Using Thematic Analysis in Healthcare HCI at CHI: A Scoping Review

Robert Bowman
Trinity College Dublin
Dublin, Ireland
rbowman@tcd.ie

Camille Nadal
Trinity College Dublin
Dublin, Ireland
nadalc@tcd.ie

Kellie Morrissey
University of Limerick
Limerick, Ireland
kellie.morrissey@ul.ie

Anja Thieme
Microsoft Research
Cambridge, United Kingdom
anthie@microsoft.com

Gavin Doherty
Trinity College Dublin
Dublin, Ireland
gavin.doherty@tcd.ie

ABSTRACT

CHI papers researching healthcare human-computer interaction (HCI) are increasingly reporting the use of "thematic analysis" (TA). TA refers to a range of flexible and evolving approaches for qualitative data analysis. Its increased use demonstrates a change in research practices, and with that the emergence of new local standards. We need to understand and reflect upon these emerging local practices, including departures from what is advocated as quality TA practice more generally. Toward this, we conducted a scoping review of a decade of CHI publications (2012 – 2021) that researched healthcare and termed their analysis approach "thematic analysis"; 78 papers reporting a total of 100 TAs were included. We contribute a description of 1) the contexts in which TA is being used, 2) the TA approaches being conducted, and 3) how TA is being reported. Drawing on this, we discuss opportunities to improve research practice when using TA in healthcare HCI.

CCS CONCEPTS

• **Human-centered computing → HCI design and evaluation methods**.

KEYWORDS

thematic analysis, qualitative research, healthcare

1 INTRODUCTION

Qualitative research serves to help us understand people's experiences, including those with technology. In human-computer interaction (HCI) we draw on this understanding to inform the design

of future technology [9]. Critically, this knowledge production is influenced by the position of the researcher(s). For example, when analysing data, a researcher's own life experience, as well as their prevailing beliefs and attitudes, can influence their interpretation of the data [17]. This is particularly pertinent within healthcare HCI, as people's experiences of health and healthcare vary hugely. Most HCI papers dealing with healthcare also impact on some level upon participants who are made vulnerable by periods of ill health, who care for others who are ill, or who are otherwise concerned about their own health. In turn, as healthcare HCI researchers, we attempt to capture, understand, analyse and re-represent these participants' lived experiences in a way which helps to direct the future design of technologies, which in turn may further impact their lives. The opportunity for our research to be successful — whereby it positively influences people's experiences of health and healthcare — is impacted by the capabilities of our analysis methods. We need methods that allow us to effectively understand and represent the experiences of our participants, as well as to develop and communicate design insight from them.

While a range of qualitative methods have been applied in HCI for healthcare, in recent years the number of CHI[1] papers focusing on the design and evaluation of healthcare technologies while reporting the use of "thematic analysis" (TA) for qualitative data analysis has grown rapidly (Figure 1). This increased adoption signifies an important change in healthcare HCI qualitative research practices. Through the publication of these works, the community has constructed local standards[2] for the use of TA in healthcare HCI at CHI which in turn influence the conduct and publication of future research. However, local standards, or norms, are not necessarily good practice and can emerge without deep discussion or shared reflection [80]. As a community, we have a responsibility to play an active role in the development of these norms. That is to say, we should strive to shape the norms, rather than let the norms shape us.

Seeking to understand the emergent norms of thematic analysis in healthcare HCI at CHI, and from this to identify opportunities to improve TA practice going forward, we conducted a scoping review. Our review covered a decade of CHI publications (2012 – 2021) including papers that researched healthcare and termed their

---

[1] ACM CHI Conference on Human Factors in Computing Systems
[2] "Local standards are guidelines based on similar or analogous studies that have already been published" [26, p. 983]
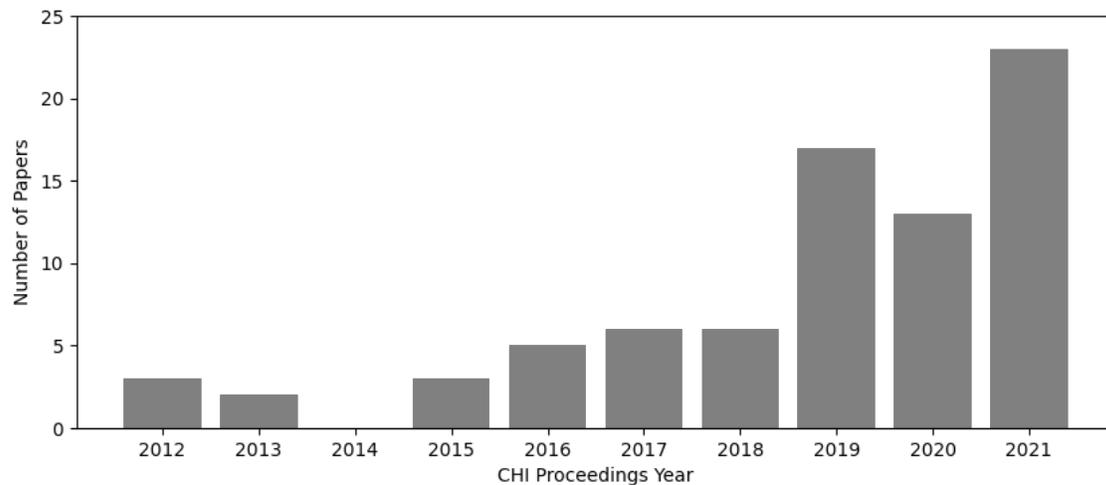
Figure 1: CHI healthcare papers reporting the use of "thematic analysis".

analysis approach "thematic analysis". 78 papers were included that reported a total of 100 TAs, which we reviewed individually. The term "thematic analysis" is used to describe, not one, but a range of analysis approaches; these approaches can differ considerably by their practices and underlying philosophies [20]. Through this review of 10 years of healthcare publications at CHI, we present a rich description of the use of thematic analysis at CHI, in the particular context of healthcare, where researchers are dealing with sensitive data and vulnerable populations. Specifically, we analysed 1) the contexts in which TA is being used, 2) the TA approaches being conducted, and 3) how TA is being reported.

Our review demonstrates that TA is often used to analyze interview data and to generate design insights. However, we observe that descriptions of TA are frequently light in explanation, or imprecise. This is a major cause for concern as it leaves it uncertain how the data was analyzed, which inhibits the reader's ability to evaluate the validity of the research. It also sets precedent for future work to report light or imprecise descriptions thereby inhibiting the communities ability to develop an understanding of how different TA approaches and practices can benefit our research. We also notice that the reflexive TA approach is prevalent, cited in 59% of papers, yet we found little reporting of researchers' positions or reflexive practices. Finally, TA is often framed as team activity, however it is often unclear how collaboration between researchers happens.

From this description, we identify opportunities for the CHI community to improve its practices going forward:

Reporting TA We argue for the importance of reporting detailed description of TAs, and that such rich descriptions are opportunities to identify and develop TA approaches that more effectively meet our research needs, particularly in the domain of healthcare HCI.

Reflexivity & Positionality We show that, while it is essential that researchers' reflect on their position when conducting and reporting TA in healthcare, this is difficult to put into practice. We offer some starting directions for researchers in healthcare HCI.

Using TA as a Team We observe that the healthcare CHI community is more often than not using TA as a team, yet how such collaboration takes place is unclear. To inform future team-based TA, we contribute a discussion of the opportunities and challenges of different collaboration approaches.

Research Fit We discuss the research fit of TA for healthcare HCI research, considering whether alternative approaches to, and instead of, TA may better address our research needs and the development of design implications.

The description we contribute affords our community an opportunity to reflect upon and identify ways to improve our TA practices, a process we initiate through our discussion; as such, our work extends the developing discourse around the use of qualitative methods in HCI [50, 53, 80].

2 BACKGROUND

2.1 Qualitative Research in Healthcare HCI

> "With qualitative research, the emphasis is not on measuring and producing numbers but instead on understanding the qualities of a particular technology and how people use it in their lives, how they think about it and how they feel about it" [2, p. 138]

Qualitative analysis is commonly carried out through techniques that aid researchers to identify or construct meanings or patterns from a given dataset. As emphasized by Blandford et al. [9, p.1]:

> "Qualitative methods play an important role in Human–Computer Interaction (HCI): in requirements gathering, in acquiring an understanding of the situations in which technology is used and might be used and in evaluating how technologies are used in practice."

Indeed, it was identified that the primary analysis method of over a third of CHI papers from 2016 − 2018 was qualitative [80]. Qualitative methods are of huge importance to research on health and

healthcare, offering ways of understanding people's lived experiences [18]. With technologies for prevention, diagnosis, treatment, and monitoring greatly influencing healthcare delivery [104], qualitative HCI research is needed to understand people's experiences of healthcare.

2.1.1 Study context matters: Individuals' – researchers and participants – experiences of healthcare can be significantly impacted by characteristics (including geographical background, ethnicity, class, and gender) due to the globally unequal nature of healthcare. Someone attempting to access healthcare in rural Ghana is liable to face different barriers than someone in a metropolitan area in Japan. African Americans face significant barriers as well as systematic racism in accessing care across multiple sectors [32]. Women experiencing chronic pain and illness are labeled 'emotional' while their male counterparts are described as 'brave' [103].

2.1.2 Researcher positionality matters: When performing qualitative research, a researcher's interpretation of the data is influenced by their own individual life experiences, as well as their knowledge, beliefs and attitudes [11, 38, 52, 58]. A relevant example of this comes from two researchers who were analysing focus group data regarding peoples' experiences with physical activity [17, p. 205]. While one researcher noticed that participants framed physical activity as a chore with no pleasure, the other did not notice this negativity. The researchers related this difference in interpretation to the difference between their own experiences of physical activity and levels of enthusiasm for it. A concept used to describe this relation between the researcher(s) and the research is *positionality*. As Holmes [62, p.3] put it:

> "Positionality implies that the social-historical-political location of a researcher influences their orientations, i.e., that they are not separate from the social processes they study."

Ames et al. [4], who explore technology use among participants of different class backgrounds, write that, without reflecting on our positionality:

> "many of us find ourselves designing for those most like ourselves … or we may make incorrect assumptions based on stereotypes about those who are different." [4, p. 64]

In the design of healthcare products and systems, these stereotypes may have very real, and very harmful effects if they persist into the implementation of technology design itself. Examples of discriminatory design consequences in healthcare HCI, as cited by Nadal [86], include misdiagnosed users [117], no access to treatment [100], inaccessible technology [6], under-represented groups of users [43], negated users' identities [83], and the perpetuation of stereotypes [115].

2.1.3 Trends to impose quantitative paradigms onto qualitative research: Researchers hold a range of perspectives about the role of subjectivity within qualitative research. These perspectives stem from researcher's beliefs about what is real (i.e., ontology) and what are valid ways of generating knowledge (i.e., epistemology)[3]. A simple and useful distinction is that of small-q and big-Q [69]

---

[3]See an introductory qualitative research textbook (e.g., [17, 22, 63]) for further detail.

which has been adopted within prominent thematic analysis literature (e.g., [17, 22]). *Small q* qualitative research is where qualitative research is conducted within a *quantitative* paradigm. It concerns itself with achieving an objective and unbiased analysis of the data (i.e., it seeks to minimise the impact of researcher subjectivity). *Big Q* qualitative research is where qualitative research is conducted within a *qualitative* paradigm, it is some times referred to as "fully qualitative". It values researcher subjectivity and views it as fundamental part of knowledge construction [17, 22]. Although qualitative research is an accepted form of HCI research, the concerns of the historically dominant quantitative paradigm remain influential. A relevant occurrence here is *positivism creep* whereby the thinking of the quantitative paradigm (i.e., positivism), for instance the need to control bias, is imposed upon qualitative research pulling it into the small q paradigm [17, 22]. For example, qualitative researchers may feel compelled to use reliability techniques (e.g., multiple people code the data toward developing a reproducible analysis) in-order to demonstrate the method is rigorous [50, 80]. This is of concern to healthcare HCI as prioritising reliability risks the marginalization or minimisation of perspectives [80].

2.1.4 Reflexivity as active interrogation of how research impacts the researched: An important concept for biq Q research is *reflexivity* which Campbell et al. [27, p. 2016] describes as an:

> "ongoing activity to situate the researcher within the analytic process including acknowledgment of social locations and positionalities, such as age, gender identification, ethnicity, and race"

Reflexivity is going beyond acknowledging that our positionality affects our research to actively engaging with how it is affecting the research as part of the research process. A recommended technique for supporting reflexivity is keeping a journal in which the researcher reflects upon their experiences of performing the research in part to interrogate how these are impacted by their positionality [22, 34].

2.1.5 Local standards, not good practice, leading the way: As healthcare HCI researchers we have a responsibility to pay close attention to our research methods to ensure that the analytic process best serves these participants when the outcome, at some level, is the generation of new technology. Through the peer-review process, the CHI community constructs standards of what is appropriate and rigorous by setting precedents for future CHI publications. These *local standards* are represented by the research which is accepted for publication and influence the acceptance and publication of future research [26]. For instance, the ways in which qualitative methodology is reported in one published paper may become the justification for future papers to reproduce this reporting of methodology. To our collective detriment, local standards can emerge that are not 'good' practice [80] thus it is important that we are critical in our acceptance of them. To examine local standards, reviews are commonly conducted. For example, Caine [26] reviewed CHI sample sizes, Linxen et al. [77] reviewed how WEIRD (Western, Educated, Industrialised, Rich, and Democratic) CHI research participants are, and Abbott et al. [1] reviewed anonymization practices in health, wellness, accessibility, and aging research at CHI. More relevant to

qualitative *analysis*, McDonald et al. [80] reviewed the use of reliability and inter-rater reliability in HCI and CSCW[4]. The authors show that:

> "CSCW and HCI qualitative researchers use the same terms and concepts in multiple, complex ways, and that readers and authors themselves may have little consensus about what was done and why" [80, p. 2]

They continue on to emphasize the importance of developing a shared understanding of how qualitative research should be written about and evaluated. Fiesler et al. [50] echo these concerns within a discussion of the contemporary challenges and opportunities for qualitative research in CSCW. A particular difficulty faced by researchers wanting to conduct qualitative research within healthcare HCI is that the research goals differ from those of the social sciences — which largely influenced the development of qualitative methods. Unlike social sciences, whose focus is understanding social phenomena, the focus of HCI is typically to contribute to the design of technology [9].

