2018 IEEE 31st International Symposium on Computer-Based Medical Systems

mHealth: a Privacy Threat Analysis for Public Health Surveillance Systems

Leonardo Horn Iwaya*, Simone Fischer-Hübner*, Rose-Mharie Åhlfeldt† and Leonardo A. Martucci*
*Department of Mathematics and Computer Science, Karlstad University, Karlstad, Sweden 651-88
Email: {leoniway,simofihu,leonmatu}@kau.se
†School of Informatics, University of Skövde, Skövde, Sweden 541-28
Email: rose-mharie.ahlfeldt@his.se

***Abstract*—Community Health Workers (CHWs) have been using Mobile Health Data Collection Systems (MDCSs) for supporting the delivery of primary healthcare and carrying out public health surveys, feeding national-level databases with families' personal data. Such systems are used for public surveillance and to manage sensitive data (i.e., health data), so addressing the privacy issues is crucial for successfully deploying MDCSs. In this paper we present a comprehensive privacy threat analysis for MDCSs, discuss the privacy challenges and provide recommendations that are specially useful to health managers and developers. We ground our analysis on a large-scale MDCS used for primary care (GeoHealth) and a well-known Privacy Impact Assessment (PIA) methodology. The threat analysis is based on a compilation of relevant privacy threats from the literature as well as brainstorming sessions with privacy and security experts. Among the main findings, we observe that existing MDCSs do not employ adequate controls for achieving transparency and interveinability. Thus, threatening fundamental privacy principles regarded as data quality, right to access and right to object. Furthermore, it is noticeable that although there has been significant research to deal with data security issues, the attention with privacy in its multiple dimensions is prominently lacking.**

Keywords–mHealth; privacy; data protection; threat analysis; mHealth data collection system; public healthcare.

I. Introduction

Mobile Health (mHealth) applications for *health surveys and surveillance* play a crucial role in developing countries, creating rich data repositories for public health decision-making. Applications for health surveys are usually known as mHealth Data Collection Systems (MDCSs), used by Community Health Workers (CHWs), replacing less efficient/reliable paper-based approaches [1]. The CHWs main task is to visit families at their homes in order to provide primary healthcare, but they also carry out surveys, collect the family's data and report it to the government. The main problem is that MDCSs initiatives (and many other mHealth systems) grow in a hopeful atmosphere but without much concern about the privacy implications [2]. There is not much debate about privacy and it remains unclear how to deal with the privacy issues inherent to health surveillance systems.

Such systems are used to collect, process and share sensitive data (i.e., personal health data), making informational *privacy* ***and*** *security* of paramount importance. In fact, in the past years much research has focused on the information security aspects of MDCSs [3]–[6]. That is, dealing with the key concepts of confidentiality, integrity and availability; commonly addressed by means of security mechanisms for encryption, authentication, secure storage and access control. Privacy, in turn, stands for fundamental rights and freedom of individuals to have their right to privacy with regards to the manipulation and processing of personal data. It overlaps with security, specially regarding confidentiality, but many other key properties have to be addressed (e.g., purpose specification, transparency, data minimisation, interveinability, accountability, information rights and consent) – fundamental differences that are further discussed in this paper. Which means, although privacy-preserving systems require strong security, security by itself is not enough.

Hence, this paper presents a privacy threat analysis for MDCSs, and also, discusses the main challenges and possible recommendations to counter the identified threats. To do so, a Privacy Impact Assessment (PIA) methodology was adopted (i.e., [7]) to evaluate the GeoHealth MDCS [1] as main use case. Among the main findings, we point out the privacy challenges related to data subject's[1] right to access and object (transparency and interveinability) are most problematic for existing MDCSs. Currently MDCSs consider just CHWs and health managers as *system users*. So, if the public were enabled access to their data through a personalised interface, a major redesign would be entailed.

II. Related Work

Many authors have already addressed the *security* issues in MDCSs (i.e., confidentiality, integrity and availability). In [6], the authors identified a range of security threats to MDCSs (i.e., ODK[2]) by means of a detailed threat modeling exercise based on surveys and interviews with technology experts. Also on security, the work of [4] and [5] propose two distinct security frameworks, specifically designed to cope with the networking and processing constraints that are inherent to mobile computing. For both frameworks the authors carried out independent security analysis but they greatly converge to the same security issues identified in [6]. Notwithstanding, there is still no comprehensive *privacy* analysis for MDCSs (i.e., distinct from a mere 'data security' analysis).

