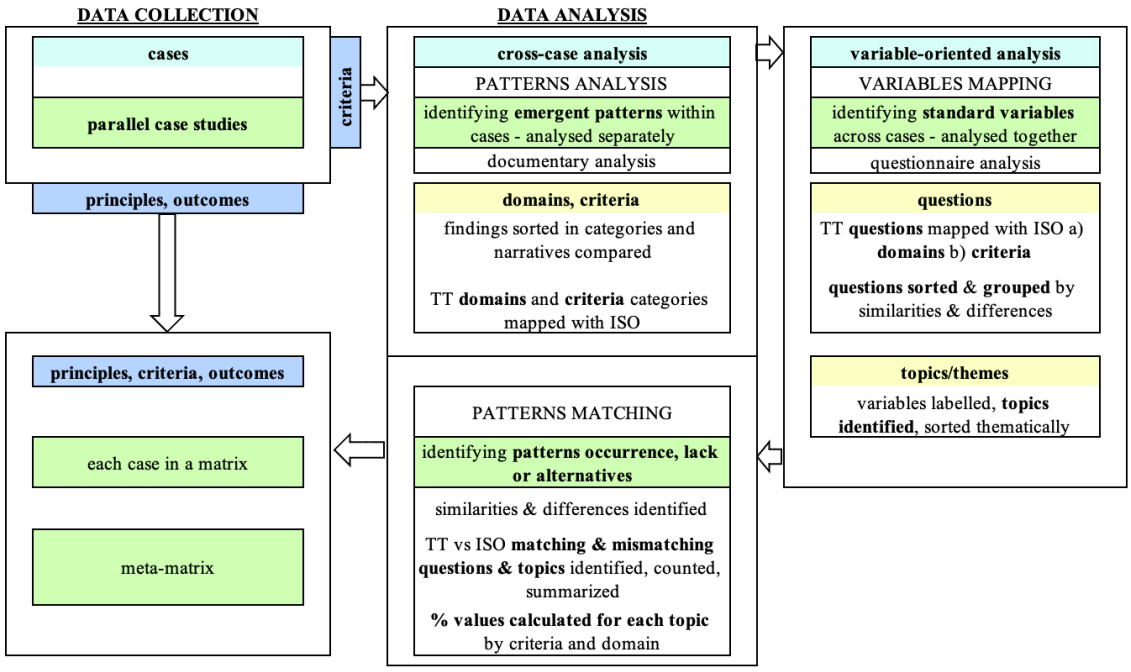

Figure 1. Data collection and analysis during cross-case study.

Source: Compiled by author.

The criteria analysis is based on the health apps evaluation questionnaires that are collected from both frameworks. The gathered data on quality evaluation domains and criteria are transferred to Excel, evaluation questions from questionnaires are sorted thematically, coded, and labelled for both cases. The initial examination helps to structure and analyse the data, and to identify present evaluation principles and criteria for further investigation.

As a part of the cross-case analysis at first pattern analysis is performed and emergent patterns within cases are studied [67], [68]. In this study during cross-case analysis pattern analysis phase the Taltech Health App Evaluator domains and criteria categories are analysed against the CEN ISO/TS 82304-2:2021, and outcomes are sorted into categories to create a basis for comparison between the cases.

During variable-oriented analysis standard variables across the cases and in themes are identified [67], [68]. In this study variable-oriented analysis is used for providing additional depth for the overview. During the questionnaire analysis, at first the Taltech Health App Evaluator questions are mapped with the CEN ISO/TS 82304-2:2021 domains and criteria, and questions are sorted and grouped by similarities and differences. Thereafter variables are labelled, standard topics are identified and sorted

thematically. The standard variables mapping enables detailed content comparison and analysis.

In patterns matching phase patterns occurrence, lack or alternatives; and recurring and complementing items are identified. Common patterns, converging evidence and overlaps in the featured criteria are searched for [68], [70], [67]. Pattern matching is one of the most preferred techniques for case studies [66]. The predicted pattern made before data collection is compared to an empirically based pattern on the findings of the case study [66]. Internal validity of the study is established during the pattern matching phase [66].

During patterns matching the Taltech Health App Evaluator and the CEN ISO/TS 82304-2:2021 matching and mismatching topics are identified, counted, and summarized. Percentage values are calculated for each topic by criteria and domain. Matching percentage is calculated as a ratio of topics covered in both frameworks divided by all the CEN ISO/TS 82304-2:2021 topics amount. This shows the Taltech Health App Evaluator content coverage compared to the CEN ISO/TS 82304-2:2021. From identified recurring and complementing items common patterns and converging evidence is searched.

Case study research design can support investigation of existing theory's "gaps and holes" with the final aim of developing theoretical explanations [71]. In this thesis gaps and holes analysis is used for complementing and concluding the results of the criteria analysis and adding profoundness to the analysis by highlighting strengths and limitations across compared frameworks in the cases.

As a part of cross-case analysis each case is written up in a detailed matrix referencing standard variables, and a meta-matrix is created by synthesis and compilation of variables from the cases [68], [70], [67]. The matrices present compiled cross-case data displayed in a content analytic summary table organized by concepts and served during the study as a foundation for examinations. Conceptually organized "cross-case content analytic summary tables can illuminate how concepts play out in different cases [72]". In this study, the meta-matrix created by the author is used as a study database.

To expand the understanding, information beyond the evidence collected from the cases is complimented by reviewed literature and published data. This provides data on both,

on the cases and context. Contextual conditions are part of the investigation in case studies, and these are not delineated and controlled [71].

To ensure quality of the study and to increase validity and reliability of research statements three major triangulation principles are followed in case studies: multiple sources are used, a case study database is created, and a chain of evidence is maintained [66]. During the data collection phase, using multiple sources and establishing the chain of evidence supports the study's construct validity and use of case study protocol and database enhance study's reliability [66].

For ensuring reliability of this study, during the data collection phase:

- a case study protocol, and
- a case study database was created and used by the author.

Construct validity of this study is ensured during the data collection phase by

- maintaining the chain of evidence by the author.
- However, the study is based only on documentary sources and no multiple sources of info were used as no interviews, observations or surveys were carried out, limiting the construct validity.
    - Also, for documentary sources as well grey literature was used due to the lack of published data in the scientific literature on this quickly evolving field.
    - The used research and frameworks present a non-exhaustive overview of the existing digital health evaluation literature, as this was not the aim of this study.

In single case studies, limitation may commonly be low generalizability. Cross-case analysis and synthesis is a basic method that helps to generalize beyond a single case. Multiple-case study performed on two or more cases is preferable over single-case studies as this approach gives possibility of direct replication of findings and substantial analytic benefits by providing possibility for more powerful analytic conclusions and enhanced generalizability. Also, a stronger test of theory is provided while studying

issues across cases [66], [68]. Replication logic of multiple-case study supports external validity of the study by allowing greater generalizability [66].

External validity of this study is ensured during data analysis phase by

- using analytic generalization and replication logic of multiple case studies.

Internal validity of this study is ensured during data analysis phase by

- using patterns matching.

3 Results

3.1 Evaluation principles in cases

The following paragraphs describe the case studies results on evaluation principles used in the CEN-ISO/TS 82304-2:2021 “Health and wellness apps - Quality and reliability” and the TalTech Health App Evaluator with focus on the purpose, value, creator, creation time, geographical application area, optionality, stakeholders, and covered technologies of both frameworks. Summary of principles is in Table 15 on page 53.

The TalTech Health App Evaluator

The first case, Taltech Health App Evaluator, was launched in March 2021 by TalTech E-Medicine Centre. The framework was created in cooperation with Digital Health Curriculum lecturers, doctoral students, analysts, and experts. The purpose of the TalTech Health App Evaluator is to assess and give a comprehensive informed overview about apps having an impact on health and health behaviour that are used in Estonia [16].

There are several target groups of the TalTech Health App Evaluator. The main ones are the people using the health app and who are seeking for more information to support their health either on their own or by the recommendation of a healthcare professional. The second major target group are healthcare professionals who will get support before recommending apps to their patients. Successful completion of the evaluation process creates confidence among the users, the recommenders, and the creators, and demonstrates apps benefits in the healthcare system by making it more efficient [16].

The Health App Evaluator provides comprehensive information on these apps that have an impact on health and are processing health data [16]. The evaluation process confirms that a chosen health app works for health as promised, is safe, complies with technological standards and is suitable for widespread use [65]. The homepage is available only in Estonian language [65].

The CEN-ISO/TS 82304-2:2021

Second case, the CEN ISO/TS 82304-2:2021 “Health and wellness apps - Quality and reliability” technical specification, was launched in July 2021. The framework was created by the European standardization technical committee CEN/TC 251 Health Informatics in cooperation with the ISO and the International Electrotechnical Commission IEC/TC 62 Electrical equipment in medical practice, including a project team from 14 countries (Australia, Belgium, China, Finland, France, Germany, Ireland, Italy, Japan, Netherlands, Nigeria, Sweden, United Kingdom, and United States), and is available in English language. The CEN ISO/TS 82304-2:2021 was developed based on an overview of already existing standards and evaluation frameworks, and on Delphi consensus study, surveys, and interviews. The European Commission has been supporting and financing the project. The technical specification is approved by CEN for provisional application for the initial period of three years and may be converted into a European Standard thereafter [17], [73], [9].

The purpose of the CEN ISO/TS 82304-2:2021 is to create a standardized health and wellness apps quality criteria and a reliable apps rating system. These apps could assist people managing, maintaining, or improving physical, mental, or emotional health, as well as care delivery. The technical specification can be beneficial for supporting wider adoption of high-quality apps in the context of limited healthcare resources and complicated access to healthcare. The health apps have high potential to support management of chronic health conditions, healthier lifestyles, and ageing people. The CEN ISO/TS 82304-2:2021 is an international framework with worldwide application scope based on different local, national, and international initiatives, guidelines and standards, like ISO, IEC, HL7 [17], [73], [9], [20].