2.2 Thematic Analysis

As illustrated by Figure 1, in healthcare HCI at CHI, the number of papers reporting the use of "thematic analysis" has grown rapidly. Although thematic analysis is often presented a single method, the term – thematic analysis – is understood to refer to a range of approaches for qualitative analysis [22]. The origin of these approaches is somewhat unclear, however a reasonable explanation is that they largely developed out of qualitative refinements of content analysis [22, 63] which is a method that seeks to quantitatively describe content [2]. These TA approaches all seek to interpret data by identifying themes using coding techniques, yet the procedures and theoretical underpinnings of these approaches can differ significantly [20].

In a 2006 article, Braun and Clarke [15, p. 77] argued that:

> "Thematic analysis is a poorly demarcated, rarely acknowledged, yet widely used qualitative analytic method"

Within this article [15], the authors presented their approach for TA which was later branded as *reflexive TA* [19]. The 2006 article proved seminal. Following the article's publication, many works started citing the article and claiming to use its approach. At the time of writing the article has has garnered over 130,000 citations. Despite the huge adoption of TA, it is argued that TA remains unclearly demarcated in part due to the differences between approaches being poorly understood [22].

It has been conceptualized that there are three main types of TA [20, 22]:

- Coding Reliability TA,
- Codebook TA,
- and Reflexive TA.

A critical way in which these differ is their position within the qualitative and quantitative research paradigms. *Coding reliability TA* are a set of small q approaches concerned with objective and reliable data coding; multiple people code the data and their codings are then numerically compared to assess reliability, before

[4]Computer-Supported Cooperative Work.

identifying themes. *Reflexive TA* is of a big Q position that values researcher subjectivity, and considers researchers to construct themes. Finally, sometimes associated with 'medium q', *Codebook TA* approaches refer to a set of methods that combine qualitative values with more structured processes, they differ from reflexive TA in using structured frameworks for coding and early theme development [22].

As this paper will demonstrate, the reflexive TA approach is highly influential within healthcare HCI at CHI, and there may be good reason for this. Reflexive TA suits applied research as its outputs can be accessible for people outside of academia, and works well for research teams that vary in qualitative research experience [18]. Furthermore, the reflexive TA approach encourages researchers to reflect on their position and decisions, what they might enable or exclude, as "their disciplinary, theoretical and personal assumptions and their design choices shape and delimit the knowledge they produce" [22, p. 294]. Given the importance of researcher positionality and reflexivity in healthcare HCI, this provides an argument for reflexive TA being a well suited approach.

In the 2006 article [15], the approach was described as a 6-phase process. Since the article's publication, the authors' thinking about TA has developed and so have some of the phase names. In recent work [22] the phases are:

(1) Familiarizing yourself with the dataset
(2) Coding
(3) Generating initial themes
(4) Developing and reviewing themes
(5) Refining, defining and naming themes
(6) Writing up

The approach is described as flexible. It can have a more inductive (bottom-up) or deductive (top-down) orientation to the data. It can also focus on surface level (i.e., semantic) or underlying (i.e., latent) meaning, as well as have more experiential or critical aims. As such there are a number of decisions that must be made when using reflexive TA which affect its research fit [22]. For an accessible and detailed introduction to reflexive TA (as well as TA more broadly) see Braun and Clarke's recent textbook *Thematic Analysis: A Practical Guide* [22].

The application of reflexive TA is contested in broader qualitative research, most of all by Braun and Clarke themselves. For instance, the authors recently published an article [21] discussing 10 'problematic' practices they observed in published research, and presenting a 20-question guide to evaluate research using TA. The practices in question include 1) assuming TA is one single approach, 2) assuming that TA is atheoretical, 3) combining TA with other approaches in incompatible ways, as well as 4) citing their 2006 article [15] without reading or following it.

3 RESEARCH AIMS

Qualitative research is crucial for understanding peoples' experiences with, and designing, healthcare technology. The number of healthcare HCI papers published at CHI reporting the use of "thematic analysis" has increased rapidly demonstrating an important change in the community's qualitative research practices. Through this scoping review, we aim to map the emergent norms of TA

within healthcare HCI at CHI. We intend for this mapping to contribute an understanding of:

(1) what thematic analysis is being used for,
(2) how research using thematic analysis is designed and reported.

Our aspiration is to then identify ways the CHI healthcare HCI community can improve its use of thematic analysis going forward. With this paper, we specifically aim to contribute guidance for researchers and reviewers, by building on the understanding of how thematic analysis is used, to provide recommendations for research design and reporting. Accordingly, within the context of healthcare HCI at CHI, our research questions are the following:

RQ 1 What is thematic analysis being used for?
- i.e., what is being studied, what data is being analyzed, what is the output used for?

RQ 2 How is thematic analysis being conducted?
- e.g., what publications are cited for method, how is data coded, how themes are constructed, how do researchers collaborate?

RQ 3 How is thematic analysis being reported?
- e.g., is the method specified and justified, what aspects of the method are detailed, is the researcher's position considered?

RQ 4 What areas for improvement exist in the design and reporting of research involving thematic analysis?

RQ 5 How should the peer-review of thematic analysis be approached?

As this research is a literature review, our analysis will rely on what was reported; consequently, our analysis in relation to RQ2 will be indirect and limited by the reporting of the included studies, thus RQ2 and RQ3 are entangled. We choose to maintain the distinction between RQ2 and RQ3 to emphasize our interest in both the method used and the reporting of the analysis.

4 METHOD

To address our research aims, we chose to use a scoping review method. A scoping review is a rigorous and transparent literature review approach that serves to provide a map, or overview, of a topic. An established purpose of scoping reviews is to examine research practices [84], thus the method aligns well with our research aims. As with similar scoping reviews conducted in HCI (e.g., [90]), our method was informed by the established scoping review framework of Arksey and O'Malley [5] and the PRISMA-ScR reporting guidelines [118] based upon it. Our work has also been guided by the study design and guidance derived by McDonald et al. [80], who studied reliability and inter-rater reliability in qualitative research.

4.1 Protocol Design and Development

To develop our review protocol, we conducted preliminary searches of the Association of Computer Machinery Digital Library (ACM-DL) to assess the feasibility of the review under different eligibility criteria. Additionally, we piloted a data extraction procedure. Authors 1 - 3 each extracted data from 6 different papers. Through

analysis of these pilot papers and all-author discussion, we developed a initial protocol [13].

We wanted to distribute the charting between authors 1 - 3 to give us the capacity to chart more TAs. To ensure our interpretations would be sufficiently aligned, we conducted further piloting to test and develop our agreement [80]. To this end, each of the three authors independently extracted the same 5 papers — we then compared and discussed the extractions. The included papers used and reported TAs in ways we had not anticipated or encountered during our previous pilot. We adapted our extraction process to account for such papers and repeated the process with a further 5 papers. Through comparison and discussion, we identified a few ways to further improve extraction, and concluded that there was sufficient agreement among us to proceed to extracting the entire corpus. At this point, we registered an updated protocol with the Open Science Foundation to reflect its development [12].

4.2 Search

Our aim was to examine the local standards for TA use in healthcare HCI at CHI that have emerged from its recent rapid adoption. Accordingly, we searched the previous 5 proceedings of the ACM Conference on Human Factors in Computing Systems (CHI) for papers on the topic of health or healthcare and that termed their analysis approach "thematic analysis". We conducted this search on April 8th 2022, which therefore included the proceedings for the years 2017 – 2021. We then extracted and analysed data from this corpus. Wishing to increase the scope of our review, we extended the search back 5 proceedings (2012 – 2016); we conducted the second search on November 14th 2022. Our combined search included a decade of CHI proceedings (2012 – 2021).

4.2.1 Eligibility. Studies were included if they satisfied the following eligibility criteria:

EC 1 Uses an analysis method described as "thematic analysis".
EC 2 Focuses on the topic of health or healthcare:
- includes physical and mental health topics,
- excludes wellness and wellbeing topics,
- literature that consider health or healthcare will be excluded if the topic is not an essential part of the work's focus.
EC 3 Is a full length research paper.
- Excludes short form papers (e.g., extended abstracts).
EC 4 Was published in the 2012 – 2021 proceedings of the ACM Conference on Human Factors in Computing Systems (CHI).

4.2.2 Source & Terms. The reviews source was the the ACM Digital Library (ACM DL), the database in which CHI papers are published. Targeting EC 1 and 2, the ACM DL was searched for items that include "*thematic analysis*" in the full paper and *health** in the title or abstract using the query:

```
Fulltext:("thematic analysis") AND
(Title:(health*)) OR Abstract:(health*))
```

The asterisk (*) used for the health term specifies a number of unknown characters, this means the term *health** will match terms including 'healthcare' in addition to 'health'. The first search (April 8th 2022) used the publication date filter "Publication Date: (01/01/2017 TO *)" to include the 2017 – 2021 proceedings. The

second search (November 14th 2022) used the publication data filter "Publication Date: (01/01/2012 TO 12/31/2016)" to include the 2012 – 2016 proceedings.

4.2.3 Screening. The references of the search results were exported as BIB files which were then computationally parsed into CSV files. For EC 4, the CSV files were filtered on the field *booktitle* to include only the articles published in CHI proceedings. Authors 1 and 2 independently screened the remaining articles against the eligibility criteria. There was 84% initial agreement for the first search and 87.5% for the second search. Disagreements (typically regarding if the paper had sufficient healthcare focus) were resolved through discussion between the two authors. Two papers included during the first search were later excluded as on closer reading they did not report a thematic analysis[5].

4.3 Data Extraction

Toward addressing our research questions (Section 3), we sought to extract data using items closely coupled to our first three research questions regarding what TA is used for, how it is conducted, and how it is reported. Critically, our approach — like that of McDonald et al. [80] — was to only code that which was explicit in the paper and to avoid making inferences about what was, or was not, done. Our process for data extraction was iteratively developed starting from recent thematic analysis methods literature and adapted to incorporate features we observed through piloting. Our scoping review protocol [12], which is available on the Open Science Foundation, details our charting factors. We used spreadsheets to record the extracted data. The corpus was split among authors 1-3 who proceeded to extract data. Throughout the extraction, the authors attempted to flag interesting examples, with regard to our full set of research questions, that may otherwise not be captured by our extraction process.

4.4 Data Analysis

The papers included in the first search were analysed prior to the second search being conducted. To map the use, conduct, and reporting of TA, authors 1-3 used different synthesis techniques. For data extracted in a categorical form (e.g., using a tick-box) we directly described it numerically. Much of the extracted data was in the form of text snippets copied from the papers; to summarize these, we coded these snippets inductively to allow us to count the occurrences of characteristics. For most fields, a simple coding involving little interpretation was required (e.g., the method source cited and the tool used); consistent with recent guidelines [80], each of these was coded by a single author.

The text snippets describing coding and theme generation processes were more complex and required a greater level of analysis in order to develop an effective description of this data. For each, one author first inductively developed an initial codebook which took the form of a list of codes, such that for each snippet (i.e., included TA) we marked if the characteristic represented by the code was present. Next, a second author applied the codebook, looking to identify areas for improvement. The two authors then discussed

disagreements and used these to further strengthen the codebook. Once satisfied with the codebook, the two authors applied the it to a fresh copy of the data. As this analysis was more complex and we wanted to report results quantitatively [80], we calculated inter-rater reliability using Cohen's Kappa for each code; the average kappas for the coding and theme generation snippets were 0.81 and 0.87 respectively, indicating a good level of agreement. The two authors then settled the disagreements to reach the final codings.

We coded the second search data using the codebooks we had developed when analysing the data from the first search. We looked for features not captured by the codebook during the coding and created a small number of new codes (e.g., previously unseen method citations). Consistent with the first analysis, a second author coded the data related to papers' reporting of their approach to *coding* and *theme generation*, resulting in only a small number of disagreements.

To report our analysis, we primarily present counts of the number of TAs in which we identified each characteristic or practices. Additionally, we highlight examples that we perceive to be particularly insightful, in an attempt to understand departures from existing guidelines for TA and uncover opportunities to provide more suitable methods.

4.5 Positionality Statement

We were motivated to conduct this research by our own experiences — as well as anecdotal accounts — of TA being used in diverse ways rendering it difficult to navigate using and reviewing TA research. Our aspiration is to contribute an understanding of TA practices that can help the healthcare HCI community to navigate and improve its use and review of TA going forward. To conduct this research, we formed a team whose position varied in ways pertinent to the topic. While we all research healthcare HCI in Europe, we differ in education and career stage. Our educations include studies in computer science, psychology and HCI. We include early career researchers as well as senior researchers working in both academia and industry. These variations contribute to us having different knowledge of and relationship with research methodology. We have sought throughout this scoping review to use the variation in our position as a resource to develop our analysis, by identifying and reflecting on differences in our beliefs and interpretations. During this work, we have held the view that there are good reasons for the different research practices we encounter. Research practices are evolving entities, rather than attempting to hold work accountable to today's documented procedures, we view variations as potential opportunities for the community to develop its practices. [6]

5 ANALYSIS

Analyzing the identified corpus allowed us to map recent HCI research practices involving thematic analysis within healthcare. We describe here the context of these studies, their methods and the TA process they employed.