Regarding privacy for mHealth, the work of [3] proposes a broader *threat taxonomy* for mHealth privacy, categorising threats into three groups: (a) identity threats; (b) access threats; and (c) disclosure threats. However, this paper still addresses privacy in a rather narrow way. The resulting taxonomy is composed by privacy-related threats that greatly overlap with security (i.e., threats to confidentiality, integrity and availability). Thus, not contemplating privacy in its broader dimension

[1]A data subject is a natural person about whom personal data is processed.
[2]Open Data Kit (ODK) is a free and open-source set of tools which help organisations author, field, and manage mobile data collection solutions. (https://opendatakit.org)

2372-9198/18/$31.00 ©2018 IEEE
DOI 10.1109/CBMS.2018.00015

and overlooking many important Privacy Principles (listed in Section V-A).

III. Methodology

As aforementioned, this privacy threat analysis follows the PIA methodology defined by [8]. In brief, this PIA methodology supports project managers and developers to integrate privacy-by-design in their system development life-cycle. The method consists of a number of steps starting with a detailed system characterisation, followed by the definition of Privacy Principles and Privacy Targets derived from legal regulations.

The system characterisation is mainly grounded on the Brazilian GeoHealth MDCS [1]. Yet also includes other similar solutions (e.g., [4]). During the threat analysis stakeholders have to identify threats associated to each of the Privacy Targets. Lastly, all threats should be addressed with respective technical and/or non-technical control measures; the residual risk should be analysed; and steps towards its implementation should be specified. All in all, it is very similar to the NIST Special Publication 800-30 [9] for risk management. It is noteworthy however, that a full scale PIA greatly exceeds the scope of this paper and would lead to a rather extensive report. So, this paper focuses on the core of a PIA, the threat analysis and recommendation of countermeasures.

As legal regulation we adopt the upcoming European General Data Protection Regulation (GDPR) [10] (to be enacted in May 2018). This choice is based on two reasons: (1) scientifically, the GDPR can be considered as the state-of-the-art in privacy regulations and it can be also mapped to the work "A Taxonomy of Privacy" [11], regarded as *"the most complete list of privacy threats"* [8]; and, (2) the current draft of the Brazilian data protection regulation, in a broad way, is akin to the EU Directive 95/46/EC [12]. And even though the health and medical fields often have their own privacy-related regulations, GDPR compliance addresses the privacy problems to a great extent.

IV. System Characterisation - GeoHealth for Public Health Surveillance

MDCSs are mainly used as a tool for gathering primary health care information and tracking existing diseases, which drive health promotion initiatives in affected communities [13]. In brief, MDCSs support health surveys and surveillance, in which paper forms traditionally employed for data collection are replaced by a mobile device (see Figure 1). Using the device's communication capabilities data can be delivered in a faster and more reliable manner, speeding up the whole process of decision making. Key actors in this scenario are the CHWs whom are responsible for visiting families in their homes and for acquiring health-related information. During those visits the CHWs fill out electronic forms containing several questions designed for this specific purpose and loaded into the mobile device. Partially filled forms (i.e., forms lacking mandatory information) are stored in the device's memory and after completion are delivered to a server (e.g., via 3G/4G or Wi-Fi) together with the corresponding family's address. The server stores all received data in a database which can then be accessed and analysed using health management systems. (Due to the page restrictions, for further details we refer the reader to [1] and [14].)

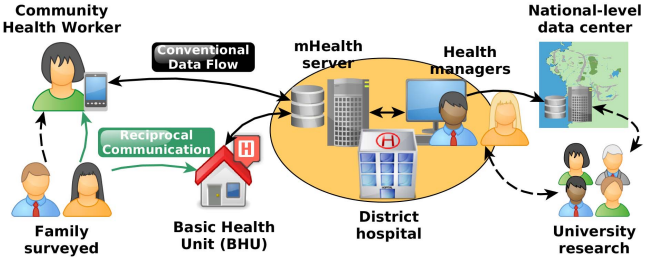

Figure 1. General process and data flow for MDCSs.

Since MDCSs are used for health surveillance of entire communities and deal with sensitive information, all data in-transit and at-rest should be protected from unauthorised access or modification. Privacy violations, such as illegal data disclosure to third-parties (e.g., health insurance or pharmaceutical companies) or even the theft or loss of smartphones can result in data breaches. Such incidents can affect the public trust on the system and discourage people's participation in health programs [15], or lead to lawsuits against those responsible for the MDCS. Therefore, it is key to understand and address privacy issues in order to successfully deploy MDCSs in public health programs.