The focus of the technical specification is on a special form of health software - health and wellness apps. It is applicable to solutions that are marketed as health apps or to health software that is released as an app [17]. The CEN ISO/TS 82304-2:2021 follows principles defined in IEC 82304-1:2016 according to which all health apps are classified as health software and “any health software that is an app is a health app” [57]. However, the borders of classifying software as apps and not apps are blurred and constantly changing. The title, “Health and wellness apps - Quality and reliability” indicates that the scope of this specification is wider than only medical and clinical apps recommended by health professionals or healthcare providers, including prevalent

classes in use in app stores and libraries. Covered apps could be as well medical devices, though the guidelines for achieving medical device regulation compliance remain out of the scope of this framework [17].

The CEN ISO/TS 82304-2:2021 supports app creators throughout the entire life cycle of health apps designing, developing, testing, releasing, and updating processes resulting in better apps for users. The document is created as well for health app assessment organizations to support evaluation, quality labelling and reporting, as well consulting or offering additional assessment services. Also, specification development organizations can utilize this document either for establishing specific user cases or for creating additional context-specific questionnaires on quality and reliability assessment related aspects like interoperability or local legislation. These results could be shared to users as additional assessment outcomes. National and regional authorities and digital marketplace providers may require apps promoted in their area to have a quality label from a trusted app evaluation organization. The health app quality label and quality report support users, customers and recommenders while choosing the most suitable high-quality reliable apps [17].

The uniform assessment principles and quality requirements supplemented with third-party evaluation possibility and visualized results will enhance the health apps uptake [17]. This standard is anticipated to support app quality requirements international harmonization and reduce the creation of new diverse evaluation systems in each country [20]. An overview of the principles used in both cases, the TalTech Health App Evaluator and the CEN-ISO/TS 82304-2:2021, is presented in Table 15 on page 53.

The focus in the next paragraphs is on evaluation criteria used in the CEN-ISO/TS 82304-2:2021 and in the TalTech Health App Evaluator. At first, patterns in domains and criteria, identified during cross-case analysis, are introduced. Thereafter, standard variables in the form of topics identified from questions in both cases questionnaires during variable-oriented analysis, are presented. Subsequently, similarities and differences across the topics in cases revealed during pattern matching are introduced. For additional comparison gaps and holes analysis is used for complementing and concluding the results of the criteria analysis.

3.1.1 Domains and criteria

The CEN-ISO/TS 82304-2:2021 is divided into five assessment domains: “Product information”; “Healthy and safe”; “Easy to use”; “Secure data”; and “Robust build” sections. In the TalTech Health App Evaluator the four domains are: “General information”; “Level of user and clinical research”; “Privacy and security”; and “Interoperability and data quality”. One of the domains, “Level of user and clinical research” overlaps with the CEN-ISO/TS 82304-2:2021 “Healthy and safe” and “Easy to use” domains. In the CEN-ISO/TS 82304-2:2021 framework 13 assessment criteria are used while in the TalTech Health App Evaluator 19 subtopics are identified. In the Table 2 on page 36 evaluation domains and criteria from both frameworks are listed.

Table 2. Evaluation domains and criteria.

CEN-ISO/TS 82304-2:2021	TalTech Health App Evaluator
Product information Product; App manufacturer	**General information** Technical description Purpose and content Use
Healthy and safe Health requirements Health risks Ethics Health benefit Societal benefit	**Level of user and clinical research** Involvement of relevant organizations and professionals Background research Research with the app
Easy to use Accessibility Usability	
Secure data Privacy Security	**Privacy and security** Privacy Policy; Data collection; Consent; User rights; Data sharing; Cookies; User age; Marketing; Notification; Authentication; Data storage Security issues
Robust build Technical robustness Interoperability	**Interoperability and data quality** Interoperability and data quality

Source: Results of the case study compiled by author [16], [17].

According to the results of the pattern analysis on Table 2 on page 36, despite similar structure and domains setup, the depth of the criteria coverage in both frameworks varies considerably. While the CEN-ISO/TS 82304-2:2021 is more profound in the

"Healthy and safe" and "Easy to use" domains, the TalTech Health App Evaluator has more detailed coverage on privacy-related criteria on "Privacy and security" domain according to this high-level analysis.

3.1.2 Topics

The following paragraphs present results from the profound content analysis conducted on the evaluation questionnaires from both cases. The CEN-ISO/TS 82304-2:2021 questionnaire contains 81 and the TalTech Health App Evaluator questionnaire a similar amount - 82 questions, altogether 163 questions. During variable-oriented analysis standards variables in questionnaires were identified and individual questions in both frameworks were grouped by similar topics based on overlapping content to enable comparability and generalizability across the cases. The questions used in the TalTech Health App Evaluator were matched to the CEN-ISO/TS 82304-2:2021 questions and domain structure to categorize similarities and differences between the frameworks. Altogether 92 unique topics were identified from 163 questions by the author. The following Table 3 on page 37 compiles together an overview of domains, criteria, questions, and topics amount in cases.

Table 3. Domains, criteria, questions, and topics.

	CEN-ISO/TS 82304-2:2021	**TalTech Health App Evaluator**	**Total**
Domains	5	4	
Criteria	13	19	
Questions	81	82	**163**
Topics	64	49	**92**

Source: Results of the case study compiled by author [16], [17].

On the following paragraphs detailed comparison results from matching the TalTech Health App Evaluator topics to the CEN-ISO/TS 82304-2:2021 are presented across the five domains and in three categories, at first overlapping topics and their amount is shown, and thereafter unique topics in both frameworks are presented.

Product information

In the first, "Product information" domain, **product** and app **manufacturer** related criteria are covered, as illustrated in Table 4 on page 38.

Table 4. Product information.

Criteria	Topics in both	Only in ISO/TS	Only in TalTech
Product	**2**	**2**	**4**
	operating systems / platforms	name / icon	needs an external medical or non-medical device to work
	languages	app access instructions for the app assessment organization	recommenders
			functionalities
			instructions for use
Manufacturer	**1**	**0**	**0**
	manufacturer		

Source: Results of the case study compiled by author [16], [17].

While both frameworks ask for the information regarding manufacturer and product's operating systems and languages; the CEN-ISO/TS 82304-2:2021 requires name, icon, and access instructions as well; and the TalTech Health App Evaluator inquires additionally about the app's main, unique, and additional functionalities; user instructions; prescribers and recommending organizations; and the app's need for an external medical or non-medical device to work.

Healthy and safe

In the "Healthy and safe" domain, the focus is on health risks and benefits. In this domain health requirements, health risks, ethics, health benefit and societal benefit criteria are covered. At first, **health requirements, health risks,** and **ethics** overview is presented in Table 5 on page 39.

Table 5. Health requirements and health risks, ethics.

Criteria	Topics in both	Only in ISO/TS	Only in TalTech
Health requirements	**5**	**1**	**0**
	intended users / target audience	peer reviewed scientific literature used in the development	
	app age-appropriateness		
	app purpose		
	medical device		
	professionals involvement to development		
Health risks	**0**	**4**	**0**
		risks of the health app analysed, controlled, accepted	
		health professional approval before use	
		info to potential customers about health risks, contra-indications and limitations of use	
		process to collect and review safety concerns	
Ethics	**0**	**1**	**0**
		ethics	

Source: Results of the case study compiled by author [16], [17].

In **health requirements** criteria five topics are covered similarly by both frameworks, dealing with the app's intended users and use, health issues, purpose, and age-appropriateness; EU MDR applicability and risk class [25]; and health professionals, medical institutions, and governmental organizations involvement to the development of the app. The CEN-ISO/TS 82304-2:2021 adds the use of peer reviewed scientific literature during the app development. At the same time, **health risks** criteria are not covered in the TalTech Health App Evaluator including four topics on analysing, controlling, and accepting residual risks; approving the use by the health professional; informing customers on health risks; and establishing processes dealing with safety concerns. Likewise, **ethics** criteria related to the assessment and approval of ethical challenges is not included into the TalTech Health App Evaluator.

Remaining criteria on "Healthy and safe" domain, **Health and societal benefits** comparison results are shown in Table 6 on page 40.

Table 6. Health and societal benefits.

Criteria	Topics in both	Only in ISO/TS	Only in TalTech
Health benefit	**5**	**4**	**4**
	health benefit / clinical goals of the app	info about the health interventions applied and the need for support of a health professional	journals where articles about the app are published
	evidence available to support the health benefit	level of the evidence is appropriate - peer reviewed abstracts	measurements
	peer reviewed research with the app	health information in the app	periodical users feedback asked
	use of advertising mechanisms in the app disclosed and clearly distinguishable	funding sources	the NICE Digital Health Technology Assessment Framework evidence
	costs awareness by users		
Societal benefit	**0**	**1**	**0**
		evidence of a societal benefit	

Source: Results of the case study compiled by author [16], [17].

While the **societal benefits** criteria are present only in the CEN-ISO/TS 82304-2:2021, the approach towards **health benefits** criteria is more diverse. Five topics on health benefits that are present in both frameworks cover the app-related peer reviewed and previous research, evidence on health benefits, and users' awareness of related advertisements and costs. The CEN-ISO/TS 82304-2:2021 adds four additional themes related to users' awareness, evidence level appropriateness, health information, and disclosure of funding sources. In the TalTech Health App Evaluator questionnaire as well additional four more topics are covered related to measurements made in the app, their accuracy, and published studies; app's evidence level by the NICE Framework [4], availability and meeting of best practice standard on proof of effectiveness; and users' feedback methodological collection. **Health benefit** is defined in the CEN-ISO/TS 82304-2:2021 as "positive impact or desirable outcome of the use of health software on the health of an individual [17]".

Easy to use

The third, "Easy to use", domain covers **accessibility** and **usability** criteria to be evaluated during the health apps assessment, see Table 7 on page 41.

Table 7. Accessibility and usability.