---

[5]The first stated a TA was first conducted, but an alternative method was then deemed more appropriate and the paper only reported the alternative method. The second described the TA conducted in a different paper.

[6]Positionality statements are not typically included in scoping reviews or suggested by scoping review reporting guidelines [118]. However, given the topic of this review, we consider it of value to include one.

5.1 Search Results

Following the steps described in Section 4.2, our search of the ACM digital library resulted in a corpus of 431 articles. After filtering out papers published in a different venue, we assessed eligibility of the CHI papers (n=104). In most cases, a paper's eligibility couldn't be determined by solely considering its title or abstract, and required looking at the full text. A total of **78 papers** were included in our scoping review, the details of which can be found in the supplementary material. Close to a quarter of the sample (n=18) reported multiple thematic analyses. Therefore, analyzing this corpus of 78 papers equated to reviewing a total of **100 thematic analyses**. Among the papers reporting several analyses, the large majority (n=13) reported two TAs, while a smaller number (n=4) reported three. Mostly, researchers reported conducting multiple TAs to analyze data gathered from different sources —- e.g. focus groups and usage logs [71], or co-design sessions and surveys [120] –– or different populations [124]. On rare occasions, papers reported a sequence of studies [72, 119].

5.2 Research Contexts

5.2.1 Study Characteristics. Healthcare HCI research using TA mainly seeks to *study demographics*, *experiences with technology*, and to *inform technology design* (see Table 1). A range of health and mental health topics are considered, particularly focusing on women's health[7], affective disorders and Diabetes (a more detailed view is given in Appendix 2). Most included studies (papers=53) relate to health. While the large majority of technologies investigated target *populations receiving care*, TA seems to be less employed in research exploring solutions for *health professionals* and *caregivers*. Examples of these include a study exploring a new clinical decision support tool [124], and a study looking at the case of technology supporting dementia caregivers [56]. It is interesting to note that TA is sometimes used (13% of cases) to explore ecologies of care.

Table 1: Characteristics of the studies reviewed.

Characteristic	Number of papers (n=78)
HCI topic	
Understand demographics & behaviors	31
Capture experiences with new technology	24
Capture experiences with existing technology	12
Inform the design process	8
Review existing literature or artifacts	3
Health topic	
Health	53
Mental health	18
Both	7
Persons technology is intended for[8]	
Persons receiving care	55
Health professionals	12
Ecologies	10
Caregivers	5

[7]The term '*women's health*' is used by the papers, but we acknowledge this language is not gender-inclusive.

5.2.2 Location. Among the 78 papers included, 53 explicitly reported a study location. In most cases, data subjected to TA was collected in the USA (papers=21), followed by the UK (papers=11) and India (papers=5), as described by Table 2. While contextual information is important to grasp the different realities of healthcare, almost a third of the corpus did not report the geographical origin of the data analyzed.

Table 2: Reported locations of the studies reviewed.

Countries	Number of Papers
USA	21
UK	11
India	5
Bangladesh	2
China	2
Austria	1
Barbados	1
Belgium	1
Canada	1
Ghana	1
Ireland	1
Kenya	1
Malaysia	1
Pakistan	1
Singapore	1
South Korea	1
Sweden	1

5.2.3 Participants. The majority of analyses reviewed were concerned with *people receiving care*, with 44 TAs focusing solely on their experiences. Next, *health professionals* were addressed by 21 TAs, while 3 TAs solely focused on *caregivers*. Our 'other' category illustrates the diversity of TAs in the sample reviewed, which focused on participants who fell outside of these categories. Some TAs involved people who cannot be easily categorized as 'receiving care' (e.g., members of an American Black Church [94]), or as 'health professionals' or 'caregivers' (e.g., health advocates [71], witches [112]). However, the majority of TAs represented under 'other' here involved groups of people with differing roles (e.g., people receiving care interviewed alongside their caregivers; health educators along with their students).

5.3 Research Design

5.3.1 Data Collection. Most TAs (TAs=63) used TA to analyze data generated through *interviews*. It is interesting to note that 90% of TAs were conducted on data generated by the research. This **created data** include *field notes and observations* (TAs=18), *design sessions and workshops* (TAs=13), *focus groups* (TAs=10), *surveys* (TA=7), *participants' diaries and biographies* (TAs=4), *non-textual media*, e.g., videos, photographs, voicemails (TAs=5), workshops artifacts (TAs=7), *chat logs* (TAs=3), and *usability tests* (TAs=2).

[8]Some technologies target several populations.

In opposition, a minority of studies analyze **real-world data**, comprising *mobile applications* (TAs=2), *forum posts/comments and social media data* (TAs=4), and *usage logs* (TAs=3). About a third of the TAs reviewed (TAs=30) employed multiple methods to collect data, most commonly qualitative data methods such as *interviews* (TAs=25), *researchers notes* (TAs=8), and *surveys* (TAs=7). Other methods typically more suited to quantitative data collection include *usage logs* (TAs=7), *usability questionnaires* (TAs=3) and *clinical measures* (TAs=3). Among the 100 TAs reviewed, one analyzed single-subject data corresponding to teaching materials produced by a kinaesthetics professional [47]. In the rest of the corpus, sample size ranged from 3 (interviews, [72, 81]) to 1,308 (survey responses, [98]), with an average of 46 participants, and median of 20 participants.

5.3.2 Mixed-Method Analysis. We observed that about a quarter of the corpus (TAs=23) combined thematic analysis with other methods of data analysis. In most cases, quantitative methods such as *statistics* (TAs=14) were used to analyze additional quantitative data. Such approaches align with recent calls for 'mixed methods' or methods combining quantitative and qualitative data in HCI for healthcare [51, 86]. Other methods were used to complement thematic analysis findings, including data classification or charting [30, 101, 106], interpretive analysis and description of interactions [75], content analysis [65, 123], and affinity diagrams [124].

5.3.3 TA Output. The output of thematic analysis was typically used to *generate design guidance*. Indeed, over half of the papers reviewed (papers=57) used TA findings to explicitly name future opportunities for design of technologies, such as Doyle et al. [46] who identify design requirements for a self-management tool to support older adults with multiple chronic conditions. TA outcomes were also often used to uncover and present more information about a particular phenomenon (papers=31), therefore *extending existing knowledge*. For instance, a study used TA output to "understand the clinical contexts of eating disorders and social media use" [96, p.2]. A smaller number of studies drew on TA findings to *evaluate a product, service, or system* named or proposed by the authors (papers=23). For example, Hauser et al. [60] assessed a device for simulating maternal skin-to-skin holding for premature infants. TA outcomes sometimes served to explicitly *name future research opportunities* (papers=16), for instance Ding et al. [42] argued for a reconsideration of data engagement in self-tracking practices, and O'Leary et al. [93] reported implications for research methods. Finally, more often than not, TA findings led to a combination of these takeaways. For instance, Tachtler et al. [113] sought to understand the interplay of mental health apps with social ecological systems in which unaccompanied migrant youth were embedded, before formulating design implications.

5.3.4 Method Name. Over half of the TAs were simply named *thematic analysis* (TAs=53, papers=45). Other TAs were named with the addition of a range of descriptive terms. The only term that was very common, being used for 29% of the TAs, was *inductive* (TAs=29, papers=25). Synonymous to inductive, the term *bottom-up* was also used (TAs=2, papers=2). Other descriptors that occurred in more than one paper were: *iterative* (TAs=4, papers=4), *deductive* (TAs=4,

papers=2), and *open coding* (TAs=2, papers=2). The phrasing thematic analysis *approach* also reoccurred (TAs=7, papers=6). One paper termed the approach they used "Braun and Clarke's thematic analysis approach" [121, p. 2633].

5.3.5 Method Citations. About a quarter of the papers (papers=18) did not cite a source in relation to the use of thematic analysis. Over half of the corpus (TAs=51, papers=46) cited a publication by Braun and Clarke [15–17, 19, 24, 31]; most of these papers (papers=38) cited Braun and Clarke's seminal 2006 paper *Using Thematic Analysis in Psychology* [15]. Braun and Clarke's writing on TA was also cited indirectly through citation to a textbook describing it [9]. Publications by other authors on TA were cited a small number of times, these were *Applied Thematic Analysis* by Guest et al. [57] (papers=2), *Transforming Qualitative Information: Thematic Analysis and Code Development* by Boyatzis [14] (papers=2) and *The Identification and Analysis of Themes and Patterns by* Luborsky [78] (papers=1).

Publications relating to grounded theory by Strauss and Corbin [33–35, 108–110][9] (papers=8) were also cited. Some of these citations were in conjunction with TA citation by Braun and Clarke (papers=4) or Boyatzis (papers=1). Additionally, one paper reported the use of *memoing* [61], citing a Charmaz and Belgrave [28] publication on grounded theory.

Some papers cited sources describing more generalized approaches to qualitative analysis. These sources were: *A General Inductive Approach for Qualitative Data Analysis* by Thomas [114] (papers=2), *Qualitative Evaluation and Research Methods* by Patton [97] (papers=2), *Qualitative Interviewing: Understanding Qualitative Research* by [25] (papers=1), and *Qualitative Data Analysis: A Methods Sourcebook* by Miles and Huberman [82][10] (papers=1).

A small amount of citations were used to refer to papers using a similar method or with regard to specific techniques (e.g., assessing data saturation and theme comprehensiveness). Additionally, a few sources were cited which, on inspection, appear to lack relevance.

5.3.6 Justification for Using TA. In the majority of cases, no form of justification for the use of TA was given (TAs=68, papers=48). Two reasons for using TA reoccurred, *to identify themes* (TAs=8, papers=8) or *to organise and understand the data* (TAs=10, papers=8). Some reasons were *coupled to the research objective* (TAs=9, papers=9), for example to describe design strategies, insights, or goals. A few reasons were given that were more related to the characteristics of TA. One TA was phrased allowing *for inductive analysis* and another as allowing *for deductive analysis*. One paper chose TA because it acknowledges subjectivity and another because of its flexibility. A further TA was used to *characterize the diversity of ideas*.

5.3.7 Theoretical Position. Few papers provided details regarding epistemology and ontology. Rooksby et al. [101] describe a theoretical perspective of *realism* and give reason for its suitability:

> "The theoretical perspective underlying our analysis is one of "realism", simply meaning that we take the participant's opinions at face value (as opposed to looking for underlying motives or social constructs) [24].

---

[9]Numerous errors were found within these citations including incorrect author ordering, years and DOIs.

[10]Unclear which edition due to a contradiction in the reference.

This is appropriate for studying acceptability where subjective opinions are of importance, even if these are mistaken or underdeveloped. This perspective acknowledges that an aspect of making interventions more acceptable may be to educate and explain."

Doyle et al. [46] describe their analysis as *semantic* which we interpret as indicating that, similar to the realist position, the analysis will focus on what is explicitly expressed in the data. Two papers described the analysis as *constructivist* [76, 94]; notably, both of these explicitly acknowledge making use of grounded theory techniques.

Multiple papers describe variations of *feminist position* [7, 64, 94, 106, 120, 123]. A couple of papers describe an emancipatory action [120], including in conjunction with a social-justice oriented research practice [92].

5.4 TA Conduct & Reporting

5.4.1 Positionality Statement. Positionality statements were found in 14 papers of the corpus reviewed. These statements varied somewhat, but typically included descriptions of the authors' countries of origin and/or where they now live and work: "She [the author] works as a researcher at an English university" [106, p.4]. Many position statements described authors' professional experiences, including their educational background, previous and/or current role(s), where they might have worked in the past, and for how long:

> "Two authors have 7+ years of experience studying CHWs in South Asia and Africa" [92, p.5];

> "Of the four researchers who completed the qualitative analysis, two have background in HCI, one has a background in psychology, and one in bioethics" [96, p.5].

Many statements referenced gender identity:

> "VS is a white, able-bodied, non-binary German person in their early thirties" [106, p.4];

> "All authors for this work are Muslim and five out of six authors identify as female and one as male" [85, p.5].

Several report on their authors' own personal views or epistemology:

> "Our analysis in this paper is shaped by our postcolonial feminist leanings and a growing sensitization to the marginalizations resulting from intersectional factors such as gender, religion, caste, and class that surfaced in our fieldwork" [64, p.5].

Other factors reported on include the authors' ability/disability status; their ethnicity; and their age. The position statement is sometimes a place where authors explain elements of their research process in relation to their authors' own positionality. One paper [85, p.5] explains the presence of a male on their research team as providing "an unbiased perspective"; another [96, p.5] describes selecting its authors professional background as

> "purposeful by design to ensure that different perspectives would be reflected within the analysis."

Finally, another paper [30, p.6] suggests that its authors' background acted as

> "a possible reason for the skew in demographics of the participants recruited for the interviews as we used snowball and purposive sampling".