V. Privacy Threat Analysis

This section presents all the components of our privacy threat analysis, i.e., the definition of Privacy Principles from the GDPR, the Privacy Targets to be achieved, followed by a compilation of threats that could prevent us from doing so.

A. Privacy Targets

Table I presents a list of Privacy Principles and Privacy Targets originally proposed by [8]. This list was then further examined regarding its applicability, meaning and exhaustiveness of the targets in the context of GeoHealth. For our analysis, we added two new targets P4.5 and P4.6, to emphasise consent withdraw and (electronic) copy of data. Targets P5.2 and P5.4 do not apply to MDCSs.

B. Identification of Privacy Threats

For each Privacy Target, we now systematically identify the threats that could impede us from achieving them. In a PIA, threats are primarily failures to comply with privacy laws or sector standards, which are outlined in the Privacy Targets. Threats can materialise when stakeholders are ignorant about privacy practices; when technologies do not have adequate privacy functionality; or, when processes and governance practices fail to protect privacy.

As aforementioned, privacy threats have been investigated before; both for mHealth in general [3] or specifically for MDCSs [4], [6]. This threat analysis is not only based on the opinion of privacy experts but also on existing scientific literature in the field. Furthermore, many of the threats are also found in the work of [7], evidencing that their privacy threat analysis (for RFID) also applies to other application scenarios. These known threats must be nonetheless reviewed and contextualised for our scenario. In what follows, we present an extensive compilation of privacy threats (applicable to GeoHealth) linked to Associated Privacy Targets (APTs).

TABLE I. List of Privacy Principles and Privacy Targets (adapted from [8]).

P1 - Data Quality
P1.1 - Ensuring fair and lawful processing through transparency
P1.2 - Ensuring processing only for legitimate purposes
P1.3 - Providing purpose specification
P1.4 - Ensuring limited processing for specified purpose
P1.5 - Ensuring data avoidance
P1.6 - Ensuring data minimisation
P1.7 - Ensuring data quality, accuracy and integrity
P1.8 - Ensuring limited storage
P2 - Processing Legitimacy (and Informed Consent)
P2.1 - Ensuring legitimacy of personal data processing
P2.2 - Ensuring legitimacy of sensitive personal data processing
P3 - Information Right of Data Subject (*ex ante* Transparency)
P3.1 - Providing adequate information in cases of direct collection of data from the data subject
P3.2 - Providing adequate information where data has not been obtained directly from the data subject (e.g., from third parties)
P4 - Access Right of Data Subject (*ex post* Transparency)
P4.1 - Facilitating the provision of information about processed data and purpose
P4.2 - Facilitating the rectification, erasure or blocking of data
P4.3 - Facilitating the portability of data
P4.4 - Facilitating the notification to third parties about rectification, erasure and blocking of data
P4.5 - Providing the ability to withdraw consent
P4.6 - Facilitating the provision of an (electronic) copy of data
P5 - Data Subject's Right to Object
P5.1 - Facilitating the objection to the processing of personal data
P5.2 - Facilitating the objection to direct marketing activities
P5.3 - Facilitating the objection to disclosure of data to third parties
P5.4 - Facilitating the objection to decisions that are solely based on automated processing of data
P5.5 - Facilitating the data subject's right to dispute the correctness of machine conclusions
P6 - Security of Data
P6.1 - Ensuring the confidentiality, integrity and availability of personal data storage, processing and transmission
P6.2 - Ensuring the detection of personal data breaches and their communication to data subjects
P7 - Accountability
P7.1 - Ensuring the accountability of personal data storage, processing and transmission

1) Threats to data quality: T1.1 Lack of transparency, missing or insufficient service information (APTs P1.1, P1.3) – This threat refers not only to (a) missing/insufficient information about the service, but also applies if, (b) the information is not easily accessible, (c) not easy to understand, (d) is outdated, or (e) the service's basic concept and purpose are not clearly explained.

T1.2 Lack of transparency, missing or insufficient privacy statement (APTs P1.1, P3.1, P4.1) – This threat occurs if (a) there is no *privacy statement*[3], or if (b) it does not sufficiently explain how data subject's data is processed. The privacy statement is also insufficient if, (c) it misses contact information about how to reach the operators, (d) it is difficult to read or to find, or if (e) there is no information about third parties that may receive data subject's information.