Criteria	Topics in both	Only in ISO/TS	Only in TalTech
Accessibility	**0**	**2**	**0**
		app is WCAG 2.1 AA or AAA compliant	
		app age-appropriateness	
Usability	**1**	**4**	**0**
	intended users involvement	understanding of users, tasks and environment	
		error & misuse measures	
		product information, instructions, help available to potential customers and users	
		collection of data on usability	

Source: Results of the case study compiled by author [16], [17].

While the topic on the app's intended user's involvement to the design and development is covered by both frameworks, the remaining themes under **accessibility** and **usability** criteria are covered only by the CEN-ISO/TS 82304-2:2021. The accessibility topics besides app's age-appropriateness check are mainly based on references to the compliance with the Web Content Accessibility Guidelines WCAG 2.1 [74]. The usability criteria focus is on understanding and supporting users, avoiding misuse, and collecting data on usability throughout the app's lifetime. Definitions in the CEN-ISO/TS 82304-2:2021 are referred as follows: "**Accessibility**: extent to which products, systems, services, environments and facilities can be used by people from a population with the widest range of user needs, characteristics and capabilities to achieve identified goals in identified contexts of use [75], p. 3.2.2". "**Usability:** extent to which a system, product or service can be used by specified users to achieve specified goals with effectiveness, efficiency and satisfaction in a specified context of use [76], p. 3.13".

Secure data

The fourth, "Secure data" domain incorporates **privacy** and **security** criteria. At first, the **privacy** criteria coverage is presented on Table 8, page 42.

Table 8. Privacy.

Criteria	Topics in both	Only in ISO/TS	Only in TalTech
Privacy	**4**	**5**	**10**
	PII collection, processing	data minimization	Consent
	PII - health related	stored data erasing, reviewing	PII - users
	privacy policy / statement	security controls and privacy protection	data storage - deletion after being inactive
	compliance officer	default opt-in	PII - data sharing
		breaches reporting	cookies
			minors
			consent from an adult
			PII protection
			location of data storage
			data storage duration

Source: Results of the case study compiled by author [16], [17].

Under privacy, both frameworks cover four topics related to Personally identifiable information (PII) collection and processing; availability of privacy policy or statement; and presence of a data protection specialist. There are five themes present only in the CEN-ISO/TS 82304-2:2021 dealing with data minimization; stored data review and deletion; delivery of promises on security controls and privacy protection; PII sharing with third parties and PII breaches reporting. The TalTech Health App Evaluator adds ten more topics related to consent; accessing, changing, sharing, protecting, and deleting PII; cookies; user's age and adult's consent; data storage location and duration. For **privacy** definition in the CEN-ISO/TS 82304-2:2021 reference to ISO/TS 27790:2009, 3.56 [77] is used as follows: "…freedom from intrusion into the private life or affairs of an individual when that intrusion results from undue or illegal gathering and use of data about that individual [17]".

Security criteria coverage comparison is presented in Table 9, page 43.

Table 9. Security.

Criteria	Topics in both	Only in ISO/TS	Only in TalTech
Security	**2**	**9**	**0**
	authentication	ISO/IEC 27001 or a recognized equivalent	
	security vulnerabilities reported	information security risk assessment	
		a secure by design process	
		reliable & maintained app components	
		unauthorized access, modifications	
		PII - processing compatible with privacy statement	
		PII - encryption	
		security vulnerabilities tested	
		information security policy available	

Source: Results of the case study compiled by author [16], [17].

Both frameworks cover secure user authentication and security vulnerabilities management topics. The CEN-ISO/TS 82304-2:2021 additionally refers to nine topics related to information security risk assessment; secure design; reliable and maintained third-party software libraries and components; existing app source code protection process; PII legitimate processing and adequate encryption; regular security testing; and availability of information security policy. **Security** is defined in the CEN-ISO/TS 82304-2:2021 as a "condition that results from the establishment and maintenance of protective measures that ensure a state of inviolability from hostile acts or influences. Hostile acts or influences could be intentional or unintentional [17]".

.

Robust build

The last domain, "Robust build" entails criteria of **technical robustness** and **interoperability**. **Technical robustness** criteria overview is on Table 10, page 44.

Table 10. Technical robustness.

Criteria	Topics in both	Only in ISO/TS	Only in TalTech
Technical robustness	**0**	**8**	**0**
		product requirements documented	
		development process covers the standards, methods and tools to be used	
		secure coding standard followed	
		configuration management plan established	
		a significant increase or spike in demand process	
		a validation and verification plan used	
		a validation and verification process established	
		a maintenance process established	

Source: Results of the case study compiled by author [16], [17].

This technical robustness field is covered only by the CEN-ISO/TS 82304-2:2021 in eight topics, inquiring info on product requirements documentation; standard-based software development; plan and processes for configuration management; dealing with significant increase or spike in demand, validation, and verification; and maintenance.

Interoperability criteria comparison is presented in Table 11, page 45.

Table 11. Interoperability.

Criteria	Topics in both	Only in ISO/TS	Only in TalTech
Inter-operability	**1**	**2**	**10**
	data export and exchange	specifications and implementation guides for all the APIs	standard terminology use
		specifications and implementation guides for the terminology	standard clinical data models
		data validation	API's for exchanging data with other apps
			data compatibility with other wireless devices
			messages exchange formats
			automatic data entry via direct import
			workflows implementation
			documents formats
			GDPR compliant cloud service to store data
			interoperability and connectivity strategy

Source: Results of the case study compiled by author [16], [17].

While both frameworks contain questions regarding data export and exchange, three topics are covered only in the CEN-ISO/TS 82304-2:2021 – users' access to the specifications and implementation guides for the APIs (application programming interfaces) and to the terminologies used; and the app data validation. The TalTech Health App Evaluator is including ten additional topics related to standard terminology use on data collection; standard clinical data models use; providing API's for exchanging data with other apps; data compatibility with other wireless devices; message exchange formats; automatic data entry via direct import; workflows implementation; documents formats; GDPR compliant cloud service use for storing data; and interoperability and connectivity strategy for interfacing with other health information systems. **Interoperability** is defined in the CEN ISO/TS 82304-2:2021 based on the "IEEE standard computer dictionary" [78] as "ability of two or more systems or components to exchange information and to use the information that has been exchanged [78]".

3.1.3 Topics comparison in domains and criteria

The overview of the TalTech Health App Evaluator topics coverage comparison to the CEN ISO/TS 82304-2:2021 in Table 12 on page 45 illustrates results from the analysis across frameworks, domains, and criteria. At first, similarities and differences summary revealed during pattern matching is introduced. The percentage match is calculated as a ratio of topics existent in both frameworks divided over total number of the CEN-ISO/TS 82304-2:2021 topics in respective categories, while higher percentage shows higher similarity in corresponding domain and criteria.

Table 12. Topics analysis summary.

Criteria/Questions	In both / In ISO/TS	In both	Only in ISO/TS	Only in TalTech	ISO/TS total	TalTech total
Total topics	**33%**	**21**	**43**	**28**	**64**	**49**
Product information	**60%**	**3**	**2**	**4**	**5**	**7**
Product	50%	2	2	4	4	6
App manufacturer	100%	1	0	0	1	1
Healthy and safe	**48%**	**10**	**11**	**4**	**21**	**14**
Health requirements	83%	5	1	0	6	5
Health risks	0%	0	4	0	4	0
Ethics	0%	0	1	0	1	0
Health benefit	56%	5	4	4	9	9
Societal benefit	0%	0	1	0	1	0
Easy to use	**14%**	**1**	**6**	**0**	**7**	**1**
Accessibility	0%	0	2	0	2	0
Usability	20%	1	4	0	5	1
Secure data	**32%**	**6**	**13**	**10**	**19**	**16**
Privacy	44%	4	5	10	9	14
Security	18%	2	9	0	11	2
Robust build	**8%**	**1**	**11**	**10**	**12**	**11**
Technical robustness	0%	0	8	0	8	0
Interoperability	33%	1	2	10	3	11

Source: Compiled by author [16], [17].

The overview of the topics distribution across frameworks highlights the presence of the TalTech Health App Evaluator topics in the CEN ISO/TS 82304-2:2021 framework. According to the detailed questionnaire and topics content analysis, the highest similarity across the frameworks, 60% is on the “Product information” domain followed by 48 % on “Healthy and safe” and 32% on the “Secure data” group. The lowest

presence of the TalTech Health App Evaluator topics compared to the CEN ISO/TS 82304-2:2021 is 14% on “Easy to use” and 8% on “Robust build” domains. The overall presence of the TalTech Health App Evaluator topics in the CEN-ISO/TS 82304-2:2021 is 33%.

While the previous topics analysis highlighted patterns and content occurrence and absence in cases, the following criteria analysis illustrates profoundness of the coverage in both frameworks. Gaps and holes analysis is used for the additional comparison that summarizes the results in Table 13 on page 47 by the number of topics and count of matching criteria in both frameworks.

Table 13. Criteria analysis summary.

Criteria/Questions	**ISO/TS higher coverage by criteria**	**TalTech higher coverage by criteria**	**Same coverage in both by criteria**
Total criteria	**7**	**3**	**3**
Product information	**0**	**1**	**1**
Product		TalTech	
App manufacturer			Both
Healthy and safe	**3**	**0**	**2**
Health requirements			Both
Health risks	ISO/TS		
Ethics	ISO/TS		
Health benefit			Both
Societal benefit	ISO/TS		
Easy to use	**2**	**0**	**0**
Accessibility	ISO/TS		
Usability	ISO/TS		
Secure data	**1**	**1**	**0**
Privacy		TalTech	
Security	ISO/TS		
Robust build	**1**	**1**	**0**
Technical robustness	ISO/TS		
Interoperability		TalTech	

Source: Compiled by author [16], [17].

The gaps and holes analysis highlights evaluation criteria where the CEN ISO/TS 82304-2:2021 is more profound as well where the TalTech Health App Evaluator current questionnaire is more detailed, and where both frameworks present similar coverage. In the following overview in Table 14 on page 48 final summary on the criteria coverage comparison in domains in both frameworks is presented.

Table 14. Criteria coverage comparison in domains.