5.4.2 Collaboration. While 22 TAs did not explicitly describe or use language implying how the work of the TA was potentially distributed, the rest (TAs=78) framed the TA as a team activity. For 43 TAs, the involvement of multiple researchers was explicit with specific people (e.g., "the first author", authors identified by initials) or a specific number of people (e.g., "two researchers") being described as conducting components of the analysis (i.e., coding or theme development). For a further 7 TAs, collaboration was limited to group discussions with a single person conducting the majority of the analysis. Many TAs (TAs=28) were described using *we* terminology (e.g., *we* analyzed the data, *we* conducted thematic analysis). With such cases we consider it unclear if the TA was conducted collaboratively as the use of such language may be convention, potentially fueled in part by a stigmatization of qualitative research being conducted by a single coder. These findings resemble those of McDonald et al. [80] when reviewing qualitative analysis more broadly.

An area where collaboration practices were sometimes reported more explicitly is coding. In a small number of TAs (TAs=13), coding was completed by a single researcher. An interesting consideration is how coding was conducted when there were multiple coders. Among those, some studies (TAs=15) reported coders working on the same dataset (e.g. 'the first and second authors independently coded the data'). In a minority of TAs (TAs=3), once the codebook was developed, the dataset was split between several coders (e.g., 'the dataset was divided among the five coders'). One study [125, p.4] mentioned that two researchers "coded four [interview] transcripts together' and another reported that the "lead author coded 2 transcripts (25%) collaboratively with a second author" [55, p.8]. It is important to note that in almost half of the TAs (TAs=41), the role of the coder(s) is either unclear or not reported.

5.4.3 Data Preparation. Many papers report *transcribing* interviews in preparation for analysis. While the majority of these papers do not provide details on the transcription process, some (papers=5) state that the transcription was manual [66], outsourced (e.g., "all interviews were transcribed by a HIPAA-compliant vendor" [49, p.644]), or verified (e.g., "transcripts were checked against the audiotapes for accuracy" [81, p.565]). For some studies where the data was not recorded in English, *translation into English* prior to analysis is reported [40, 47, 66, 67, 73, 92, 105, 112]. It is interesting to note that one study conducted TA on the original Hindi transcriptions "to understand the local terms and associated nuances better" [123, p. 5]. As translation might be considered adding another layer of interpretation to the data, there might be value in reporting whether translation was automated, manual (as detailed in one study [66]), and verified.

5.4.4 Tools. For most TAs (TAs=80), it is not described what tools and materials were used to conduct the analysis. Less than a quarter of TAs (TAs=19) report using specialist qualitative analysis tools

(NVivo, 9; Atlas.ti, 4; Dedoose, 3; Taguette, 1; MAXQDA, 1; Re-framer, 1). Two of these TAs also noted coding the data manually. An additional paper reported using a spreadsheet. The features of specialist tools, such as those for code organization, may be impactful for the analysis. Yet, how these tools are used is currently unclear.

5.4.5 Familiarization. Two thirds of the TAs (TAs=67) describe steps taken by the researchers to familiarize themselves with the data set prior to coding. By far, the most commonly noted was *transcription*, with 48 TAs reporting transcription occurring as part of their method. Beyond this, *reading* was also noted as occurring in 17 TAs, and *translation* in 9 TAs. Other TAs (TAs=5) reported note-taking as a method of familiarization [8, 71, 74, 91, 92], while 6 others simply noted that authors 'familiarized' themselves with the dataset. Two TAs described how researchers listened back to recordings. Another TA had their authors use an apps (later to be subject to analysis) for a period of time in order to become familiar with them [41].

5.4.6 Coding. Data coding constitutes an important step of thematic analysis, and is described by Braun and Clarke [22, p.53] as:

> "the process of exploring the diversity and patterning of meaning from the dataset, developing codes, and applying code labels to specific segments of each data item".

In this part of the analysis we examined explicit descriptions of the coding process. While the majority (TAs=62) of TAs in the corpus include information on the coding process — beyond 'the data was coded' — a significant number (TAs=30) do not provide a description of the approach followed. We note that for this part of the analysis, we will not infer a TA named, for example, "inductive thematic analysis" to code inductively, unless explicitly stated[11].

Information on the origin of codes was reported in over half of the corpus of TAs (TAs=53). Almost half of the TAs (TAs=43) describe a one-stage coding process. Some of these report using *open coding* (TAs=8), with some papers referencing grounded theory publications. A publication by Braun and Clarke [15] was also cited in relation to open coding, however, open coding is not part of the method. Some TAs report using *inductive* or *bottom-up* coding (TAs=9), which we interpret as "analysis 'grounded in' the data" [15]. About a quarter of the corpus (TAs=20) describes a multi-stage coding process, with 14 TAs referring to an 'iterative' development of codes. A small number of TAs (TAs=9) report coding data through axial coding, a technique anchored in grounded theory. *Axial coding*, introduced by Strauss [111], is defined by Strauss and Corbin as "a set of procedures whereby data are put back together in new ways after open coding, by making connections between categories" [110, p.96]. Some TAs in this subset align with this, employing axial coding after a round of open coding [30, 36, 125, 126]. Early coding was reported in two TAs, whereby the coding process started as soon as the first pieces of data were collected [76], and in order to "define[d] areas that we needed to collect more data,

and refine[d] our interview questions accordingly' [95, p.4]. Finally, over a quarter of the TA reports (TAs=29) indicate involving other team members to discuss the codes.

In some cases (TAs=18), a codebook was developed and then applied during the coding process. Studies relying on a codebook assessed agreement between coders, either informally through discussions, or statistically by calculation of inter-rater reliability (n=6). Inter-code agreement was then used as part of the codebook development (n=7) or in order to finalize the coding (n=7). These processes begin to resemble codebook and coding reliability approaches to TA (see Section 2.2).

5.4.7 Creating Themes. Part of thematic analysis is creating themes through and from coding of the data. Indeed, 3 of the 6 phases of thematic analysis described by Braun and Clarke [22] regard this: generating initial themes; developing and reviewing themes; refining, defining and naming themes.

40% of the TAs (TAs=40) did not provide any description of the theme generation. Of the 60 TAs that reported information, 11 TAs provided descriptions limited to stating that themes emerged, were identified, or similar.

Over a third of the descriptions (TAs=22) described a process of *codes being organized or grouped into themes*. A small number of TAs (TAs=8) described using *axial coding* (defined in the previous section) to achieve themes. Seven TAs described generating themes *deductively* by drawing on the themes of prior TAs or relevant theories. *Early generation* of themes, during data collection or before much of the data had been coded, was reported by 5 TAs.

A small number of the theme generation descriptions (TAs=8) explicitly reported the process as *iterative*. Sixteen TAs described some form of *theme review*, such as comparing the themes to the data or to other themes. Terms such as 'exhaustiveness', 'comprehension', as well as 'saturation' were used. The terms 'constant comparison' and 'theoretical saturation' — which originate from grounded theory — were also present. Related to this, 2 TAs described calculating *per-theme inter-rater reliability*. A few studies (TAs=5) stated a step of *defining or naming* themes, which aligns with the stages of thematic analysis described by Braun and Clarke [15]. Three TAs reported using a thematic map. Some TAs (TAs=19) reported using group discussion to develop themes.

A challenge we found when attempting to chart descriptions of theme creation was that they often used imprecise language, such that it was difficult to understand the relationship between codes and themes. Sometimes, it seemed the terms 'codes' and 'themes' were used near interchangeably, a characteristic that has been identified as occurring in research using TA beyond HCI [21].

5.4.8 Saturation. Eight TAs adopted a data collection relying on data saturation (e.g. 'data was collected until saturation was reached'). A small number of studies considered saturation when coding the data (TAs=4). Particularly, a study referred to "code saturation" citing a publication by Aldiabat and Le Navenec [3], and reported they "considered saturation to have occurred when no more codes emerged" [88, p.6]. Saturation was also mentioned in relation to theme generation (TAs=2) to support 'themes refinement' [76] and 'theme comprehensiveness across participants' [55], citing the works of Fusch and Ness [54] and Weiss [122].

---

[11]Our reason for this is that the method name is too ambiguous. For example"inductive thematic analysis" could be taken to mean themes directly created from the data without a coding process taking place.

5.4.9 Reporting Themes. The large majority of TAs (TAs=80) reported the set of themes resultant from the analysis. When first developing our study protocol, we had intended to categorize themes by what Braun and Clarke [22] describe as the two predominant conceptualizations of *patterns of shared meaning* and *topic summaries*. However, from piloting this, we concluded the approach was unsuitable as often papers would report themes presenting characteristics of both conceptualizations. We instead charted themes as being clear examples of patterns of shared meaning (TAs=25), clear examples of topic summaries (TAs=23), or other (TAs=32). The TAs were quite equally distributed among these categories. We would expect publications citing a reflexive TA approach to report themes in the form of *patterns of shared meaning*. However, we recognized that to be clearly the case in only 35% of this subset.

Within the reporting of themes, the use of participant quotes was prevalent. Over half of the TAs (TAs=68) included a quote for *every* theme, while 9 TAs included a quote for *some* themes. We noted that 3 TAs did not use quotes when reporting themes. While the majority of TAs (TAs=52) did not report participants count in their description of themes, close to a third of the corpus (TAs=28) did so, with 11 TAs reporting counts of participants represented by *every* theme, and 17 TAs reported counts for *some* themes. Interesting examples of reporting included a participant incidence matrix [123] and summary tables (e.g., [75, 102]).

Initially we assumed that papers using TA would explicitly report the resultant themes. However, we realised during the piloting that this was not the case. We found that 20% of the TAs (TAs=20) in our corpus did not explicitly report themes. For these TAs, we charted descriptions of what was reported. We noted general descriptions of findings [68, 71], sometimes combining the results of multiple studies or analysis approaches [120], meta analysis of the themes [124], and design guidelines [70]. Some studies grouped TA findings by research question [47, 79].

5.4.10 Alignment with Cited Methods. While the objective of this work — a scoping review — is not evaluate the quality of results (e.g., to evaluate how well texts citing a method adhere to the procedure), our analysis highlighted a clear presence of grounded theory techniques reported within TAs citing Braun and Clarke [15]. Indeed, a significant proportion of the corpus (TAs=51) report following Braun and Clarke's approach. Among this subset, a third of studies (TAs=15) reported using non-reflexive TA practices, or grounded theory techniques, such as open coding (TAs=10), axial coding (TAs=9), inter-rater reliability (TAs=2), saturation (TAs=1), and early coding (TAs=1). Most of these (TAs=9) did not provide a rationale for combining approaches. However, a small number explicitly reported complementing their reflexive TA with further analysis relying on grounded theory (3 TAs, [30, 74, 99]) or saturation principles (2 TAs, [55, 95]. Although we do not argue against using grounded theory techniques when conducting TA, it can be problematic to use such techniques when *claiming* to use — in other words citing — reflexive thematic analysis. This relates to the writing of Braun and Clarke [21, p. 10] which describes the unwarranted use of grounded theory processes as a common problem, as these are embedded in located and particular meanings and theories which "do not always translate (well) to, or cohere with, TA". This observation also corresponds with the findings of McDonald

et al. [80] that open and axial coding are often used within wider qualitative HCI research without being related to grounded theory.

6 DISCUSSION

Our research aim was to understand the emergent norms of thematic analysis in healthcare HCI at CHI, and to identify opportunities for improvement going forward. We conducted a scoping review of a decade (2012 – 2021) of healthcare research published at CHI; this included 78 papers reporting 100 TAs. We anticipated this would be straightforward given the commonalities of the research, publication norms, and general reporting guidance. On the contrary, the activity was challenging and required repeated adaptation of our analytical approach, to account for the diversity of practices we were identifying.

For the most part, we found descriptions of TA practices be light in explanation or imprecise. We identified little reporting of researcher(s)' positionality or reflexive practice, despite the majority of papers citing a reflexive TA approach. We observe that the healthcare CHI community is often using TA as a team, but that how such collaboration takes place is unclear. Our analysis also shows TA is both commonly used to analyze interview data and to generate design insight. We also highlight a considerable presence of grounded theory techniques, including when a reflexive TA source is cited which it is argued can be a problematic practice [21].

In this section we discuss opportunities to improve research practice when using thematic analysis in healthcare HCI.

6.1 TA Practices Should Be Reported

Overall, we observe descriptions of TA practices to be often light in explanation or imprecise — particularly on theoretical position (Section 5.3.7), coding process (Section 5.4.6) and theme creation (Section 5.4.7). This may in part be the result of space limitations and the hourglass model of paper writing whereby details of the analysis process are restricted to short sections [53]. An avenue for mitigating these challenges could be the increased use of supplementary materials. Echoing recent work on qualitative research in HCI [80] and on TA more broadly [21], we argue that reporting the research process and its rationale is of significant value. Explanations serve to support other researchers and reviewers with interpreting and evaluating research [89]. This is particularly the case with TA, given variability both between and within approaches [21]. As highlighted by McDonald et al. [80] writing on HCI, detailed methodological accounts are valuable as training resources, particularly for researchers receiving less guidance from those more experienced. Furthermore, the absence of detailed accounts of how people are doing TA could itself be a key reason for the substantial variation in the method's use [116]. Extending this notion, by not providing detailed accounts of our research practices we may, as a community, be missing out on opportunities to identify and develop TA approaches that more effectively meet our research needs, particularly in the domain of healthcare. For instance, the outcome of reflexive TA is usually a set of themes, yet the target analytic output for HCI researchers is often to form design implications or develop new products and systems [44, 45]. To support this, authors and reviewers can make use of recently published guidance for evaluating TA manuscripts. Motivated by

observing similar TA reporting issues in scholarship more widely, a twenty question tool for the assessment of TA manuscripts, applicable to all TA broadly, was created [21]. Although a valuable starting point, this guidance should be treated as partial and modifiable to the concerns of healthcare HCI which we shall touch upon in the next sections.