T1.3 Unspecified and unlimited purpose (APTs P1.1, P1.2, P1.3, P1.4) – This threat occurs when (a) the *purposes*[4] have not been specified before data processing starts, or if, (b) the purposes are not specific, limited and explicitly defined. Also, they have to be (c) adequately documented. It is also a threat if (d) data is not processed only for specific purpose, or if (e) the data processing does not strictly follow the specified purposes, (f) data processing leads to discriminatory profiling, (g) data processing facilitates mass surveillance unrelated to healthcare, or it (h) processing of data for another new purpose is incompatible to the original ones.

T1.4 Collection and/or combination of data exceeding purpose (APTs P1.4, P1.5, P1.6) – This threat happens when (a) the data collection is inadequate, irrelevant, or excessive in relation to the specified purposes, or (b) unnecessary to fulfil the overall aim of the processing. It is also a problem when (c) there are no measures in place to ensure *data minimisation*, or if (d) there are no measures to prevent *linking* of data sets.

T1.5 Missing quality assurance of data (APTs P1.2, P1.4, P1.7) – This threat arises when (a) data subjects are not thoroughly identified, (b) data is used without certainty of its accuracy and up-to-dateness, or (c) data is not properly validated before consolidation. Other threats may also happen if (d) data is updated without legitimate purpose, (e) there are no regular procedures to check if data is accurate and up-to-date, or if (f) personal data or profiles that are enriched by probabilistic algorithms that may lead to false judgements.

T1.6 Unlimited data storage (APTs P1.5, P1.6, P1.8) – This threat appears when (a) data is kept stored beyond the necessary time, or (b) kept stored when no longer needed in a form that permits identification of data subjects (i.e., not anonymous). Besides, other threatening cases are (c) the absence of erasure policies, and (d) the inability to exclude data that is no longer needed from "regular data processing operations" due to data retention rules.

2) Threats to processing legitimacy & informed consent: T2.1 Invalidation or non-existence of consent (APTs P2.1, P2.2) – This threat occurs if (a) there is no informed *consent*[5] made by data subjects. The consent can also be invalidated if (b) it was obtained based on incomplete/incorrect information, or (c) obtained on an offer of advantage or threat of disadvantage. Other particular cases are (d) the transgression of a relevant legal basis (e.g., consent, contract, legal obligation), or (e) the denial of processing in case of exemptions (e.g., "break-glass" emergency).

3) Threats to data subject's information right: T3.1 No or insufficient information concerning collection of data from the data subject (APTs P3.1, P3.2) – This threat emerges when (a) at the time of data collection, data subjects are not informed (e.g., orally or in writing) about the controllers identity and contact details, recipients of data (e.g., third-parties), what data is optional and consequences of not providing, and the existence of rights to access and rectify data. Other threats may occur if, at the time of data collection, the information is (b) not clearly visible (e.g., small pop-up box), (c) not easy to find (or hidden), or (d) not easy to understand.

4) Threats to data subjects right to access: T4.1 Inability to provide individualised information about processed data and purpose (APTs P4.1, P6.1, P6.2) – This threat arises when (a) at the time of processing, the operator does not provide an

[3]A privacy statement is a document that discloses some or all of the ways a party gathers, uses, discloses, and manages a customer or client's data.

[4]The purposes for personal data collection and processing.

[5]GDPR Art. 4(11): *'consent' of the data subject means any freely given, specific, informed and unambiguous indication of the data subject's wishes by which he or she, by a statement or by a clear affirmative action, signifies agreement to the processing of personal data relating to him or her [...]*

interface to the data subjects to efficiently identify what data about them is processed and used for. It is also a problem if (b) access is possible, but relevant information is missing (e.g., data being processed and purpose, data categories, recipients, logic involved in automated processing and decisions). And it is also a threat if (c) access is allowed with insufficient authentication, or (d) access and disclosure of personal data is not logged.

T4.2 Inability to rectify, erase or block individual data (APTs P1.7, P4.2, P4.4, P4.5, P6.1, P6.2) – This threat happens when (a) there is no implemented procedure for data rectification, erasure or block of individual data, or if (b) such operations cannot be done by controllers in case of illegal processing (e.g., consent withdrawn, data is no longer needed to fulfil the purposes). . Other more particular threats occur when (c) errors are not automatically rectified, (d) there is no procedure to delete data in stored backups, (e) rectification is allowed with insufficient authentication, or (f) such operations are not logged.