Domain / Criteria	Equal coverage	TalTech profounder	ISO/TS profounder
Product information	App manufacturer	Product	
Healthy and safe	Health requirements Health benefit		Health risks Ethics Societal benefit
Easy to use			Accessibility Usability
Secure data		Privacy	Security
Robust build		Interoperability	Technical robustness

Source: Compiled by author [16], [17].

From the total 13 assessment criteria the TalTech Health Apps Evaluator has more profound coverage in three: “Product information”, “Privacy” and “Interoperability”. The CEN ISO/TS 82304-2:2021 has more extensively covered the criteria in the areas of “Health risks”, “Ethics”, “Societal benefit”, “Accessibility”, “Usability”, “Security”, and “Technical robustness”. Similar coverage in both frameworks is in “App manufacturer”, “Health requirements” and “Health benefit” criteria.

3.2 Evaluation outcomes in cases

The TalTech Health App Evaluator

The evaluation process is voluntary for app developers. The created health apps registry webpage lists the apps that have passed evaluation and describes their compliance to assessment criteria. The Taltech Health App Evaluator provides assessment results on the Evaluator webpage as a traffic lights system, as of April 2021 however only in Estonian language. Three major assessment domains are: Level of user and clinical research, Privacy and security, and Interoperability and data quality. Each is evaluated in four levels – high, medium, low, or not applicable as shown in the Figure 2, page 49. Besides the colour-scale also a comprehensive overview of questions and answers to the

evaluation questionnaire are provided in the Evaluator webpage for each evaluated app, however without comprehensive published guidelines and calculation principles [16].

Figure 2. The Taltech Health App Evaluator outcomes visualisation [16].

The CEN-ISO/TS 82304-2:2021

As a result of the voluntary the CEN-ISO/TS 82304-2:2021 evaluation process, a quality label designed similarly to the EU Energy label, and an assessment report are created to support informed decision making of health apps customers and end-users. An example of the standardized label is shown on Figure 3, page 49 [17].

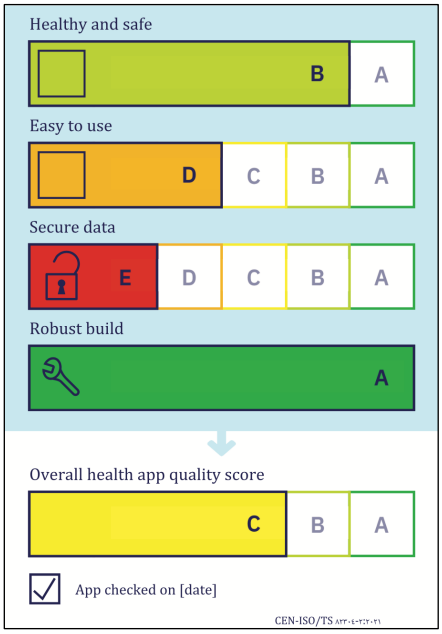

Figure 3. The CEN-ISO/TS 82304-2:2021 evaluation outcomes visualisation [17].

The label visualizes the outcomes of quality assessment in four domains: Healthy and safe, Easy to use, Secure data and Robust build. Five scaling levels are using letters from A to E and colours from green to red respectively. For calculating the quality label and colour code score, the questions in the assessment questionnaire have been assigned numeric values and thresholds are established for the four domains and five scaling levels. The label can be used in a repository or digital marketplace [17].

4 Discussion

Previous chapters provided input for the localization of the international knowledge in Estonian context and for the comparison of similarities and differences in selected cases. In this chapter results from the cross-case study and analysis are discussed. The approaches in three focus areas - health apps evaluation principles, criteria and outcomes are compared in the Estonian TalTech Health App Evaluator and in the international health apps evaluation harmonization initiative, the CEN ISO/TS 82304-2:2021 "Health and wellness apps - Quality and reliability" technical specification to understand similarities and differences and analyse what these may entail, as the aim of this research stated. The results are complemented with input from the theoretical background, and generalizations and suggestions for the future are made.

4.1 Discussion of key findings

4.1.1 Evaluation principles

According to the evaluation theory, the focus of the assessment may be on the process, achievement of objectives [18], determining value, efficiency, effectiveness, impact, and sustainability [19]. For reaching to valuable results all the suitable, applicable, necessary, and useful methods may be used, though it's relevant to establish goals that are appropriate, relevant, and are focusing on stakeholders needs [18]. However, instead of following the evaluation theory, in evaluation practice pragmatic approaches that are based on extensive experiences and have proven to work well are often applied [18]. In European health apps evaluation frameworks, variable heterogeneous and diverse approaches towards health apps evaluation are used [9]. The following discussion on evaluation principles is focusing on the assessment purpose, value, creator, creation time, geographical application area, stakeholders, and covered technologies.

The purpose of both compared frameworks, the CEN-ISO/TS 82304-2:2021 and TalTech Health App Evaluator is similar - to establish a consistent set of quality criteria and rating system for health apps evaluation and create a profound informed overview. Evaluated health apps can support people's physical, mental, or emotional health

management, maintenance, improvement, or delivery of care. Though in different wordings, the intention is to support adoption of quality health apps dealing with healthier lifestyles, chronic health conditions, or ageing of people; to have a supportive impact on health and health behaviours, and thereby lower burden on limited healthcare resources. For comparison, additional purposes in the highlighted European frameworks are supporting the incorporation of digital health solutions within healthcare systems and enabling their reimbursement.

The evaluation can create value by gathering and generating quality information for decision-making [19]. The value of the Taltech Health App Evaluator is in the confidence provided to the health app users and recommenders by the comprehensive informed overview of the mobile health and health behaviour apps created and used in Estonia, which ultimately benefits the healthcare system. The CEN-ISO/TS 82304-2:2021 has a wider scope, including health apps evaluation frameworks creators and providing to them quality requirements international harmonization opportunity instead of the creation of new diverse additional evaluation systems. The harmonization approach is relevant as well for the Estonia and the Taltech Health App Evaluator.

Whereas formative evaluations are used during the development process as a prospective and proactive opportunity for ensuring and improving quality; summative retrospective evaluations are establishing accountability by informing stakeholders on quality and safety of assessed products and services [18]. While the CEN-ISO/TS 82304-2:2021 is providing support to developers already during the health app development phase, the Taltech Health App Evaluator may become more valuable for developers if the framework offers as well formative evaluation possibility, and the requirements could thereby be considered already on the development process for quality improvements.

The CEN-ISO/TS 82304-2:2021 has been created in cooperation with the ISO, European standardisation technical committee, International Electrotechnical Commission, and a project team from 14 countries with support by the European Commission. The Taltech Health App Evaluator was launched by the Tallinn University of Technology’s E-Medicine Centre in cooperation with Digital Health Curriculum lecturers, doctoral students, analysts, and experts. Both frameworks are originating from the year 2021, though the Taltech Health App Evaluator was launched already in

March, while the CEN-ISO/TS 82304-2:2021 was released in July. Likewise, highlighted European frameworks have as well been released only recently, in 2019-2021. Whereas the CEN-ISO/TS 82304-2:2021 is an international framework with worldwide application scope, the Taltech Health App Evaluator is for apps that are used in Estonia; similarly, the other European frameworks have national focuses.

Contrary to the CEN-ISO/TS 82304-2:2021 and majority of national frameworks that are available as well in English, the Taltech Health App Evaluator is available only in Estonian language. The frameworks that are using for assessment and results publication only the local language may limit their scalability [9]. Therefore, expanding the language scope can have a supporting impact on the framework's applicability on wider selection of international health apps.

The stakeholder focus varies across European frameworks, primarily it is on supporting end-users - people and healthcare professionals [9]. Both studied frameworks focus on creating a quality label and a comprehensive report that can serve the people using the health apps - patients and caregivers; and health professionals recommending the apps. Additionally, the CEN-ISO/TS 82304-2:2021 creates the value of the technical specification being a guideline for other interested counterparties and target groups. The document may benefit app creators, developers, and manufacturers, also app assessment organizations, health authorities and policy makers. The health app creation is supported across the full lifecycle leading to higher-quality solutions for users. The Taltech Health App Evaluator may add stronger manufacturers and policy makers focus into the scope.

The approach in both frameworks towards assessed technologies is similar. The CEN ISO/TS 82304-2:2021 has specially highlighted that the focus is on both, health and wellness apps, including not only the apps that are recommended by healthcare professionals, but as well the apps that are available from app stores and libraries. The Taltech Health App Evaluator focus is on apps that are having an impact on health and health behaviour. The primary focus of both frameworks is on non-medical devices, however neither of them excludes apps that may also be medical devices.

Comparison of the principles used in studied cases and recommendations are presented in Table 15 on page 53.

Table 15. Evaluation principles in cases.

Name	CEN-ISO/TS 82304-2:2021	TalTech Health App Evaluator	Conclusion
Purpose	standardized quality criteria, reliable apps rating system	comprehensive informed overview	similar
Value	health and wellness apps evaluation framework; international harmonization and reduction of the creation of new diverse evaluation systems; support to users, customers and recommenders	confidence among the users, the recommenders and the creators, benefits in the healthcare system: health app works for health as promised, is safe, complies with technological standards and is suitable for widespread use	harmonization and formative evaluation possibilities
Creator	European standardization technical committee with ISO, *International Electrotechnical Commission,* a project team from 14 countries. The European Commission has been supporting and financing the project.	TalTech E-Medicine Centre - Digital Health Curriculum lecturers, doctoral students, analysts, and experts	
Year	July 2021	March 2021	similar
Area	an international framework with worldwide application scope	framework for apps that are used in Estonia	wider geographic applicability, more languages
Stakeholders	app creators, assessment organizations, specification development organizations, national and regional authorities and digital marketplace providers	people using the health app; healthcare professionals recommending apps	include manufacturers and policy makers
Technology	health and wellness apps, applicable to solutions that are marketed as health apps or to health software that is released as an app; apps could be medical devices	apps having an impact on health and health behaviour; may be medical devices	similar

Source: Results of the case study compiled by author [16], [17].