6.2 Using TA Reflexively and Positionality

Although a focus of reflexive TA, reflexivity and positionality are critical to not only all TA approaches, but to all research: both qualitative and quantitative. This is especially the case for healthcare HCI which, as discussed in Section 2.1, studies and shapes healthcare systems for populations often vulnerable, stigmatized, or discriminated against. In addition to during data analysis, reflexivity is also key throughout research design and data collection, particularly when interviews are used (which our analysis shows is often the case), as researchers' positions might affect participant recruitment [53] and study conduct [22]. Yet, few papers in our corpus reported on the researchers' position or reflexive practices; what researchers chose to report also varied from one paper to another (see Section 5.4.1). This warrants the question of what aspects of a researcher's positionality are beneficial to report? Indeed, it is one we are ourselves grappling with as we write this paper.

Many factors contribute to researchers' positions. When addressing healthcare issues, interesting elements might include researchers' own health and healthcare experiences, and personal characteristics influencing those, such as gender, ethnicity, (dis)abilities, and wealth. Some approaches to reporting on positionality consist of descriptions of the researchers' characteristics, such as aspects of their identity, educational and research background, and personal views. Such approaches have been referred to as 'shopping list' positionality whereby:

> "engagement is purely descriptive, providing a 'shopping list' of characteristics and stating if these are shared or not with participants" [52]

These lists risk being long, particularly in the case of research teams, while lacking in nuance. The value such reporting offers the reader is unclear — should we interpret these lists as disclosures, or even as qualifiers to conduct research? A concern with this approach, particularly in healthcare, is that it could construct the (problematic) expectation that researchers *should* self-disclose sensitive information (e.g., those researching mental health should disclose about their own mental health). Focused reporting of key ways in which the research team sought to be reflexive and reflections upon its efficacy could offer greater value. Work could discuss differences in researchers' interpretations; how researchers' views developed through conducting the research; and specific ways in which the researchers' position may have supported or challenged their analysis of certain content (e.g., experience, or a lack of, with a healthcare topic). For instance, referring back to the physical activity example [17] discussed in Section 2.1, reporting of the researchers' different interpretations may not only benefit the reporting of the analysis, but also support researchers to conduct related research reflexively. Discussion could extend to describing practices used to protect the wellbeing of researchers engaging with sensitive content, as is often the case in healthcare research [39].

To support such reflexivity researchers can use techniques such as reflexive journaling [22, 34, 53].

The peer review system poses challenges for researchers wanting to report on their positionality and reflexive practices. One issue is that such reporting can conflict with current norms [80]. Disciplines that historically value positivist 'ways of knowing' (including HCI) may struggle with explicit declarations of researcher characteristics and demographics, and may disagree or misunderstand these statements as introducing bias rather than attempting to account for the researcher as an active part of the process [22]. A second issue is that of satisfying the anonymization policy of the double-blind review processes used by CHI and other publication venues. The anonymization policy of CHI 2023 [29] states:

> "Make sure that no description that can easily reveal authors' names and/or affiliations is included in the submission (e.g., too detailed descriptions of where user studies were conducted)."

Reporting informative descriptions of the study context, author positionality, and reflexive practice clearly risks violating this. While authors can remove content and mark its absence for review processes (e.g., 'Removed for Anonymization'), such removals may be considerable, and in turn negatively impact the paper's prospects of acceptance. These challenges may in part explain the absence of greater reporting of reflexivity and positionality; given the publish-or-perish reality of research, research practices will typically be shaped toward maximizing the likelihood of paper acceptance.

6.3 Using TA as a Team

Our analysis (Section 5.4.2) demonstrates that TA is widely framed as a team activity within healthcare HCI, yet it is often unclear how researchers work together to code data and generate themes. Group discussion is commonly reported as a practice to support this, but the nature of these discussions is given little description. In fact, researchers considering the challenges and opportunities of qualitative methods within HCI have raised the question of how to conduct research collaboratively [50].

Although reflexive TA (as cited by the majority of papers) is suited to team approaches [22], which resonates with the interdisciplinary nature of healthcare HCI, it is argued that reflexive TA "works especially well with a single researcher" [22, p. 248]. The closest TA accounts to being single researcher were the small number that reported a single coder along with group discussion. We emphasize that there is no problem with reflexive TA being conducted by a single researcher, and acknowledge that this is also a matter of logistics and equity: not all research will have the resource for multiple researchers to be heavily involved in the analysis [80].

Braun and Clarke [22] describe two broad approaches to coding as a team: consensus and collaborative. *Consensus* coding involves developing a 'best fit' coding of the data; it occupies a small q position, concerned with accurate and reliable coding — such approaches are characterized by settling disagreements and aiming for high levels of inter-rater reliability. The second is termed *collaborative* coding and involves developing the analysis with emphasis on reflexivity. Each researcher is involved with coding the data, together the researchers discuss and reflect on their ideas and assumptions. This approach aligns with reflexive TA [22] which

appears the most influential TA approach in healthcare HCI (Section 5.3.5). While there exists guidance on consensus coding techniques (such as for coding reliability [80]), advice for using reflexive TA collaboratively seems more scarce. Informed by descriptions in our corpus and writing on reflexive TA, we next discuss the challenges and opportunities of different ways of collaborating to inform research teams going forward.

6.3.1 Collaborative Coding & Theme Generation. Researchers could code the data *together* to produce a single coding (e.g., two researchers sat together using the same interface in a manner resembling pair programming); this is potentially a practice being reported with work describing researchers coding together [125] and collaboratively [55]. Alternatively, researchers could code the data more *independently* whereby they work on separate copies of the data, and meet to regularly discuss and iteratively develop the analysis [23]. Crucially, researchers should be engaging with the same data and not splitting the dataset between them. Researchers analyzing different subsets of the data means each is being analyzed by a single perspective and that different parts of the data are receiving different analytic attention; this largely undermines the intent of collaborative approaches to involve multiple interpretations and perspectives. Often within our corpus research reports 'multiple researchers coded the data independently' from which it is ambiguous what data each researcher engaged with; we argue this is an important detail that should be explicitly reported.

Multiple researchers working with the entire dataset is however resource intensive and therefore may not always be an option. More feasible approaches could be for additional researchers to contribute to analyzing certain portions of the data, or to be involved only at certain stages of the analysis. For instance, an additional researcher could be involved at the start of the analysis in attempt to bootstrap the identification of alternative interpretations. Similarly, another researcher could apply the primary researcher's codebook to support identifying differences in interpretation that could be used to develop the analysis. We emphasize the purpose of this is not to reach consensus or a high coding reliability, but to develop a more diversified analysis through the inclusion of multiple perspectives. Another technique is that expert researchers (e.g., ones with greater clinical or design expertise), or researchers whose position differs in a way pertinent to the content, could contribute focused analysis on pertinent extracts.

Multiple researchers developing — although in collaboration — multiple codings of the data could be perceived as complicating the process(es) of theme development. As our work demonstrates, theme development is commonly equated to a process of code organization. In reflexive TA approaches, whereby coding is an analytic *process* rather than an analytic *product* [22], researchers can collaboratively develop themes from multiple codings — a process which goes beyond simple code organization. Braun et al. [23] describe a process they informally term as a "theme off":

> "Gareth and David again worked independently and collaboratively in the early stages of theme construction, meeting regularly to discuss candidate themes. Their meetings took the form of a kind of 'theme off': each presented their candidate themes, including preliminary theme names and definitions (discussed

soon); they then 'tussled' with each theme, and the collection of themes, to identify the most meaningful potential themes, the ones that collectively told the best story of the data." [23, p 855]

6.3.2 Group Discussion. Group discussion was a commonly reported practice to support coding and theme generation. Yet, the nature of these group discussions again was given little description. We believe that the role of group discussion within analysis is a complex topic, presenting both challenges and opportunities. Within a collaborative analysis approach, talking with others can contribute to clarifying the analysis and introducing alternative interpretations, therefore supporting reflexivity [22]. A possible concern, particularly with inductive approaches, is that researchers involved in the discussion may lack familiarity with the data (e.g., have not directly engaged with or coded it) and thus might exert an influence that is not founded in the data. This may be particularly the case when more expert (e.g., clinicians, more experienced HCI researchers) or higher ranking (e.g., supervisors) persons are contributing from a position of lower data familiarity. Teams should take extreme care with the inclusion of this input, and be sure to diligently evaluate its relation to the data (e.g., through reflection on existing, and the introduction of new, codes). An alternative way to use group discussions — potentially suiting groups who are less intimate with the data — is to focus on supporting the analyst(s) with producing a high quality analysis. Such group work could concentrate on asking questions of the analysis being produced (e.g., inspecting the relation between themes and the data associated with them) and seeking to prevent premature closure of the analysis [22].

Logistically, how can group discussions be conducted effectively? We currently see an abundance of tools for gathering and preparing data (e.g., to record, transcribe, or translate), as well as an increasing range of specialist tools for conducting thematic analysis (see Section 5.4.4). We thus wonder how tools can support group discussion during TA, potentially by promoting the involvement of data in the discussion process.

Furthering Section 6.1, we argue future HCI work should describe how researchers conduct analysis as an individual or as a team, so as to support external evaluation and inform future practices.

6.4 The Research Fit of TA

Despite prevalent TA use within healthcare HCI (Section 2.2), it is most often unclear why researchers choose TA; nor if the use of TA fits the theoretical assumptions of the research (see Section 5.3). Long existing guidelines for reflexive TA have argued that "the assumptions about, and specific approach to, thematic analysis [should be] clearly explicated" [15, p.96]. A development of these guidelines applicable to TA approaches more broadly, published since the TAs in our corpus would have been conducted, argue that it should also be explained why TA was chosen [21]. Hopefully going forward, reporting of the reasons for choosing TA and theoretical assumptions will become the norm, enabling us to develop our understanding of the research fit of TA for healthcare HCI.

6.4.1 TA Approaches. We demonstrate with our analysis that in healthcare HCI, we often collect data in fairly structured ways

such as semi-structured interviews, questionnaires, and through technology evaluations. This structure is a key characteristic of the research, and impacts the fit of analysis approaches. For instance, a study in our corpus writes:

> "As the interviews were semi-structured and typically included the same list of questions focused around a fairly narrow set of topics, codes were generally structured around responses to individual questions or categories of similar questions." [96, p.5]

Furthermore, our analysis shows TA is being most often used to generate design guidance; thus, often 'themes' (or even 'topic summaries') are not the ultimate goal of the analysis. Indeed, for studies involving more structured data collection and aiming to generate design guidance — particularly those evaluating technologies — inductive reflexive TA may not be a strong research fit.

Research may benefit from more deductive approaches informed by relevant HCI models (e.g., technology acceptance [87] or personal informatics [48]) or the design insights of closely related work. This echoes the writing of Furniss et al. [53] on using grounded theory in HCI, which argues for the value of a top-down approach that benefits from existing theoretical concepts and structures. Approaches combining inductive and deductive elements, which benefit from existing knowledge while enabling data driven extension, could be particularly effective. What's more, they may better describe some approaches framed as purely inductive because researchers will often be knowledgeable of, and thus influenced by, prior work.

It is worth reflecting upon if theme development is a valuable analysis step, or whether the codes and (grouped) qualitative data could be subjected to a different form of analysis, better suited for the generation of design insights. If a coding process offers sufficient description and interpretation of a dataset (e.g., responses to some specific interview questions), we should report it as such and consider it an acceptable practice. However, to construct design insights, a creative or interpretive leap is likely required. Although we believe that reflexive thematic analysis is flexible enough to incorporate these leaps, we suggest that researchers should attempt to describe such leaps. Reflexive TA and theme generation is a process that requires creativity, and we believe there may be opportunity to identify or develop effective techniques specifically for generating design insights. To do this, HCI researchers might borrow more strongly from needs-finding and ideation methods commonly used by designers. Needs-finding approaches in healthcare design already exist — for instance, Stanford Biodesign's observations-problems-needs framework offers a simple and highly verbal way to move from research findings to ideation statements [107]. It could be that such frameworks may be too specified and reductive within a reflexive TA framing. However, simple visual methods such as brainstorming and spider diagramming help designers to ground and make tangible their movements from research findings to beginning design ideas [10]. As the process evolves further, ideation methods such as lateral thinking exercises (provocation, fractionation and more) help meld visuals and text to further ground and explicate a designers' ideation process [37], while tools such as SCAMPER [59] help designers to diversify their design insights in order to achieve wider applicability. Such creative exercises, with

their capacity for reflexivity and flexibility, could fit well within reflexive TA, and are already present to some degree in HCI-adjacent publishing, such as DIS'[12] pictorials submission route.

Our considerations have focused on the reflexive approach to TA by virtue of it being the most influential approach within healthcare HCI. As noted in Section 2.2, alternative approaches including codebook and coding reliability may better fit certain research. Further work is required to determine for what healthcare HCI research these approaches would be a better fit.

6.4.2 Alternatives to TA. Our analysis demonstrates a clear presence of grounded theory techniques, such as open and axial coding, as well as saturation. While it is argued that this can be a problematic practice in the context of reflexive TA [21], it could suggest that grounded theory techniques offer researchers something they find lacking with TA. For example, axial coding is being reported as a process for theme generation potentially because it offers greater instruction for how to move from codes to analytic outputs than reflexive TA. A possibility is that some healthcare HCI research which is using TA — for instance research seeking to develop knowledge about populations and healthcare processes rather than design guidelines — may benefit from alternative methods such as grounded theory or interpretative phenomenological analysis. Determining this would however require a broader understanding of the contemporary use of qualitative methods in healthcare HCI.