T4.3 Inability to notify third parties about rectification, erasure and blocking of individual data (APT P4.4) – This threat appears when (a) the operator has not implemented any procedure to notify third-parties when personal data has been rectified, erased or blocked.

T4.4 Inability to support data portability for individual data (APTs P4.3, P4.6) – This threat occurs when data subjects (a) cannot obtain a portable version of their personal data, (b) cannot request direct transmission to another service (if possible), or (c) cannot request an individualised (electronic) copy of their own data.

5) Threats to data subjects right to object: T5.1 Inability to allow objection to the processing of personal data (APT P5.1) – This threat takes place when (a) there is no implemented procedure that allows data subjects to object to the processing of personal data (i.e., based on compelling legitimate grounds).

T5.2 Inability to allow objection to the disclosure of data to third parties (APTs P4.4, P5.3) – This threat occurs when (a) the right to object is not raised to data subjects before data is made available, (b) the data subject does not have the opportunity to object before the disclosure, (c) the data subject is not informed about disclosure to third-parties, or (d) the operator has no procedure to notify third-parties about data subject's objections to data processing.

T5.3 Inability to allow objection to being subject to decisions that are solely based on automated processing of data (APTs P5.5) – This threat emerges when (a) the data subject cannot object automated decision procedures that are in the realm of the offered services (e.g., a wrong GPS location cause incorrect analysis or service provision).

6) Threats to data security: T6.1 Identity threats, misuse and leakage of data subject identities [3] (APTs P1.6, P6.1) – This threat happens if (a) insiders misuse data subjects' identity by forging data entries (e.g., to avoid work, steal identity, deny service). Other threats can also happen if (b) outsiders can re-identify individuals from anonymized datasets, or if (c) outsiders can eavesdrop communication channels to observe data subject's identity and personal data.

T6.2 Access threats, unauthorized access and modification of Personal Health Information (PHI) or Personal Health Record (PHR) [3] (APTs P1.4, P1.6, P2.1, P2.2, P6.1) – This threat occurs when (a) data subjects fail to express their consent consistently with their actual system preferences. In addition, it can occur if (b) insiders access or modify data subject's data by mistake (e.g., over-privilege or inadequate controls), or if (c) insiders commit intentional unauthorised access or modification. And similarly if (d) outsiders commit intentional unauthorised access or modification of data (i.e., curiosity, fraud or malice), or if (e) outsiders demand legal access to the data (e.g., through subpoena).

T6.3 Disclosure threats, unauthorized disclosure and data leaks of Personally Identifiable Information (PII) and PHI [3] (APTs P1.4, P1.6, P6.1, P6.2) – This threat emerges when (a) insiders accidentally disclose personal data (i.e., due to malwares, misconfigured application, improper password management), or if (b) insiders intentionally disclose personal data for profit or malice. It also happens if (c) outsiders intentionally disclose personal data for profit or malice (i.e., after unauthorised acces, T6.2), or if (d) outsiders eavesdrop the communication channel between mobile and server for traffic analysis and/or content decryption. Or even, if (e) the used mobile devices are lost or stolen, or if (f) the used mobile devices are improperly disposed (i.e., allowing outsiders to expose personal data, credentials and key materials).

T6.4 Denial-of-Service threats [4], [6] (APT P6.1) – This threat occurs when (a) users are unable to use the application (i.e., GeoHealth) due to misconfiguration or lack of network connectivity, or when (b) an denial-of-service attack prevents data from being uploaded (e.g., insiders/outsiders disrupts the communication channels, device or server functions).

T6.5 Inability to detect personal data breaches and communicate them to data subjects (APT P6.2) – This threat appears when (a) operators fail to detect infringements of data security in the system (e.g., no internal procedure for managing and reporting data breaches), or when (b) they fail to notify data breaches to the data subjects, or (c) they fail to notify competent authorities.

7) Threats to accountability: T7.1 Lack of accountability of personal data storage, processing and transmission (APT P7.1) – This threat occurs when (a) controllers fail to implement the measures to protect and safeguard data in their processing activities, or if (b) they are not able to demonstrate (at any time) compliance to data subjects, general public or competent authorities.

VI. Results and Discussion

This assessment led to the enumeration of the Privacy Principles and challenges pertinent to GeoHealth and other MDCSs that are summarised and discussed in Table II. The PIA RFID framework [8] supported us to vastly expand the taxonomy of mHealth privacy threat introduced by [3]. Most importantly, the current compilation of threats adheres to all dimensions of privacy as established in the GDPR, following a well-structured and reproducible methodology (i.e., using the Principles, Targets and Threats association scheme). As a result, our threat analysis offers a solid examination of the privacy issues for MDCSs used for health surveys and surveillance, thereby helping project managers and developers to understand and deal with privacy issues fairly.