To **conclude**, the approach towards evaluation principles in both compared frameworks, the CEN-ISO/TS 82304-2:2021 and the TalTech Health App Evaluator is essentially similar. Both frameworks share **similar purposes**, **end-user,** and **technological scope**, and are originating from the same **year**, 2021. The Taltech Health App Evaluator could add **manufacturers and policy makers focus** into the scope, and additionally consider

providing an **earlier formative evaluation possibility** for manufacturers for quality improvements already during the development process. Selecting the **framework's harmonization** approach and using **more languages** can widen the geographic applicability of the evaluations and support wider access to international health apps. Evaluation principles that follow internationally recognized health apps quality requirements will help to ensure the TalTech Health App Evaluator assessments results reliability.

4.1.2 Evaluation criteria

The discussion on **evaluation criteria** is covering domains, criteria, questions, and topics in studied cases, the TalTech Health App Evaluator, and the CEN ISO/TS 82304-2:2021 (Table 3, page 37).

Domains and criteria in both frameworks were compared and analysed during the cross-case pattern analysis, detailed results of which are presented in paragraph 3.1.1 "Domains and criteria" on page 36. The compared frameworks are structurally similar and covered domains are largely overlapping; however, the further analysis highlighted considerable criteria coverage differences. Nevertheless, the similar baseline structure in both frameworks established a good basis for further comparison, analysis, and discussion. Also, similar structure of both frameworks lays a solid ground for further possible adjustments of the TalTech Health App Evaluator.

The variable-oriented analysis investigated standard variables across the cases and studied the comparable **topics** based on the **questions** mapping from both framework questionnaires. Similarities and differences between frameworks were highlighted in detail. Besides similar domains amount, five and four in both frameworks, also the number of questions in both questionnaires is almost identical, 81 in the CEN-ISO/TS 82304-2:2021 and 82 in the TalTech Health App Evaluator. From a total 163 questions altogether 92 unique topics were identified. The CEN ISO/TS 82304-2:2021 questions cover 64 topics compared to the TalTech Health App Evaluator's 49. The topics analysis introduced the comparison results and presence of each topic either in both frameworks, or only in the CEN-ISO/TS 82304-2:2021 or the TalTech Health App Evaluator, mapped against the CEN-ISO/TS 82304-2:2021 five domains and 13 criteria. Detailed results are shown in paragraph 3.1.2 "Topics" from page 37 onward. These are discussed together with the next, pattern analysis results.

Pattern matching was used for **topics** further review and **comparison**. For getting measurable results percentage values were calculated for each topic by criteria and domain. The presence of the TalTech Health App Evaluator topics in the CEN ISO/TS 82304-2:2021 framework was investigated. The analysis revealed in which criteria what topics match and what need to be adjusted. Summary overview is presented in Table 16 on page 57.

The overall TalTech Health App Evaluator topics presence in the CEN-ISO/TS 82304-2:2021 resulted in 33% match, meaning that 67% additional topics could be added to the TalTech Health App Evaluator for achieving equal coverage. In the following paragraph similarities and differences are highlighted at domains, criteria, and topics level. The highest match by domains, 60% is in the “Product information”, followed by 48% in “Healthy and safe”, 32% in “Secure data”, 14% in “Easy to use” and 8% in “Robust build.

In the “Product information” domain the information regarding manufacturer and product are included (Table 4, page 38). In this domain altogether three topics are covered by both frameworks, while the CEN-ISO/TS 82304-2:2021 adds two more themes to the mutually covered three topics. Under the “Healthy and safe” domain, the focus is on health requirements, health risks, ethics, health and societal benefits (Table 5, page 39 and Table 6, page 40). In both frameworks altogether 10 topics are the same in this domain, however the CEN-ISO/TS 82304-2:2021 introduces additional 11 ones in total to be considered for the TalTech Health App Evaluator. In the “Easy to use” domain accessibility and usability criteria is covered (Table 7, page 41). There is one common theme, under accessibility, and the CEN-ISO/TS 82304-2:2021 brings in additionally six topics that are missing in the TalTech Health App Evaluator. “Secure data” domain includes privacy (Table 8, page 42) and security (Table 9, page 43) criteria. In these domains six topics are present in both frameworks, however the CEN-ISO/TS 82304-2:2021 adds 13 extra topics altogether to these fields. “Robust build” domain covers technical robustness (Table 10, page 44) and interoperability (Table 11, page 45) criteria. Though profoundly present in both frameworks, the CEN-ISO/TS 82304-2:2021 has stronger focus with eight unique topics in technical robustness, and the TalTech Health App Evaluator has greater coverage with unique themes in the interoperability field. However, only one theme is covered by both frameworks under

this domain mutually, and one additional topic is introduced by the CEN-ISO/TS 82304-2:2021.

Whereas the previous overview presented the TalTech Health App Evaluator topics presence and coverage in the CEN-ISO/TS 82304-2:2021, the next, gaps and holes analysis looked at these frameworks separately, highlighted the strengths and weaknesses of both and illustrated profoundness of the criteria coverage in each individually (Table 13, page 47 and Table 14, page 48). According to this analysis, **the TalTech Health Apps Evaluator needs to be complemented within seven criteria**: "Health risks", "Ethics", "Societal benefit", "Accessibility", "Usability", "Security", and "Technical robustness". Similar coverage is in "App manufacturer", "Health requirements" and "Health benefit" fields and the TalTech Health Apps Evaluator is more profound in three: "Product information", "Privacy" and "Interoperability" categories.

To **conclude**, according to pattern analysis, the structure in compared frameworks at **domains level is similar**. The TalTech Health App Evaluator can add one additional domain, subsequently both frameworks will contain five domains. Similarity in both frameworks structures establishes a good basis for further possible adjustments of the TalTech Health App Evaluator. During the variable-oriented analyses individual 81 and 82 questions in both frameworks were compiled and mapped to 92 unique comparable topics; altogether 64 topics present in the CEN ISO/TS 82304-2:2021 and 49 in the TalTech Health App Evaluator's questionnaires were identified. Altogether the TalTech Health App Evaluator can **add 43 topics to the current 21** that are already present in both frameworks. The topics that can be added are presented in detail in paragraph 3.1.2. Pattern matching supported identifying patterns correspondence across frameworks and highlighted the differences in criteria and topics where considerable differences exist. The **overall coverage of the Taltech Health App Evaluator** content compared to the CEN ISO/TS 82304-2:2021 **is 33%**. For achieving equal coverage across frameworks questionnaires, **67% of additional missing topics can be added** to the TalTech Health App Evaluator.

The final summary of the evaluation topics and criteria comparison is presented in Table 16 on page 57.

Table 16. Evaluation topics and criteria in cases.

	Topics in both / In ISO/TS	All topics in ISO/TS	Topics in both	Unique topics only in ISO/TS	Criteria profounder
Total	**33%**	**64**	**21**	**43**	
Product information	**60%**	**5**	**3**	**2**	
Product	50%	4	2	2	TalTech
App manufacturer	100%	1	1	0	Similar
Healthy and safe	**48%**	**21**	**10**	**11**	
Health requirements	83%	6	5	1	Similar
Health risks	0%	4	0	4	ISO/TS
Ethics	0%	1	0	1	ISO/TS
Health benefit	56%	9	5	4	Similar
Societal benefit	0%	1	0	1	ISO/TS
Easy to use	**14%**	**7**	**1**	**6**	
Accessibility	0%	2	0	2	ISO/TS
Usability	20%	5	1	4	ISO/TS
Secure data	**32%**	**19**	**6**	**13**	
Privacy	44%	9	4	5	TalTech
Security	18%	11	2	9	ISO/TS
Robust build	**8%**	**12**	**1**	**11**	
Technical robustness	0%	8	0	8	ISO/TS
Interoperability	33%	3	1	2	TalTech

Source: Results of the case study compiled by author [16], [17].

Overall, the evaluation criteria that are based on an internationally recognized health apps quality requirements, ensures the TalTech Health App Evaluator assessments results quality.

4.1.3 Evaluation outcome

The following discussion on evaluation outcomes is focusing on the visualization and presentation of assessment results. Approaches towards quality labels, scores, reports, published principles, publishing results in repositories, providing guidelines, and frameworks optionality are covered.

The evaluation outcomes presentation has over the last five years become enriched by a variety of approaches like scores, labels, quality marks and vignettes; and new evaluated health apps repository webpages have evolved [20], [9]. Frameworks in investigated cases have followed the same pattern, and both aim to offer similar outcomes:

visualization of evaluation results, and a detailed presentation of assessment conclusions.

For the evaluation outcomes visualization, the CEN ISO/TS 82304-2:2021 quality label is designed in a similar way to the EU Energy label with an aim to improve transparency on the quality and reliability of health apps. The related scores for the four domains and five scaling levels are calculated based on published principles. The TalTech Health App Evaluator uses a traffic lights system for the results compliance assessment visualization in three assessment domains and at four levels – high, medium, low, or not applicable. However, the scores calculation principles are not made public. The TalTech Health App Evaluator could benefit from enhanced transparency if the used scores calculation principles are publicly available to interested stakeholders.

For outcomes presentation, the CEN ISO/TS 82304-2:2021 evaluation report combines answers to the questions in the assessment categories. The report may be applicable for informing decisions on the appropriate health app selection by users and recommenders to the adoption in care guidelines, contracts, and pathways. The TalTech Health App Evaluator report provides a comprehensive overview of questions and answers for assessed health apps' users and recommenders. Both frameworks provide the same reporting approach.

It is necessary to pay attention as well to the timeliness of effective and accurate presentation of the evaluation results to the proper audiences [18, p. 15]. The convenient publication of evaluation outcomes in an online repository, library or in an evaluator's webpage has become a prevalent approach in recent years [20]. The investigated cases have different approaches towards results publication. The role of the CEN-ISO/TS 82304-2:2021 is to establish a basis for the creation of a common label that can be used in a uniform way across the countries and repositories. However, the creation of a common repository for evaluated health apps has not been in the scope of this standard. The assessment results of the Taltech Health App Evaluator are available on the TalTech Health App Evaluator webpage, and these are listed by assessed app's names with added outcomes visualizations for an easier end-user's overview, and answers to evaluation questions are presented. Considering the potential future growth in the number of assessed health apps, the categorization by app types and specialization according to the final purpose in an easily findable form can benefit the Taltech Health

App Evaluator results end-users, especially if this is done in comparable format across frameworks and countries repositories.