7 LIMITATIONS

This research aimed to map healthcare HCI research practices for TA at CHI; this was to include contexts of use, conduct, and reporting. Our analysis of how TA is conducted is through the lens of what is reported, but the reporting of the analysis procedure is often short. It should also be kept in mind that this reporting is the constructed accounts of researchers who are motivated to get their work accepted, and are influenced by norms. Future work should engage with researchers directly to study their practices, with focus on aspects such as collaboration.

Our search included only papers published at CHI that explicitly use the term "thematic analysis". We highlight that local standards for TA may differ for other publication venues. An area for future work would be to examine the local standards of TA in healthcare HCI at other venues. Papers were not included if they described processes resembling TA (i.e., using coding to identify themes) without explicitly terming the analysis "thematic analysis". This decision was made to ensure that we did not include analyses that were not considered to be TA by the researchers who conducted them.

Our discussion considered the relation of TA to other analysis methods due in part to the presence of grounded theory techniques. However, our ability to reflect on this is limited by our review focusing on TA and it being currently unclear how other methods are used within healthcare HCI. Future work should seek to study the use of a range of qualitative research methods, potentially including grounded theory and interpretative phenomenological analysis.

---

[12]ACM SIGCHI Conference on Designing Interactive Systems (DIS)

8 CONCLUSION

This study aimed to understand the emergent norms of TA in healthcare HCI at CHI, and from this to identify opportunities to improve going forward. To do this, we conducted a scoping review of 10 years (2012 - 2021) of healthcare publications at CHI. Our analysis serves to develop a shared understanding of research practices upon which we can reflect, and use to identify opportunities for improvement. To this end, we contribute a discussion of opportunities for the CHI healthcare HCI community to improve its TA practices. We argue the importance of reporting TA practices, and discuss using — and the challenges of reporting — TA reflexively. We identify the opportunities and challenges of different approaches to conducting TA as a team. Finally, we call for the careful assessment of TA's research fit for different forms healthcare HCI research, and for investigation into how TA could better address our research needs. We hope this work will be valuable for both new and experienced researchers using TA, and that it will motivate further research on the use of qualitative methods in healthcare HCI.

ACKNOWLEDGMENTS

REFERENCES

[1] Jacob Abbott, Haley MacLeod, Novia Nurain, Gustave Ekobe, and Sameer Patil. 2019. Local Standards for Anonymization Practices in Health, Wellness, Accessibility, and Aging Research at CHI. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems* (Glasgow, Scotland Uk) *(CHI '19)*. Association for Computing Machinery, New York, NY, USA, 1–14. https://doi.org/10.1145/3290605.3300692

[2] Anne Adams, Peter Lunt, and Paul Cairns. 2008. A qualititative approach to HCI research. In *Research Methods for Human-Computer Interaction*. Cambridge University Press, Cambridge, UK, 138 – 157.

[3] Khaldoun M Aldiabat and Carole-Lynne Le Navenec. 2018. Data saturation: The mysterious step in grounded theory methodology. *The qualitative report* 23, 1 (2018), 245–261.

[4] Morgan G. Ames, Janet Go, Joseph 'Jofish' Kaye, and Mirjana Spasojevic. 2011. Understanding Technology Choices and Values through Social Class. In *Proceedings of the ACM 2011 Conference on Computer Supported Cooperative Work* (Hangzhou, China) *(CSCW '11)*. Association for Computing Machinery, New York, NY, USA, 55–64. https://doi.org/10.1145/1958824.1958834

[5] Hilary Arksey and Lisa O'Malley. 2005. Scoping studies: towards a methodological framework. *International journal of social research methodology* 8, 1 (2005), 19–32.

[6] Andreas Balaskas, Stephen M Schueller, Anna L Cox, and Gavin Doherty. 2021. The Functionality of Mobile Apps for Anxiety: Systematic Search and Analysis of Engagement and Tailoring Features. *JMIR mHealth and uHealth* 9, 10 (2021), e26712.

[7] Marguerite Barry, Kevin Doherty, Jose Marcano Belisario, Josip Car, Cecily Morrison, and Gavin Doherty. 2017. MHealth for Maternal Mental Health: Everyday Wisdom in Ethical Design. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems* (Denver, Colorado, USA) *(CHI '17)*. Association for Computing Machinery, New York, NY, USA, 2708–2756. https://doi.org/10.1145/3025453.3025918

[8] Andrew B. L. Berry, Catherine Y. Lim, Tad Hirsch, Andrea L. Hartzler, Linda M. Kiel, Zoë A. Bermet, and James D. Ralston. 2019. Supporting Communication About Values Between People with Multiple Chronic Conditions and Their Providers. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems* (Glasgow, Scotland Uk) *(CHI '19)*. Association for Computing Machinery, New York, NY, USA, 1–14. https://doi.org/10.1145/3290605.3300700

[9] Ann Blandford, Dominic Furniss, and Stephann Makri. 2016. *Qualitative HCI research: Going behind the scenes.* Morgan & Claypool Publishers, USA. 1–115 pages.

[10] Nathalie Bonnardel and John Didier. 2020. Brainstorming variants to favor creative design. *Applied ergonomics* 83 (2020), 102987.

[11] Brian Bourke. 2014. Positionality: Reflecting on the research process. *The qualitative report* 19, 33 (2014), 1–9.

[12] Robert Bowman, Camille Nadal, Kellie Morrissey, Anja Thieme, and Gavin Doherty. 2022. The State of Thematic Analysis in HCI for Healthcare: A Scoping Review Protocol. https://osf.io/7mkzn. Registered: July 4th 2022.

[13] Robert Bowman, Camille Nadal, Anja Thieme, Kellie Morrissey, and Gavin Doherty. 2022. The State of Thematic Analysis in HCI for Healthcare: A Scoping Review Protocol. https://osf.io/hndpa. Registered: April 8th 2022.

[14] Richard E Boyatzis. 1998. *Transforming qualitative information: Thematic analysis and code development.* sage.

[15] Virginia Braun and Victoria Clarke. 2006. Using thematic analysis in psychology. *Qualitative research in psychology* 3, 2 (2006), 77–101.

[16] Virginia Braun and Victoria Clarke. 2012. Thematic Analysis. In *APA handbook of research methods in psychology, Vol. 2. Research designs: Quantitative, qualitative, neuropsychological, and biological.* American Psychological Association, 57–71.

[17] Virginia Braun and Victoria Clarke. 2013. *Successful Qualitative Research: A Practical Guide for Beginners.* Sage.

[18] Virginia Braun and Victoria Clarke. 2014. What can "thematic analysis" offer health and wellbeing researchers? *International journal of qualitative studies on health and well-being* 9, 1 (2014), 26152.

[19] Virginia Braun and Victoria Clarke. 2019. Reflecting on reflexive thematic analysis. *Qualitative research in sport, exercise and health* 11, 4 (2019), 589–597.

[20] Virginia Braun and Victoria Clarke. 2021. Can I use TA? Should I use TA? Should I not use TA? Comparing reflexive thematic analysis and other pattern-based qualitative analytic approaches. *Counselling and Psychotherapy Research* 21, 1 (2021), 37–47.

[21] Virginia Braun and Victoria Clarke. 2021. One size fits all? What counts as quality practice in (reflexive) thematic analysis? *Qualitative research in psychology* 18, 3 (2021), 328–352.

[22] Virginia Braun and Victoria Clarke. 2021. *Thematic Analysis: A Practical Guide.* Sage.

[23] Virginia Braun, Victoria Clarke, Nikki Hayfield, and Gareth Terry. 2019. Thematic Analysis. In *Handbook of Research Methods in Health Social Sciences*, Pranee Liamputtong (Ed.). Springer.

[24] Virginia Braun, Victoria Clarke, and Gareth Terry. 2014. Thematic Analysis. *Qual Res Clin Health Psychol* 24 (2014), 95–114.

[25] Svend Brinkmann. 2013. *Qualitative interviewing.* Oxford university press.

[26] Kelly Caine. 2016. Local Standards for Sample Size at CHI. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems* (San Jose, California, USA) *(CHI '16)*. Association for Computing Machinery, New York, NY, USA, 981–992. https://doi.org/10.1145/2858036.2858498

[27] Karen A Campbell, Elizabeth Orr, Pamela Durepos, Linda Nguyen, Lin Li, Carly Whitmore, Paige Gehrke, Leslie Graham, and Susan M Jack. 2021. Reflexive thematic analysis for applied qualitative health research. *The Qualitative Report* 26, 6 (2021), 2011–2028.

[28] Kathy Charmaz and Linda Liska Belgrave. 2015. Grounded Theory. In *The Blackwell Encyclopedia of Sociology*. John Wiley & Sons, Ltd. https://doi.org/10.1002/9781405165518.wbeosg070.pub2

[29] CHI 2023. 2022. CHI Anonymization Policy. https://chi2023.acm.org/submission-guides/chi-anonymization-policy/. Last accessed: January 5th 2022.

[30] Shaan Chopra, Rachael Zehrung, Tamil Arasu Shanmugam, and Eun Kyoung Choe. 2021. Living with Uncertainty and Stigma: Self-Experimentation and Support-Seeking around Polycystic Ovary Syndrome. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems* (Yokohama, Japan) *(CHI '21)*. Association for Computing Machinery, New York, NY, USA, Article 202, 18 pages. https://doi.org/10.1145/3411764.3445706

[31] Victoria Clarke and Virginia Braun. 2013. Teaching thematic analysis: Overcoming challenges and developing strategies for effective learning. *The psychologist* 26, 2 (2013).

[32] Carol L Connell, Sherry C Wang, LaShaundrea Crook, and Kathy Yadrick. 2019. Barriers to healthcare seeking and provision among African American adults in the rural Mississippi Delta region: community and provider perspectives. *Journal of community health* 44, 4 (2019), 636–645.

[33] Juliet Corbin and Anselm Strauss. 2008. *Basics of qualitative research: Techniques and procedures for developing grounded theory* (3 ed.). Sage publications.

[34] Juliet Corbin and Anselm Strauss. 2014. *Basics of qualitative research: Techniques and procedures for developing grounded theory* (4 ed.). Sage publications.

[35] Juliet M Corbin and Anselm Strauss. 1990. Grounded theory research: Procedures, canons, and evaluative criteria. *Qualitative sociology* 13, 1 (1990), 3–21.

[36] Nediyana Daskalova, Eindra Kyi, Kevin Ouyang, Arthur Borem, Sally Chen, Sung Hyun Park, Nicole Nugent, and Jeff Huang. 2021. Self-E: Smartphone-Supported Guidance for Customizable Self-Experimentation. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems* (Yokohama, Japan) *(CHI '21)*. Association for Computing Machinery, New York, NY, USA, Article 227, 13 pages. https://doi.org/10.1145/3411764.3445100

[37] Edward De Bono. 2010. *Lateral thinking: a textbook of creativity.* Penguin, UK.

[38] Jon Dean, Penny Furness, Diarmuid Verrier, Henry Lennon, Cinnamon Bennett, and Stephen Spencer. 2018. Desert island data: an investigation into researcher positionality. *Qualitative Research* 18, 3 (2018), 273–289.

[39] Maria Dempsey, Sarah Foley, Nollaig Frost, Raegan Murphy, Niamh Willis, Sarah Robinson, Audrey Dunn-Galvin, Angela Veale, Carol Linehan, Nadia Pantidi, et al. 2022. Am I lazy, a drama queen or depressed? A journey through a pluralistic approach to analysing accounts of depression. *Qualitative Research in Psychology* 19, 2 (2022), 473–493.

[40] Pooja M. Desai, Elliot G. Mitchell, Maria L. Hwang, Matthew E. Levine, David J. Albers, and Lena Mamykina. 2019. Personal Health Oracle: Explorations of Personalized Predictions in Diabetes Self-Management. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems* (Glasgow, Scotland Uk) *(CHI '19)*. Association for Computing Machinery, New York, NY, USA, 1–13. https://doi.org/10.1145/3290605.3300600

[41] Anjali Devakumar, Jay Modh, Bahador Saket, Eric P. S. Baumer, and Munmun De Choudhury. 2021. A Review on Strategies for Data Collection, Reflection, and Communication in Eating Disorder Apps. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems* (Yokohama, Japan) *(CHI '21)*. Association for Computing Machinery, New York, NY, USA, Article 547, 19 pages. https://doi.org/10.1145/3411764.3445670

[42] Xianghua (Sharon) Ding, Shuhan Wei, Xinning Gui, Ning Gu, and Peng Zhang. 2021. Data Engagement Reconsidered: A Study of Automatic Stress Tracking Technology in Use. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems* (Yokohama, Japan) *(CHI '21)*. Association for Computing Machinery, New York, NY, USA, Article 535, 13 pages. https://doi.org/10.1145/3411764.3445763

[43] Marissa J Doshi. 2018. Barbies, goddesses, and entrepreneurs: Discourses of gendered digital embodiment in women's health apps. *Women's Studies in Communication* 41, 2 (2018), 183–203.