Among the main findings, we noticed that existing MDCSs fall short particularly with respect to GDPR principles of

TABLE II. Summary of Privacy Principles, Challenges and Recommendations (critical items in bold type).

Summary of Privacy Principles and Challenges	Recommendations
(P1) Data Quality refers to: (a) fair and lawful transparent processing; (b) purpose specification; (c) data minimisation; (d) data quality (accuracy and integrity); and, (e) data retention. In MDCSs (and health systems in general), the topics of purpose specification, data quality and data retention are already well-known and most health managers are aware about issues and possible solutions. However, the use of appropriate strategies for transparency and data minimisation seems to be most challenging. Transparent processing can be achieved through accurate information about the system and privacy statements. Data minimisation can be done through multiple strategies to minimize collection, disclosure, replication, centralization, linkability and retention of personal data to the minimal/necessary extent [16].	- **Transparency-Enhancing Tools** - Guidelines for purpose specification - Fine-grained access control - Anonymisation and pseudonymisation - Data validation and integrity - **Automated data deletion**
(P2) Processing Legitimacy refers to: (a) legitimate data processing of sensitive data, typically obtained by means of a valid consent (i.e., freely given, informed, specific, involving an affirmative action); and also, (b) considering other relevant legal basis for using personal data, e.g., legal obligations, individuals' vital interest ("break-glass" emergency), or public interest. Informed consent for collecting and processing personal data is a well-known topic in the healthcare services. The main challenge refers to the informed consent about the technology and its privacy impacts, so that, individuals can base their consent on complete and correct information. Also, to *withdraw* the consent should be as easy as giving it. Besides, in the case of public health systems, a "freely given" consent cannot be obtained based on a threat of disadvantage (e.g., no access to healthcare).	- **Obtain informed consent** - **Check validity of consent**
(P3) Information Right of Data Subject refers to: (a) the provision of adequate information *before* data collection. Information about the MDCSs should be incorporated in the description of primary healthcare program. CHWs are the main link with the families and should be able to give such information, orally or in writing. This information should also be easy to find (e.g., project's website) and to understand (i.e., simple language).	- Accurate and up-to-date information about data controller, purpose, recipients and non-mandatory forms - Easy to find and to understand
(P4) Access Right of Data Subject refers to:(a) the provision of adequate information *after* data collection; (b) possible rectification, erasure and blocking of personal data; (c) data portability (e.g., electronic copy); and, (d) consent withdraw. That is a major challenge for MDCSs. Existing solutions do not provide family members access to the information system (i.e., a user interface). To access their personal data they would have to talk to the CHWs or go to the Basic Health Units in their coverage area. In summary, all aspects of access rights are problematic, and the provision of such functions would likely increase the cost of development and maintenance of MDCSs.	- **TETs for individualised information (e.g., privacy dashboards)** - Timely response to data subject's information requests and rectifications
(P5) Data Subject's Right to Object refers to: (a) objection to the processing of personal data; (b) objection to direct marketing; (c) objection to disclosure of data to third-parties; (d) objection to decisions solely based on automated processing; and, (e) right to dispute the correctness of machine conclusions. Similar to P4, existing MDCSs did not fully consider such rights during the design stage. The provision of an interface to send in objections would also considerably increase the project's overall cost. Apart from that, the families' data is not used for any marketing activities. And the use of data for research purposes follows *ad hoc* legal agreements among the public health sector and universities.	- **Provide an interface to send in objections** - Timely response to data subject's objections
(P6) Security of Data refers to: (a) personal information confidentiality, integrity and availability; and, (b) the detection and communication of personal data breaches. MDCSs should follow existing standards (ISO 27001) to protect personal data and *privacy-by-default* settings are highly recommended. And as mentioned, there are already security frameworks specifically designed for MDCSs (e.g., [1], [4], [5]). Such frameworks solve the technical challenge of what and how to use security mechanisms but one may have problems if they are not *correctly* implemented.	- Authentication and authorisation - Secure communication and storage (i.e., encryption) - Logging
(P7) Accountability refers to : (a) implementation of measures to promote and safeguard data protection; and, (b) being able to demonstrate (at any time) compliance with data protection provisions to data subjects, general public and supervisory authorities. MDCSs normally implement authentication, authorisation and logging mechanisms that can support accountability to some extent.	- Compliance with notification requirements - Logging