The published CEN-ISO/TS 82304-2:2021 technical specification document itself serves as a guideline for the related stakeholders - developers and evaluation organizations and can be used for providing information about the health apps evaluation requirements and results [17]. The Taltech Health App Evaluator questions can be found from the webpage under the assessed apps pages [16], however there would be more value to developers and other interested stakeholders while these are presented as well all together in a comprehensive form. Similarly, to international practice [9] the questionnaire could serve as an additional function of being a guideline.

In Europe most of the national and regional frameworks operate on a voluntary basis and provide thereby secondary benefits [9]. Also, the ISO certification and compliance are voluntary [54]. Similarly, the TalTech Health App Evaluator assessment is a voluntary process for app manufacturers [16]. For comparison, the focus of highlighted national DHT evaluation frameworks is on these technologies integration into healthcare systems, and the evaluation process is mandatory for the apps aiming for reimbursement. The mandatory nature of few frameworks supports the health apps integration into the healthcare systems [9]. In the reimbursement-focused frameworks, like in Germany [44] and Belgium [46], the applicability precondition is being a CE-marked medical device, which already ensures the compliance to the clinical safety among other requirements [25]. The reimbursement opportunity increases health apps' relevance on having an impact on personal health [13]. Following the international trends and the EHIF initiatives [15], in a longer time horizon the TalTech Health App Evaluator has a good potential to expand its scope to support the reimbursement-focused assessments as well.

The evaluation outcomes comparison is presented in Table 17 on page 60.

Table 17. Evaluation outcomes in cases.

Name	CEN-ISO/TS 82304-2:2021	TalTech Health App Evaluator	Conclusion
Quality label	similar to the EU Energy label	a traffic lights system	similar
Scores	4 domains, 5 levels	3 domains, 4 levels	transparency
Report	yes	yes	similar
Published principles	yes	no	transparency
Repository	n/a	yes	comparability
Guideline	yes	no	guideline for manufacturers
Optionality	voluntary	voluntary	similar

Source: Results of the case study compiled by author [16], [17].

To **conclude**, the approach towards **outcomes visualization and presentation is similar** in the compared cases, however the TalTech Health App Evaluator can benefit from increased **transparency while the scores calculations principles are made publicly available**. Also, the **evaluation questionnaire can serve as a guideline for developers and other stakeholders while published** in a comprehensive format on the webpage. The TalTech Health App Evaluator a**ssessment results presentation format on the webpage**, considering the potential future growth in the number of assessed health apps, can achieve **better clarity and comparability** from the categorization by app types and specializations according to the final purpose in an easily findable form. The CEN ISO/TS 82304-2:2021 principles adoption can help to solve the questions about the TalTech Health App Evaluator report and quality label compilation principles **transparency**. The adopted framework can serve as an internationally accepted guideline for developers. **Harmonization** of the evaluation approaches can establish worldwide **comparability** and **recognizability** to the TalTech Health App Evaluator assessments results.

Author's contribution

The author has in this thesis systematized vast quickly growing international knowledge on health apps evaluation for making it beneficial for the Estonian context. The study's aim was achieved, and the research question was answered. The chosen case study and cross-case analysis and synthesis methods were suitable for the purposes of this study and enabled attainment of comparable results that formed the basis for the further discussion and recommendations development. The results provided an overview where

the Taltech Health App Evaluator stands in respect of the international health apps evaluation standardization and assessment harmonization initiative, the CEN-ISO/TS 82304-2:2021, which has been created with worldwide implementation potential.

4.2 Limitations

Due to the novelty of the health apps evaluation field that has anticipated rapid international developments over the last years, and resulting yet limited availability of scientific publications, the study has included also grey literature for background overview on selected countries and major frameworks, standards, regulations, and cases to achieve more holistic coverage of this quickly growing and changing area. As the health technologies assessment field has traditionally been used for direct informing of policies, the related materials have not been necessarily published in the scientific literature. This thesis as well presents a non-exhaustive current overview of the existing research and frameworks on digital health evaluation, as this was not the aim of this study. Though case studies may be based on multiple sources of data, this thesis is based only on documentary sources; and no interviews, observations, or surveys were a part of this study, which may possibly limit the construct validity. In this study cross-case analysis and synthesis methods are used to diminish the limitation of generalizability, strengthen validity and reliability, and improve representativeness that single-case studies otherwise may encounter. Replication logic of two cases is used due to the investigated theme's straightforwardness. The most similar case selection approach that emphasizes the comparability of characteristics was used in this multiple case study. To increase reliability and validity, case study protocol and case study database were used, and chain of evidence was maintained by the author throughout the study.

4.3 Further research

While case studies can provide detailed insights into the selected phenomenon and may expand theory, these are not used for testing it, therefore the applicability of proposed recommendations could be further examined by carrying out practical existing health apps evaluations. Whereas this study was focusing on comparison of health apps evaluation quality requirements, further studies may focus on the next, organizational

and policy related implementation steps. The relevance of health apps evaluation criteria could be investigated in greater detail; and research and analysis on practical impact assessments on target groups could be carried out to prioritize the further developments and focus on the most important ones at first. The theme could be investigated further by using larger number of cases to enable broader generalizations of the findings. As well, follow-up in international developments in this quickly changing field is crucial due to anticipated continuous changes. While this analysis focused on overall quality requirements comparison, the next investigations' focus could be on patients' adoption of new health apps, and on examining their viewpoints on what is relevant for the improvement of these innovations' uptake. Also, theoretical fields that have not yet found much inclusion either to the health apps development or evaluation, like behaviour change techniques, or improvement of usability by human–computer interaction modelling, are worthy of further investigation. Creation of additional evidence on the health apps quality and reliability can support the adoption of new innovative health technologies by empowered patients. The quickly evolving digital health technologies and health apps evaluation field provides plentiful opportunities for further investigation.

5 Summary

New evaluation initiatives have been launched internationally to ensure the quality and safety of health apps and minimize related risks in the quickly growing innovative field. These assessments provide an overview of the compliance of health apps with the established evaluation criteria and can be used from the choice of the appropriate solution by users to the decisions of policy makers to include the health app in the reimbursement scheme. The health apps evaluation can provide assurance on quality and reliability for stakeholders and build trust, which is essential for the further adoption of new solutions that have potential to empower citizens in their health and wellness care and thereby support healthcare systems sustainability. In Estonia the process of a nationwide digital health technologies evaluation framework creation was initiated by the EHIF in 2021. The TalTech Health App Evaluator, as an existing resource, has potential for being used within the framework as a tool for the baseline assessment of patient-centred health apps.

In this thesis the author has introduced different international evaluation approaches that have been relevant in the health apps quality evaluation context development. Theoretical evaluation background and practical major examples in this quickly evolving field were presented. The Estonian TalTech Health App Evaluator quality evaluation requirements were compared with the CEN ISO/TS 82304-2:2021 and similarities and differences were investigated by using case study and cross-case synthesis and analysis methods. The CEN ISO/TS 82304-2:2021 was chosen to be the most applicable comparison case, because this international health apps evaluation-focused approach is based on a variety of local, national, and international initiatives, guidelines, and standards; has a worldwide scope; is anticipated to support the international harmonization of health apps quality requirements; and to reduce the creation of new diverse assessment systems in each country.

The evaluation principles in both frameworks, the TalTech Health App Evaluator, and the CEN ISO/TS 82304-2:2021, have similar purposes, technological and end-user

groups focus. However, the evaluation criteria in the TalTech Health App Evaluator needs to be supplemented by health risks, ethics, societal benefit, accessibility, usability, security, and technical robustness categories. Also, the TalTech Health App Evaluator evaluation outcomes presentation may achieve higher transparency and worldwide recognizability while the international health app quality label introduced by the CEN ISO/TS 82304-2:2021 is adopted.

This thesis provides an opportunity to better understand the health apps quality requirements evaluation principles, criteria, and outcomes. The study addresses the gap in existing knowledge by providing an international comparison on the health apps evaluation and introduces new perspectives on the phenomenon that could benefit the Estonian digital health technologies evaluation community, policy makers, and end-users – patients and healthcare professionals. The study can be used for harmonization of health apps evaluation requirements of the TalTech Health App Evaluator based on the international quality evaluation approach for health apps, and for the further development of the Estonian digital health technologies evaluation framework. The study may have in a longer time horizon an impact on wider public and healthcare decision-making while recommendations are adopted by the policy makers; and has potential to support the sustainability of the healthcare system by enhancing the development and uptake of high-quality, reliable, secure, and trusted digital health apps by empowered citizens.

6 Acknowledgements

I would like to give my profound gratitude to my supervisor Priit Kruus for constructive criticism and keeping me on the track during the thesis writing process. Also, I would like to express my gratitude to my initial supervisor Maarja-Liis Elland for bringing me to this topic and for guidance and feedback. Furthermore, I would like to thank Carmen Mäe for valuable assistance and kind support. It has been a privilege to learn from you. In addition, I would like to thank my family and friends, who have supported me throughout my studies and master's thesis process.

References

[1] WHO, “Monitoring and evaluating digital health interventions. A practical guide to conducting research and assessment,” 12 Dec 2016. [Online]. Available: https://www.who.int/publications/i/item/9789241511766. [Accessed 6 Mar 2022].

[2] C. Vis, L. Bührmann, H. Riper and H. Ossebaard, “Health technology assessment frameworks for eHealth: A systematic review,” *International Journal of Technology Assessment in Health Care,* vol. 36, no. 3, pp. 204-216, Jun 2020.