[44] Paul Dourish. 2006. Implications for Design. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Montréal, Québec, Canada) *(CHI '06)*. Association for Computing Machinery, New York, NY, USA, 541–550. https://doi.org/10.1145/1124772.1124855

[45] Paul Dourish. 2007. Responsibilities and Implications: Further Thoughts on Ethnography and Design. In *Proceedings of the 2007 Conference on Designing for User EXperiences* (Chicago, Illinois) *(DUX '07)*. Association for Computing Machinery, New York, NY, USA, Article 25, 15 pages. https://doi.org/10.1145/1389908.1389941

[46] Julie Doyle, Emma Murphy, Janneke Kuiper, Suzanne Smith, Caoimhe Hannigan, An Jacobs, and John Dinsmore. 2019. Managing Multimorbidity: Identifying Design Requirements for a Digital Self-Management Tool to Support Older Adults with Multiple Chronic Conditions. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems* (Glasgow, Scotland Uk) *(CHI '19)*. Association for Computing Machinery, New York, NY, USA, 1–14. https://doi.org/10.1145/3290605.3300629

[47] Maximilian Dürr, Marcel Borowski, Carla Gröschel, Ulrike Pfeil, Jens Müller, and Harald Reiterer. 2021. KiTT - The Kinaesthetics Transfer Teacher: Design and Evaluation of a Tablet-Based System to Promote the Learning of Ergonomic Patient Transfers. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems* (Yokohama, Japan) *(CHI '21)*. Association for Computing Machinery, New York, NY, USA, Article 233, 16 pages. https://doi.org/10.1145/3411764.3445496

[48] Daniel A. Epstein, An Ping, James Fogarty, and Sean A. Munson. 2015. A Lived Informatics Model of Personal Informatics. In *Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing* (Osaka, Japan) *(UbiComp '15)*. Association for Computing Machinery, New York, NY, USA, 731–742. https://doi.org/10.1145/2750858.2804250

[49] Jordan Eschler, Eleanor R. Burgess, Madhu Reddy, and David C. Mohr. 2020. *Emergent Self-Regulation Practices in Technology and Social Media Use of Individuals Living with Depression.* Association for Computing Machinery, New York, NY, USA, 1–13. https://doi.org/10.1145/3313831.3376773

[50] Casey Fiesler, Jed R. Brubaker, Andrea Forte, Shion Guha, Nora McDonald, and Michael Muller. 2019. Qualitative Methods for CSCW: Challenges and Opportunities. In *Conference Companion Publication of the 2019 on Computer Supported Cooperative Work and Social Computing* (Austin, TX, USA) *(CSCW '19)*. Association for Computing Machinery, New York, NY, USA, 455–460. https://doi.org/10.1145/3311957.3359428

[51] Geraldine Fitzpatrick and Gunnar Ellingsen. 2013. A review of 25 years of CSCW research in healthcare: contributions, challenges and future agendas. *Computer Supported Cooperative Work (CSCW)* 22, 4 (2013), 609–665.

[52] Louise Folkes. 2022. Moving beyond 'shopping list' positionality: Using kitchen table reflexivity and in/visible tools to develop reflexive qualitative research. *Qualitative Research* (2022), 14687941221098922. https://doi.org/10.1177/14687941221098922

[53] Dominic Furniss, Ann Blandford, and Paul Curzon. 2011. Confessions from a Grounded Theory PhD: Experiences and Lessons Learnt. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Vancouver, BC, Canada) *(CHI '11)*. Association for Computing Machinery, New York, NY, USA,

113–122. https://doi.org/10.1145/1978942.1978960

[54] Patricia I Fusch and Lawrence R Ness. 2015. Are we there yet? Data saturation in qualitative research. *The Qualitative Report* 20 (2015), 1408–1416. Issue 9.

[55] Elliot G. Mitchell, Elizabeth M. Heitkemper, Marissa Burgermaster, Matthew E. Levine, Yishen Miao, Maria L. Hwang, Pooja M. Desai, Andrea Cassells, Jonathan N. Tobin, Esteban G. Tabak, David J. Albers, Arlene M. Smaldone, and Lena Mamykina. 2021. From Reflection to Action: Combining Machine Learning with Expert Knowledge for Nutrition Goal Recommendations. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems* (Yokohama, Japan) *(CHI '21)*. Association for Computing Machinery, New York, NY, USA, Article 206, 17 pages. https://doi.org/10.1145/3411764.3445555

[56] Connie Guan, Anya Bouzida, Ramzy M. Oncy-avila, Sanika Moharana, and Laurel D. Riek. 2021. Taking an (Embodied) Cue From Community Health: Designing Dementia Caregiver Support Technology to Advance Health Equity. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems* (Yokohama, Japan) *(CHI '21)*. Association for Computing Machinery, New York, NY, USA, Article 655, 16 pages. https://doi.org/10.1145/3411764.3445559

[57] Greg Guest, Kathleen M MacQueen, and Emily E Namey. 2011. *Applied thematic analysis.* sage publications.

[58] Kate Haddow. 2022. 'Lasses are much easier to get on with': The gendered labour of a female ethnographer in an all-male group. *Qualitative Research* 22, 2 (2022), 313–327.

[59] H. James Harrington and Frank Voehl. 2016. *The Innovation Tools Handbook, Volume 2.* Productivity Press, New York. https://doi.org/10.1201/9781315367699

[60] Sabrina Hauser, Melinda J. Suto, Liisa Holsti, Manon Ranger, and Karon E. MacLean. 2020. Designing and Evaluating Calmer, a Device for Simulating Maternal Skin-to-Skin Holding for Premature Infants. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems.* Association for Computing Machinery, New York, NY, USA, 1–15. https://doi.org/10.1145/3313831.3376539

[61] Jeremy Heyer, Zachary Schmitt, Lynn Dombrowski, and Svetlana Yarosh. 2020. *Opportunities for Enhancing Access and Efficacy of Peer Sponsorship in Substance Use Disorder Recovery.* Association for Computing Machinery, New York, NY, USA, 1–14. https://doi.org/10.1145/3313831.3376241

[62] Andrew Gary Darwin Holmes. 2020. Researcher Positionality–A Consideration of Its Influence and Place in Qualitative Research–A New Researcher Guide. *Shanlax International Journal of Education* 8, 4 (2020), 1–10.

[63] Dennis Howitt. 2019. *Introduction to qualitative research methods in psychology: Putting theory into practice.* Pearson UK.

[64] Azra Ismail and Neha Kumar. 2021. AI in Global Health: The View from the Front Lines. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems* (Yokohama, Japan) *(CHI '21)*. Association for Computing Machinery, New York, NY, USA, Article 598, 21 pages. https://doi.org/10.1145/3411764.3445130

[65] Sue Jamison-Powell, Conor Linehan, Laura Daley, Andrew Garbett, and Shaun Lawson. 2012. "I Can't Get No Sleep": Discussing #insomnia on Twitter. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Austin, Texas, USA) *(CHI '12)*. Association for Computing Machinery, New York, NY, USA, 1501–1510. https://doi.org/10.1145/2207676.2208612

[66] Naveena Karusala, David Odhiambo Seeh, Cyrus Mugo, Brandon Guthrie, Megan A Moreno, Grace John-Stewart, Irene Inwani, Richard Anderson, and Keshet Ronen. 2021. "That Courage to Encourage": Participation and Aspirations in Chat-Based Peer Support for Youth Living with HIV. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems* (Yokohama, Japan) *(CHI '21)*. Association for Computing Machinery, New York, NY, USA, Article 223, 17 pages. https://doi.org/10.1145/3411764.3445313

[67] Konstantinos Kazakos, Siddhartha Asthana, Madeline Balaam, Mona Duggal, Amey Holden, Limalemla Jamir, Nanda Kishore Kannuri, Saurabh Kumar, Amarendar Reddy Manindla, Subhashini Arcot Manikam, GVS Murthy, Papreen Nahar, Peter Phillimore, Shreyaswi Sathyanath, Pushpendra Singh, Meenu Singh, Pete Wright, Deepika Yadav, and Patrick Olivier. 2016. A Real-Time IVR Platform for Community Radio. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems* (San Jose, California, USA) *(CHI '16)*. Association for Computing Machinery, New York, NY, USA, 343–354. https://doi.org/10.1145/2858036.2858585

[68] Christina Kelley, Bongshin Lee, and Lauren Wilcox. 2017. Self-Tracking for Mental Wellness: Understanding Expert Perspectives and Student Experiences. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems* (Denver, Colorado, USA) *(CHI '17)*. Association for Computing Machinery, New York, NY, USA, 629–641. https://doi.org/10.1145/3025453.3025750

[69] Louise H Kidder and Michelle Fine. 1987. Qualitative and quantitative methods: When stories converge. *New directions for program evaluation* 1987, 35 (1987), 57–75.

[70] Yoojung Kim, Eunyoung Heo, Hyunjeong Lee, Sookyoung Ji, Jueun Choi, Jeong-Whun Kim, Joongseek Lee, and Sooyoung Yoo. 2017. Prescribing 10,000 Steps Like Aspirin: Designing a Novel Interface for Data-Driven Medical Consultations. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing*

Systems (Denver, Colorado, USA) *(CHI '17)*. Association for Computing Machinery, New York, NY, USA, 5787–5799. https://doi.org/10.1145/3025453.3025570

[71] Daniel Lambton-Howard, Emma Simpson, Kim Quimby, Ahmed Kharrufa, Heidi Hoi Ming Ng, Emma Foster, and Patrick Olivier. 2021. Blending into Everyday Life: Designing a Social Media-Based Peer Support System. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems* (Yokohama, Japan) *(CHI '21)*. Association for Computing Machinery, New York, NY, USA, Article 168, 14 pages. https://doi.org/10.1145/3411764.3445079

[72] Matthias Laschke, Christoph Braun, Robin Neuhaus, and Marc Hassenzahl. 2020. Meaningful Technology at Work - A Reflective Design Case of Improving Radiologists' Wellbeing Through Medical Technology. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) *(CHI '20)*. Association for Computing Machinery, New York, NY, USA, 1–12. https://doi.org/10.1145/3313831.3376710

[73] Celine Latulipe, Amy Gatto, Ha T. Nguyen, David P. Miller, Sara A. Quandt, Alain G. Bertoni, Alden Smith, and Thomas A. Arcury. 2015. Design Considerations for Patient Portal Adoption by Low-Income, Older Adults. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems* (Seoul, Republic of Korea) *(CHI '15)*. Association for Computing Machinery, New York, NY, USA, 3859–3868. https://doi.org/10.1145/2702123.2702392

[74] Kwangyoung Lee and Hwajung Hong. 2018. MindNavigator: Exploring the Stress and Self-Interventions for Mental Wellness. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems* (Montreal QC, Canada) *(CHI '18)*. Association for Computing Machinery, New York, NY, USA, 1–14. https://doi.org/10.1145/3173574.3174146

[75] Minha Lee, Sander Ackermans, Nena van As, Hanwen Chang, Enzo Lucas, and Wijnand IJsselsteijn. 2019. Caring for Vincent: A Chatbot for Self-Compassion. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems* (Glasgow, Scotland Uk) *(CHI '19)*. Association for Computing Machinery, New York, NY, USA, 1–13. https://doi.org/10.1145/3290605.3300932

[76] Catherine Y. Lim, Andrew B.L. Berry, Andrea L. Hartzler, Tad Hirsch, David S. Carrell, Zoë A. Bermet, and James D. Ralston. 2019. Facilitating Self-Reflection about Values and Self-Care Among Individuals with Chronic Conditions. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems* (Glasgow, Scotland Uk) *(CHI '19)*. Association for Computing Machinery, New York, NY, USA, 1–12. https://doi.org/10.1145/3290605.3300885

[77] Sebastian Linxen, Christian Sturm, Florian Brühlmann, Vincent Cassau, Klaus Opwis, and Katharina Reinecke. 2021. How WEIRD is CHI?. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems* (Yokohama, Japan) *(CHI '21)*. Association for Computing Machinery, New York, NY, USA, Article 143, 14 pages. https://doi.org/10.1145/3411764.3445488

[78] Mark R Luborsky. 1994. The identification and analysis of themes and patterns. In *Qualitative methods in aging research*, J. F. Gubrium and A Sankar (Eds.). Sage, 189–210.

[79] Yuhan Luo, Bongshin Lee, Donghee Yvette Wohn, Amanda L. Rebar, David E. Conroy, and Eun Kyoung Choe. 2018. Time for Break: Understanding Information Workers' Sedentary Behavior Through a Break Prompting System. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems* (Montreal QC, Canada) *(CHI '18)*. Association for Computing Machinery, New York, NY, USA, 1–14. https://doi.org/10.1145/3173574.3173701

[80] Nora McDonald, Sarita Schoenebeck, and Andrea Forte. 2019. Reliability and Inter-Rater Reliability in Qualitative Research: Norms and Guidelines for CSCW and HCI Practice. *Proc. ACM Hum.-Comput. Interact.* 3, CSCW, Article 72 (nov 2019), 23 pages. https://doi.org/10.1145/3359174

[81] Mollie McKillop, Lena Mamykina, and Noémie Elhadad. 2018. Designing in the Dark: Eliciting Self-Tracking Dimensions for Understanding Enigmatic Disease. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems* (Montreal QC, Canada) *(CHI '18)*. Association for Computing Machinery, New York, NY, USA, 1–15. https://doi.org/10.1145/3173574.3174139

[82] Matthew B Miles and A Michael Huberman. 1984. *Qualitative data analysis*. Ca: Sage.

[83] Thekla Morgenroth and Michelle K Ryan. 2020. The effects of gender trouble: An integrative theoretical framework of the perpetuation and disruption of the gender/sex binary. *Perspectives on Psychological Science* 16 (2020), 1745691620902442. Issue 6.

[84] Zachary Munn, Micah DJ Peters, Cindy Stern, Catalin Tufanaru, Alexa McArthur, and Edoardo Aromataris. 2018. Systematic review or scoping review? Guidance for authors when choosing between a systematic or scoping review approach. *BMC medical research methodology* 18, 1 (2018), 1–7.