transparency and interveinability (i.e., (P1) Data Quality and (P4) Access Right of Data Subject). In brief, MDCSs do not consider data subject's personalised access to their data, and in fairness, they were designed to be accessed only by CHWs and medical staff. Thus, major redesign is required to add data subjects as system users, and to support interaction with a personalised interface (e.g., a privacy dashboard); somewhat similar to existing online medical records [17]. In this line, MDCSs would benefit from emerging Transparency-Enhancing Tools (TETs) (see [18] for a survey), that help to raise privacy awareness among data subjects, by allowing them to know about what data has been collected and processed about them and what are the potential privacy risks (e.g., discriminatory profiling, data breaches and leaks). Such changes however, greatly expand the system's attack surface (i.e., a new category of users with access rights) and increase the costs of software development and underlying infrastructure. The redesign of MDCSs therefore requires further feasibility studies, especially for projects running in low- and middle-income countries.

Other than that, *explicit informed consent* also has some particularities. Consent is a well-known requisite for providing medical treatment. In MDCSs, the consent given is for the handling of personal data. It refers to the data collection, processing, and access rights to the data and for what purpose, i.e., it is about technologies and systems. And just as important as obtaining valid consent, the operators should also facilitate its revocation. Because CHWs use smartphones for data collection it is later not very easy for data subjects to withdraw the consent, as they do not have direct computer access. Asking to revoke consent via telephone is not an easy solution either, as the data subjects have to be properly identified first. There should also be routines for allowing to revoke the consent only for selected purposes (e.g., a partial agreement). Existing literature on MDCSs do not discuss much about that, but there are guidelines to help project managers (such as [19]).

Features for *automated data deletion* are also missing in the existing MDCSs. That may be seen as a technicality that

is just not explored in the MDCS literature, but it is associated to the well-known right to be forgotten. For MDCSs, families may also change address or move to other communities, that would require formal procedures for automated deletion, as well as *data portability* (i.e., to send family's data to another health unit). Data subjects may also require deletion or blocking of sensitive data that can impact their privacy. Even more, medical conditions with strong genetic components can disclose information about the patient's relatives, i.e., impacting other people's privacy. This poses challenges for executing data subject rights, as the data may refer to more than one data subject, who all may have rights by different interest (e.g., one may want the data to be deleted while the other would like it to be kept). Routines are needed to handle such disputes and situations. In some cases, it may be possible to pseudonymise the identity of the person that wants his data to be deleted (e.g., in case of infections), while in case of genetic relations, this may not be possible.

This research also shows that the literature mostly focuses on the information security issues, solving just a fraction of the problem (i.e., (P6) Security of Data). Currently, there is lack of contributions on how to engineer privacy not only in MDCSs but actually for the area of mHealth in general [2], [20]. Project leaders and developers should pay special attention on the privacy aspects in the light of upcoming privacy and data protection regulations. PIAs, as shown herein, is one of the available approaches to incorporate privacy-by-design in the systems development life-cycle. The list of recommendations (Table II) also makes it clear that privacy cannot be dealt only with technical controls. In fact, almost half of the controls are non-technical, i.e., organisational procedures that should be put in place in order to achieve privacy and data protection.

Besides that, a few remarks can be made on the paper's limitations. Firstly, it was limited to a threat analysis instead of encompassing all the steps of a PIA, mainly due to page restrictions. Secondly, the recommendations are not fully detailed but we expect that they suffice to show concrete directions on how to control, mitigate or eliminate the privacy threats. Third, although the PIA RFID framework [8] offers a sound methodology, other PIA frameworks were already published (e.g., UK ICO's "PIA Handbook", CNIL's PIA manual, ISO/IEC 29134), so future comparative analyses among frameworks would be greatly desirable. Lastly, this threat analysis is relevant for MDCSs such as GeoHealth (i.e., very specific purposes and system characteristics). Other mHealth systems would require a new analysis, for instance if they use personal data for marketing/advertising or store data in the cloud, entailing different Privacy Targets and associated privacy threats.