[3] ORCHA, “Digital Health Accreditation in the Nordics: A cross-border initiative for integrated digital health provision,” Dec 2021. [Online]. Available: https://orchahealth.com/digital-health-accreditation-in-the-nordics/. [Accessed 16 Jan 2022].

[4] NICE, “Evidence standards framework for digital health technologies,” Apr 2021. [Online]. Available: https://www.nice.org.uk/about/what-we-do/our-programmes/evidence-standards-framework-for-digital-health-technologies. [Accessed 12 Jan 2022].

[5] S. P. Rowland, J. E. Fitzgerald, T. Holme, J. Powell and A. McGregor, “What is the clinical value of mHealth for patients?,” *npj Digital Medicine,* vol. 3, no. 4, 2020.

[6] P. Llorens-Vernet and J. Miró, “Standards for Mobile Health-Related Apps: Systematic Review and Development of a Guide,” *JMIR Mhealth Uhealth,* vol. 8, no. 3, p. e13057, 3 Mar 2020.

[7] M. H. van Velthoven, J. C. Wyatt, E. Meinert, D. Brindley and G. Wells, “How standards and user involvement can improve app quality: A lifecycle approach,” *International Journal of Medical Informatics,* vol. 118, no. 10, pp. 54-57, Oct 2018.

[8] F. Woulfe, K. P. Fadahunsi, S. Smith, G. B. Chirambo, E. Larsson, P. Henn, M. Mawkin and J. J. O’ Donoghue, “Identification and Evaluation of Methodologies to Assess the Quality of Mobile Health Apps in High-, Low-, and Middle-Income Countries: Rapid Review,” *JMIR Mhealth Uhealth,* vol. 9, no. 10, p. e28384, 12 Oct 2021.

[9] European mHealth Hub, “D2.1 Knowledge Tool 1. Health apps assessment frameworks,” Oct 2021. [Online]. Available: https://mhealth-hub.org/download/d2-1-knowledge-tool-1-health-apps-assessment-frameworks#. [Accessed 19 Apr 2022].

[10] D. Neal, T. Engelsma, J. Tan, M. P. Craven, R. Marcilly, L. Peute and D. Neal, “Limitations of the new ISO standard for health and wellness apps,” *The Lancet Digital Health,* vol. 4, no. 2, pp. E80-E82, 1 Feb 2022.

[11] S. Mathews, M. J. McShea, C. L. Hanley and A. Raviz, “Digital health: a path to validation,” *npj Digital Medicine,* p. 2(1):38, May 2019.

[12] European mHealth Hub, “Assessment Frameworks,” Dec 2021. [Online]. Available: https://mhealth-hub.org/assessment-frameworks. [Accessed 2 Feb 2022].

[13] IQVIA, “Digital Health Trends 2021,” July 2021. [Online]. Available: https://www.iqvia.com/en/insights/the-iqvia-institute/reports/digital-health-trends-2021. [Accessed 11 Oct 2021].

[14] Statista, “Global digital health market by major segment 2015-2025,” Feb 2022. [Online]. Available: https://www.statista.com/statistics/387867/value-of-worldwide-digital-health-market-forecast-by-segment/. [Accessed 18 Feb 2022].

[15] M.-L. Elland, K. Kuldkepp, M. Tomson, K. Lubi, P. Ross and T. Ross, “Digitaalsete tervisetehnoloogiate hindamisraamistiku töötoad: raport,” Jan 2022. [Online]. Available: https://haigekassa.ee/haigekassa/eesti-haigekassa/uuringud-ja-analuusid. [Accessed 5 May 2022].

[16] TalTech, “Terviserakenduste hindaja,” 2022. [Online]. Available: https://taltech.ee/et/terviserakendused. [Accessed 12 Feb 2022].

[17] ISO, “ISO/TS 82304-2:2021 Health software — Part 2: Health and wellness apps — Quality and reliability,” July 2021. [Online]. Available: https://www.iso.org/standard/78182.html. [Accessed 15 Sep 2021].

[18] D. L. Stufflebeam and C. L. S. Coryn, Evaluation Theory, Models, and Applications, San Francisco: John Wiley & Sons, 2014.

[19] WHO, “WHO evaluation practice handbook,” 2013. [Online]. Available: https://apps.who.int/iris/handle/10665/96311. [Accessed 28 Mar 2022].

[20] HAS - Haute Autorité de santé, “Haute Autorité de santé,” June 2021. [Online]. Available: https://www.has-sante.fr/upload/docs/application/pdf/2021-09/assessment_of_apps_in_the_mobile_health_mhealth_sector._overview_and_quality_criteria_of_medical_content.pdf. [Accessed 12 Oct 2021].

[21] C. A. Reeves and D. A. Bednar, “Defining Quality: Alternatives and Implications,” July 1994. [Online]. Available: https://doi.org/10.2307/258934. [Accessed 28 March 2022].

[22] WHO, “Digital health: transforming and extending the delivery of health services,” Sep 2020. [Online]. Available: https://www.euro.who.int/en/health-topics/health-policy/european-programme-of-work/news/news/2020/09/digital-health-transforming-and-extending-the-delivery-of-health-services. [Accessed 14 Oct 2021].

[23] ISO, “One step ahead with mobile apps,” July 2020. [Online]. Available: https://www.iso.org/news/isofocus_141-6.html. [Accessed 12 Nov 2021].

[24] T. L. Lewis, M. A. Boissaud-Cooke, T. D. Aungst and G. Eysenbach, “Consensus on Use of the Term “App” Versus “Application” for Reporting of mHealth Research,” *Journal of Medical Internet Search,* vol. 16, no. 7, p. e174, July 2014.

[25] EC, “Medical Device Regulation,” Apr 2020. [Online]. Available: https://eur-lex.europa.eu/eli/reg/2017/745/2020-04-24. [Accessed 7 Apr 2022].

[26] T. Ricker, “Jobs: App Store launching with 500 iPhone applications, 25% free,” July 2018. [Online]. Available: https://www.engadget.com/2008-07-10-jobs-app-store-launching-with-500-iphone-applications-25-free.html?guce_referrer=aHR0cHM6Ly9lbi53aWtpcGVkaWEub3JnLw&guce_referrer_sig=AQAAAETv6ntI0KHgUdHQOC7rriutdZatQXD6OrdueeU9pG6PgDzr0RE5uKcuOEeZlo5Y85DmMyOxuIrTyDeFA. [Accessed 17 Mar 2022].

[27] B. Rohleder, “Digital Health,” July 2020. [Online]. Available: https://www.bitkom.org/sites/default/files/2020-07/prasentation_digitalhealth2020.pdf. [Accessed 29 Mar 2022].

[28] A. van Drongelen, A. de Bruijn, B. Roszek and R. Vonk, “Apps under the medical device legislation,” 1 Feb 2019. [Online]. Available: https://www.rivm.nl/publicaties/apps-under-medical-devices-legislation-apps-onder-medische-hulpmiddelen-wetgeving. [Accessed 19 Apr 2022].

[29] T. L. Lewis and J. C. Wyatt, “mHealth and Mobile Medical Apps: A Framework to Assess Risk and Promote Safer Use,” *Journal of Medical Internet Research,* vol. 16, no. 9, p. e210, Sep 2014.

[30] J. Haverinen, N. Keränen, P. Falkenbach, A. Maijala, T. Kolehmainen and J. Reponen, “Digi-HTA: Health technology assessment framework for digital healthcare services,” *Finnish Journal of EHealth and EWelfare,* vol. 11, no. 4, p. 326–341, Nov 2019.

[31] M. S. Marcolino, J. A. Q. Oliveira, M. D'Agostino, A. L. Ribeiro and M. B. M. Alkmim, “The Impact of mHealth Interventions: Systematic Review of Systematic Reviews,” *JMIR Mhealth Uhealth,* vol. 6, no. 1, p. e23, 17 Jan 2018.

[32] R. Nouri, S. R. N. Kalhori, M. Ghazisaeedi, G. Marchand and M. Yasini, “Criteria for assessing the quality of mHealth apps: a systematic review,” *Journal of the American Medical Informatics Association,* vol. 25, no. 8, p. 1089–1098, Aug 2018.

[33] C. Wild and I. Zechmeister-Koss, “Editorial: health technology assessment in Austria,” *Wiener Medizinische Wochenschrift,* vol. 169169, p. 253, 28 Aug 2019.

[34] N. Daniels, T. Porteny and J. Urritia, “Expanded HTA: Enhancing Fairness and Legitimacy,” *International Journal of Health Policy and Management,* vol. 5, no. 1, pp. 1-3, Jan 2016.

[35] F. B. Kristensen, K. Lampe, C. Wild, M. Cerbo, W. Goettsch and L. Becla, “The HTA Core Model - 10 Years of Developing an International Framework to Share Multidimensional Value Assessment,” *Value in Health,* vol. 20, no. 2, pp. 244-250, Feb 2017.

[36] EUnetHTA, “HTA Core Model®,” 2021. [Online]. Available: https://www.eunethta.eu/hta-core-model/. [Accessed 7 Apr 2022].

[37] EC, “Regulation (EU) 2021/2282 of the European Parliament and of the Council of 15 December 2021 on health technology assessment and amending Directive 2011/24/EU (Text with EEA relevance),” Dec 2021. [Online]. Available: http://data.europa.eu/eli/reg/2021/2282/oj. [Accessed 5 Apr 2022].

[38] A. von Huben, M. Howell, K. Howard, J. Carrello and S. Norris, “Health technology assessment for digital technologies that manage chronic disease: a systematic review,” *International Journal of Technology Assessment in Health Care,* vol. 37, no. 66, pp. 1-14, May 2021.

[39] R. Jeindl and C. Wild, “Framework zur Unterstützung von Refundierungsentscheidungen zu digitalen Gesundheitsanwendungen (mHealth) und dessen (retrospektive) Anwendung an ausgewählten Beispielen. AIHTA Projektbericht Nr.: 134,” 2020. [Online]. Available: https://eprints.aihta.at/1279/1/HTA-Projektbericht_Nr.134.pdf. [Accessed 29 Mar 2022].