[85] Maryam Mustafa, Kimia Tuz Zaman, Tallal Ahmad, Amna Batool, Masitah Ghazali, and Nova Ahmed. 2021. Religion and Women's Intimate Health: Towards an Inclusive Approach to Healthcare. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems* (Yokohama, Japan) *(CHI '21)*. Association for Computing Machinery, New York, NY, USA, Article 232, 13 pages. https://doi.org/10.1145/3411764.3445605

[86] Camille Nadal. 2022. *User Acceptance of Health and Mental Health Care Technologies*. Ph. D. Dissertation. Trinity College Dublin. School of Computer Science & Statistics. Discipline ….

[87] Camille Nadal, Corina Sas, and Gavin Doherty. 2020. Technology Acceptance in Mobile Health: Scoping Review of Definitions, Models, and Measurement. *J Med Internet Res* 22, 7 (6 Jul 2020), e17256. https://doi.org/10.2196/17256

[88] Timothy Neate, Aikaterini Bourazeri, Abi Roper, Simone Stumpf, and Stephanie Wilson. 2019. Co-Created Personas: Engaging and Empowering Users with Diverse Needs Within the Design Process. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems* (Glasgow, Scotland Uk) *(CHI '19)*. Association for Computing Machinery, New York, NY, USA, 1–12. https://doi.org/10.1145/3290605.3300880

[89] Helen Noble and Joanna Smith. 2015. Issues of validity and reliability in qualitative research. *Evidence-based nursing* 18, 2 (2015), 34–35.

[90] Giovanna Nunes Vilaza, Kevin Doherty, Darragh McCashin, David Coyle, Jakob Bardram, and Marguerite Barry. 2022. A Scoping Review of Ethics Across SIGCHI. In *Designing Interactive Systems Conference* (Virtual Event, Australia) *(DIS '22)*. Association for Computing Machinery, New York, NY, USA, 137–154. https://doi.org/10.1145/3532106.3533511

[91] Aisling Ann O'Kane, Abdinasir Aliomar, Rebecca Zheng, Britta Schulte, and Gianluca Trombetta. 2019. Social, Cultural and Systematic Frustrations Motivating the Formation of a DIY Hearing Loss Hacking Community. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems* (Glasgow, Scotland Uk) *(CHI '19)*. Association for Computing Machinery, New York, NY, USA, 1–14. https://doi.org/10.1145/3290605.3300531

[92] Chinasa T. Okolo, Srujana Kamath, Nicola Dell, and Aditya Vashistha. 2021. "It Cannot Do All of My Work": Community Health Worker Perceptions of AI-Enabled Mobile Health Applications in Rural India. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems* (Yokohama, Japan) *(CHI '21)*. Association for Computing Machinery, New York, NY, USA, Article 701, 20 pages. https://doi.org/10.1145/3411764.3445420

[93] Katie O'Leary, Jordan Eschler, Logan Kendall, Lisa M. Vizer, James D. Ralston, and Wanda Pratt. 2015. Understanding Design Tradeoffs for Health Technologies: A Mixed-Methods Approach. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems* (Seoul, Republic of Korea) *(CHI '15)*. Association for Computing Machinery, New York, NY, USA, 4151–4160. https://doi.org/10.1145/2702123.2702576

[94] Teresa K. O'Leary, Elizabeth Stowell, Jessica A. Hoffman, Michael Paasche-Orlow, Timothy Bickmore, and Andrea G. Parker. 2021. Examining the Intersections of Race, Religion & Community Technologies: A Photovoice Study. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems* (Yokohama, Japan) *(CHI '21)*. Association for Computing Machinery, New York, NY, USA, Article 698, 19 pages. https://doi.org/10.1145/3411764.3445418

[95] Işıl Oygür, Zhaoyuan Su, Daniel A. Epstein, and Yunan Chen. 2021. The Lived Experience of Child-Owned Wearables: Comparing Children's and Parents' Perspectives on Activity Tracking. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems* (Yokohama, Japan) *(CHI '21)*. Association for Computing Machinery, New York, NY, USA, Article 480, 12 pages. https://doi.org/10.1145/3411764.3445376

[96] Jessica Pater, Fayika Farhat Nova, Amanda Coupe, Lauren E. Reining, Connie Kerrigan, Tammy Toscos, and Elizabeth D Mynatt. 2021. Charting the Unknown: Challenges in the Clinical Assessment of Patients' Technology Use Related to Eating Disorders. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems* (Yokohama, Japan) *(CHI '21)*. Association for Computing Machinery, New York, NY, USA, Article 548, 14 pages. https://doi.org/10.1145/3411764.3445289

[97] Michael Quinn Patton. 1990. *Qualitative Evaluation and Research Methods*. SAGE Publications, inc.

[98] Claudette Pretorius, Darragh McCashin, Naoise Kavanagh, and David Coyle. 2020. Searching for Mental Health: A Mixed-Methods Study of Young People's Online Help-Seeking. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. Association for Computing Machinery, New York, NY, USA, 1–13. https://doi.org/10.1145/3313831.3376328

[99] Afsaneh Razi, Karla Badillo-Urquiola, and Pamela J. Wisniewski. 2020. Let's Talk about Sext: How Adolescents Seek Support and Advice about Their Online Sexual Experiences. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. Association for Computing Machinery, New York, NY, USA, 1–13. https://doi.org/10.1145/3313831.3376400

[100] Magdalena Rodekirchen and Sawyer Phinney. 2021. The Role of Social Infrastructures for Trans* People During the COVID-19 Pandemic. In *Volume 1: Community and Society*, Brian Doucet, Rianne van Melik, and Pierre Filion (Eds.). Bristol University Press, UK, 223–234. https://doi.org/10.51952/9781529218893.ch020

[101] John Rooksby, Alistair Morrison, and Dave Murray-Rust. 2019. Student Perspectives on Digital Phenotyping: The Acceptability of Using Smartphone Data to Assess Mental Health. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems* (Glasgow, Scotland Uk) *(CHI '19)*. Association for Computing Machinery, New York, NY, USA, 1–14. https://doi.org/10.1145/3290605.3300655

[102] Heleen Rutjes, Martijn C. Willemsen, and Wijnand A. IJsselsteijn. 2019. Beyond Behavior: The Coach's Perspective on Technology in Health Coaching. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems* (Glasgow, Scotland Uk) *(CHI '19)*. Association for Computing Machinery, New

York, NY, USA, 1–14. https://doi.org/10.1145/3290605.3300900

[103] Anke Samulowitz, Ida Gremyr, Erik Eriksson, and Gunnel Hensing. 2018. "Brave men" and "emotional women": A theory-guided literature review on gender bias in health care and gendered norms towards patients with chronic pain. *Pain Research and Management* 2018 (2018), 14 pages.

[104] Pedro Sanches, Axel Janson, Pavel Karpashevich, Camille Nadal, Chengcheng Qu, Claudia Daudén Roquet, Muhammad Umair, Charles Windlin, Gavin Doherty, Kristina Höök, and Corina Sas. 2019. HCI and Affective Health: Taking Stock of a Decade of Studies and Charting Future Research Directions. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems* (Glasgow, Scotland Uk) *(CHI '19)*. Association for Computing Machinery, New York, NY, USA, 1–17. https://doi.org/10.1145/3290605.3300475

[105] Hanna Schneider, Julia Wayrauther, Mariam Hassib, and Andreas Butz. 2019. Communicating Uncertainty in Fertility Prognosis. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems* (Glasgow, Scotland Uk) *(CHI '19)*. Association for Computing Machinery, New York, NY, USA, 1–11. https://doi.org/10.1145/3290605.3300391

[106] Velvet Spors, Hanne Gesine Wagner, Martin Flintham, Pat Brundell, and David Murphy. 2021. Selling Glossy, Easy Futures: A Feminist Exploration of Commercial Mental-Health-Focused Self-Care Apps' Descriptions in the Google Play Store. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems* (Yokohama, Japan) *(CHI '21)*. Association for Computing Machinery, New York, NY, USA, Article 404, 17 pages. https://doi.org/10.1145/3411764.3445500

[107] Standford Byers Center for Biodesign. 2022. Stanford Biodesign. https://biodesign.stanford.edu/

[108] Anselm Strauss and Juliet Corbin. 1990. *Open Coding*. Sage, 101–121.

[109] Anselm Strauss and Juliet Corbin. 1994. Grounded theory methodology: An overview. In *Handbook of Qualitative Research*, Norman Denzin and Yvonna Lincoln (Eds.). Sage Publications, Inc, 273—-285.

[110] Anselm Strauss and Juliet Corbin. 1998. *Basics of qualitative research: Techniques and procedures for developing grounded theory* (2 ed.). Sage publications.

[111] Anselm L Strauss. 1987. *Qualitative analysis for social scientists*. Cambridge university press.

[112] Sharifa Sultana and Syed Ishtiaque Ahmed. 2019. Witchcraft and HCI: Morality, Modernity, and Postcolonial Computing in Rural Bangladesh. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems* (Glasgow, Scotland Uk) *(CHI '19)*. Association for Computing Machinery, New York, NY, USA, 1–15. https://doi.org/10.1145/3290605.3300586

[113] Franziska Tachtler, Reem Talhouk, Toni Michel, Petr Slovak, and Geraldine Fitzpatrick. 2021. Unaccompanied Migrant Youth and Mental Health Technologies: A Social-Ecological Approach to Understanding and Designing. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems* (Yokohama, Japan) *(CHI '21)*. Association for Computing Machinery, New York, NY, USA, Article 541, 19 pages. https://doi.org/10.1145/3411764.3445470

[114] David R. Thomas. 2006. A General Inductive Approach for Analyzing Qualitative Evaluation Data. *American Journal of Evaluation* 27, 2 (2006), 237–246. https://doi.org/10.1177/1098214005283748

[115] Gareth M Thomas, Deborah Lupton, and Sarah Pedersen. 2018. 'The appy for a happy pappy': expectant fatherhood and pregnancy apps. *Journal of Gender Studies* 27, 7 (2018), 759–770.

[116] Lisa R Trainor and Andrea Bundon. 2021. Developing the craft: Reflexive accounts of doing reflexive thematic analysis. *Qualitative Research in Sport, Exercise and Health* 13, 5 (2021), 705–726.

[117] Sam Trendall. 2019. Gender bias concerns raised over GP app. https://www.publictechnology.net/articles/features/gender-bias-concerns-raised-over-gp-app

[118] Andrea C Tricco, Erin Lillie, Wasifa Zarin, Kelly K O'Brien, Heather Colquhoun, Danielle Levac, David Moher, Micah DJ Peters, Tanya Horsley, Laura Weeks, et al. 2018. PRISMA extension for scoping reviews (PRISMA-ScR): checklist and explanation. *Annals of internal medicine* 169, 7 (2018), 467–473.

[119] Chun-Hua Tsai, Yue You, Xinning Gui, Yubo Kou, and John M. Carroll. 2021. Exploring and Promoting Diagnostic Transparency and Explainability in Online Symptom Checkers. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems* (Yokohama, Japan) *(CHI '21)*. Association for Computing Machinery, New York, NY, USA, Article 152, 17 pages. https://doi.org/10.1145/3411764.3445101

[120] Anupriya Tuli, Shaan Chopra, Pushpendra Singh, and Neha Kumar. 2020. Menstrual (Im)Mobilities and Safe Spaces. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) *(CHI '20)*. Association for Computing Machinery, New York, NY, USA, 1–15. https://doi.org/10.1145/3313831.3376653

[121] Jayne Wallace, Anja Thieme, Gavin Wood, Guy Schofield, and Patrick Olivier. 2012. Enabling Self, Intimacy and a Sense of Home in Dementia: An Enquiry into Design in a Hospital Setting. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Austin, Texas, USA) *(CHI '12)*. Association for Computing Machinery, New York, NY, USA, 2629–2638. https://doi.org/10.1145/2207676.2208654

[122] Robert S Weiss. 1995. *Learning from strangers: The art and method of qualitative interview studies*. Simon and Schuster.

[123] Deepika Yadav, Prerna Malik, Kirti Dabas, and Pushpendra Singh. 2021. Illustrating the Gaps and Needs in the Training Support of Community Health Workers in India. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems* (Yokohama, Japan) *(CHI '21)*. Association for Computing Machinery, New York, NY, USA, Article 231, 16 pages. https://doi.org/10.1145/3411764.3445111

[124] Qian Yang, Aaron Steinfeld, and John Zimmerman. 2019. Unremarkable AI: Fitting Intelligent Decision Support into Critical, Clinical Decision-Making Processes. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems* (Glasgow, Scotland Uk) *(CHI '19)*. Association for Computing Machinery, New York, NY, USA, 1–11. https://doi.org/10.1145/3290605.3300468

[125] Renwen Zhang, Kathryn E. Ringland, Melina Paan, David C. Mohr, and Madhu Reddy. 2021. Designing for Emotional Well-Being: Integrating Persuasion and Customization into Mental Health Technologies. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems* (Yokohama, Japan) *(CHI '21)*. Association for Computing Machinery, New York, NY, USA, Article 542, 13 pages. https://doi.org/10.1145/3411764.3445771

[126] Haining Zhu, Zachary J. Moffa, Xinning Gui, and John M. Carroll. 2020. *Prehabilitation: Care Challenges and Technological Opportunities*. Association for Computing Machinery, New York, NY, USA, 1–13. https://doi.org/10.1145/3313831.3376594

A HEALTH FOCUS OF THE INCLUDED PAPERS

Figure 2: Word cloud representing the health focus of the studies reviewed.