[40] A. von Huben, M. Howell, J. Carrello, S. Norris, S. Wortley, A. Ritchie and K. Howard, “Application of a health technology assessment framework to digital health technologies that manage chronic disease: a systematic review,” *International Journal of Technology Assessment in Health Care,* vol. 38, no. 1, pp. 1-10, Dec 2021.

[41] EC, “Regulation (EU) 2016/679 of the European Parliament and of the Council,” Apr 2016. [Online]. Available: http://data.europa.eu/eli/reg/2016/679/2016-05-04. [Accessed

7 Apr 2022].

[42] EC, “Market study on telemedicine. European Union,” Oct 2018. [Online]. Available: https://ec.europa.eu/health/sites/default/files/ehealth/docs/2018_provision_marketstudy_telemedicine_en.pdf. [Accessed 19 Oct 2021].

[43] NHS, “Digital Technology Assessment Criteria (DTAC),” 2022. [Online]. Available: https://www.nhsx.nhs.uk/key-tools-and-info/digital-technology-assessment-criteria-dtac/. [Accessed 12 Jan 2002].

[44] Federal Institute for Drugs and Medical Devices, “DiGA Digital Health Application,” 2022. [Online]. Available: https://www.bfarm.de/EN/Medical-devices/Tasks/Digital-Health-Applications/_node.html. [Accessed 12 Jan 2022].

[45] Bundesinstitut für Arzneimittel und Medizinprodukte, “DiGA-Verzeichnis,” 12 Jan 2022. [Online]. Available: https://diga.bfarm.de/de. [Accessed 18 Oct 2021].

[46] mHealth Belgium, “mHealth Belgium,” 2020. [Online]. Available: https://mhealthbelgium.be. [Accessed 12 Jan 2022].

[47] mHealth Belgium, “Validation pyramid,” 2020. [Online]. Available: https://mhealthbelgium.be/validation-pyramid. [Accessed 12 Jan 2022].

[48] University of Oulu, “Centre for Health and Technology,” Jan 2022. [Online]. Available: https://www.oulu.fi/cht/node/194742. [Accessed 18 Feb 2022].

[49] FinCCHTA, “Digi-HTA,” 2022. [Online]. Available: https://www.ppshp.fi/Tutkimus-ja-opetus/FinCCHTA/Sivut/Digi-HTA.aspx. [Accessed 14 Jan 2022].

[50] ISO, “Standards,” 2021. [Online]. Available: https://www.iso.org/standards.html. [Accessed 12 Nov 2021].

[51] CEN-CENLEC, “Frequently Asked Questions (FAQs),” 2022. [Online]. Available: https://www.cencenelec.eu/bottom-navigation-pages/faqs/. [Accessed 4 March 2022].

[52] CEN-CENLEC, “CEN-CENLEC,” 2022. [Online]. Available: https://www.cencenelec.eu/european-standardization/european-standards/. [Accessed 4 March 2022].

[53] ISO, “Certification & Conformity,” 2022. [Online]. Available: https://www.iso.org/certification.html. [Accessed 4 March 2022].

[54] Reciprocity, “Guide to ISO Certification and ISO Compliance,” October 2019. [Online]. Available: https://reciprocity.com/guide-to-iso-certification-and-iso-compliance/. [Accessed 6 March 2022].

[55] eHealth Suisse, “Guide for app developers, manufacturers and distributors,” June 2020. [Online]. Available: https://www.e-health-suisse.ch/fileadmin/user_upload/Dokumente/2018/E/180731_Leitfaden_fuer_App_Entwickler_def_EN.pdf. [Accessed 6 Mar 2022].

[56] ISO, “IEC 62304:2006/AMD 1:2015 Medical device software — Software life cycle processes — Amendment 1,” Nov 2017. [Online]. Available: https://www.iso.org/standard/64686.html. [Accessed 4 March 2022].

[57] ISO, “IEC 82304-1:2016 Health software,” Oct 2016. [Online]. Available: https://www.iso.org/standard/59543.html. [Accessed 8 Oct 2021].

[58] ORCHA, “ORCHA,” 2022. [Online]. Available: https://orchahealth.com. [Accessed 14 Jan 2022].

[59] ORCHA, “How do Digital Health Standards Assess Evidence?,” Jul 2021. [Online].

Available: https://orchahealth.com/how-do-digital-health-standards-assess-evidence/. [Accessed 15 Nov 2021].

[60] BSI, “Code of Practice for wellness apps published,” May 2015. [Online]. Available: https://www.bsigroup.com/en-GB/about-bsi/media-centre/press-releases/2015/may/Code-of-Practice-for-wellness-apps-published/. [Accessed 12 March 2022].

[61] HIMSS Analytics, “eHealth Trendbarometer,” Nov 2019. [Online]. Available: https://europe.himssanalytics.org/sites/himssanalytics_europe/files/eHealth%20TRENDBAROMETER%20-%20HIMSS%20Analytics%20Annual%20European%20eHealth%20Survey%202019.pdf. [Accessed 3 Nov 2021].

[62] University of Tartu, “Health Technology Assessment in Estonia,” Feb 2022. [Online]. Available: https://tervis.ut.ee/en/health-technology-assessment/health-technology-assessment-estonia. [Accessed 17 Oct 2021].

[63] Eesti Haigekassa, “Eesti Haigekassa Arengukava 2020 – 2023,” 2020. [Online]. Available: https://www.haigekassa.ee/sites/default/files/uuringud_aruanded/EHK_arengukava_2020_2023.pdf. [Accessed 4 Nov 2021].

[64] Eesti Haigekassa, “Digitaalsete tervisetehnoloogiate hindamisraamistik,” Nov 2020. [Online]. Available: https://haigekassa.ee/sites/default/files/Digitaalsete%20tervisetehnoloogiate%20hindamisraamistik_EHK_202011.pdf. [Accessed 15 Nov 2021].

[65] TalTech, “Taltechi tervise nutirakenduste avalik hindaja annab kasutajale erapooletu ja süsteemse ülevaate,” Mar 2021. [Online]. Available: https://taltech.ee/uudised/taltechi-tervise-nutirakenduste-avalik-hindaja-annab-kasutajale-erapooletu-ja-susteemse. [Accessed 14 Sep 2021].

[66] R. K. Yin, Case study research: Design and methods, London: Sage, 2009.

[67] M. B. Miles and A. M. Huberman, Qualitative data analysis, housand Oaks, CA: Sage, 1994.

[68] A. J. Mills, G. Durepos and E. Wiebe, Encyclopedia of case study research (Vols. 1-0). Cross-Case Synthesis and Analysis, Thousand Oaks, CA: SAGE Publications, Inc, 2010.

[69] J. Gerring, Case study research: principles and practices, Cambridge: Cambridge University Press, 2017.

[70] M. Q. Patton, Qualitative evaluation and research methods, Beverly Hills, CA: Sage, 1990.

[71] H.-G. Ridder, “The theory contribution of case study research designs,” *Businesss Research,* vol. 10, p. 281–305, 16 Feb 2017.

[72] L. M. Given, , The Sage Encyclopedia of Qualitative Research Methods: A-L ; Vol. 2, M-Z Index, SAGE, 2008.

[73] CEN/TC 251 Health Informatics, “New international guidelines to help to sort the best from the rest,” Sep 2021. [Online]. Available: https://www.ehealth-standards.eu/2021/09/01/health-and-wellness-apps-new-international-guidelines-to-help-to-sort-the-best-from-the-rest/. [Accessed 17 Feb 2022].

[74] The World Wide Web Consortium (W3C), “Web Content Accessibility Guidelines (WCAG) 2.1,” Jun 2018. [Online]. Available: https://www.w3.org/TR/WCAG21/.

[Accessed 18 Feb 2022].

[75] ISO, “ISO 9241-11:2018 Ergonomics of human-system interaction — Part 11: Usability: Definitions and concepts,” Mar 2018. [Online]. Available: https://www.iso.org/standard/63500.html. [Accessed 18 Apr 2022].

[76] ISO, “ISO 9241-210:2019 Ergonomics of human-system interaction — Part 210: Human-centred design for interactive systems,” July 2019. [Online]. Available: https://www.iso.org/standard/77520.html. [Accessed 18 Apr 2022].

[77] ISO, “ISO/TS 27790:2009 Health informatics — Document registry framework,” Dec 2019. [Online]. Available: https://www.iso.org/standard/44316.html. [Accessed 18 Apr 2022].

[78] The Institute of Electrical and Electronics Engineers, IEEE standard computer dictionary: a compilation of IEEE standard computer glossaries, New York: The Institute of Electrical and Electronics Engineers, 1990.

Appendix 1 – Non-exclusive licence for reproduction and publication of a graduation thesis[1]

I Marju Tomson

1. Grant Tallinn University of Technology free licence (non-exclusive licence) for my thesis "[Thesis title]" , supervised by Priit Kruus
   1.1. to be reproduced for the purposes of preservation and electronic publication of the graduation thesis, incl. to be entered in the digital collection of the library of Tallinn University of Technology until expiry of the term of copyright;
   1.2. to be published via the web of Tallinn University of Technology, incl. to be entered in the digital collection of the library of Tallinn University of Technology until expiry of the term of copyright.
2. I am aware that the author also retains the rights specified in clause 1 of the non-exclusive licence.
3. I confirm that granting the non-exclusive licence does not infringe other persons' intellectual property rights, the rights arising from the Personal Data Protection Act or rights arising from other legislation.

09.05.2022

1 The non-exclusive licence is not valid during the validity of access restriction indicated in the student's application for restriction on access to the graduation thesis that has been signed by the school's dean, except in case of the university's right to reproduce the thesis for preservation purposes only. If a graduation thesis is based on the joint creative activity of two or more persons and the co-author(s) has/have not granted, by the set deadline, the student defending his/her graduation thesis consent to reproduce and publish the graduation thesis in compliance with clauses 1.1 and 1.2 of the non-exclusive licence, the non-exclusive license shall not be valid for the period.