VII. Conclusion

It is critical to understand the privacy implications that emerge in the advent of new technologies such MDCSs. Bridging the areas of privacy and health systems help project managers not only to be compliant with new regulations, but most importantly to respect the individuals right to privacy, essential for achieving high-quality healthcare. There is however a significant lack of research on privacy analysis for mHealth systems in general. This led us to contribute with a thorough analysis of privacy threats for the use case of MDCSs, as well as with the review and integration of existing approaches. Our future work involves the creation of more streamlined guidelines for MDCS's developers and the development of a privacy-enhancing framework, putting together mainly the recommended technical controls.

References

[1] J. H. Sá, M. S. Rebelo, A. Brentani, S. J. Grisi, L. H. Iwaya, M. A. Simplício Jr., T. C. Carvalho, and M. A. Gutierrez, "Georeferenced and secure mobile health system for large scale data collection in primary care," Int. J. of Med. Inf., vol. 94, 2016, pp. 91 – 99.

[2] A. Crowe, "Taking privacy and data protection seriously in M4D initiatives," in Proc. of 4th Int. Conf. on M4D Mobile Communication for Development - M4D 2014, General Tracks, ser. Karlstad University Studies, no. 2014:26, 2014, pp. 107–118.

[3] S. Avancha, A. Baxi, and D. Kotz, "Privacy in mobile technology for personal healthcare," ACM Comput. Surv., vol. 45, no. 1, Dec. 2012, pp. 3:1–3:54.

[4] S. Gejibo, F. Mancini, and K. A. Mughal, Mobile Data Collection: A Security Perspective. Cham: Springer Int. Pub., 2015, pp. 1015–1042.

[5] M. A. Simplício, L. H. Iwaya, B. M. Barros, T. C. M. B. Carvalho, and M. Näslund, "Secourhealth: A delay-tolerant security framework for mobile health data collection," IEEE J. Biomed. Health Inform., vol. 19, no. 2, March 2015, pp. 761–772.

[6] C. Cobb, S. Sudar, N. Reiter, R. Anderson, F. Roesner, and T. Kohno, "Computer security for data collection technologies," in Proc. of the 8th Int. Conf. on Inf. and Comm. Technologies and Development, ser. ICTD '16. New York, NY, USA: ACM, 2016, pp. 2:1–2:11.

[7] M. C. Oetzel, S. Spiekermann, I. Grüning, H. Kelter, and S. Mull, "Privacy impact assessment guideline for RFID applications," Bonn, Germany, 2011.

[8] M. C. Oetzel and S. Spiekermann, "A systematic methodology for privacy impact assessments: a design science approach," Eur. J. Inf. Syst., vol. 23, no. 2, 2014, pp. 126–150.

[9] G. Stoneburner, A. Goguen, and A. Feringa, "Risk Management Guide for Information Technology Systems," National Institute of Standards and Technology, Geneva, CH, Standard, Jun. 2002.

[10] EU Commission, "Proposal for a Regulation of the European Parliament and of the Council on the protection of individuals with regard to the processing of personal data and on the free movement of such data (General Data Protection Regulation)," 2015.

[11] D. J. Solove, "A taxonomy of privacy," U. Pa. L. Rev., vol. 154, no. 3, 2006, pp. 477–564.

[12] D. Doneda and L. S. Mendes, Data Protection in Brazil: New Developments and Current Challenges. Dordrecht: Springer Netherlands, 2014, pp. 3–20.

[13] WHO, "mhealth: new horizons for health through mobile technologies: second global survey on ehealth," The World Health Organization (WHO), Tech. Rep., 2011.

[14] L. H. Iwaya, L. A. Martucci, and S. Fischer-Hübner, "Towards a privacy impact assessment template for mobile health data collection systems," in Proc. of the 5th Int. Conf. on M4D Mobile Communication Technology for Development: M4D 2016, General Tracks, ser. Karlstad University Studies, no. 2016:40, 2016, pp. 189–200.

[15] J. G. Hodge, "Health information privacy and public health," The Journal of Law, Medicine & Ethics, vol. 31, no. 4, 2003, pp. 663–671.

[16] S. Gürses, C. Troncoso, and C. Diaz, "Engineering privacy by design reloaded," in Proc. of the Amsterdam Privacy Conf., 2015.

[17] H. Rexhepi, R.-M. Åhlfeldt, Å. Cajander, and I. Huvila, "Cancer patients' attitudes and experiences of online medical records," in Proc. of the 17th Int. Symp. on Health Inf. Management Research (ISHIMR 2015), 2015, pp. 24–26.

[18] P. Murmann and S. Fischer-Hübner, "Tools for achieving usable ex post transparency: A survey," IEEE Access, vol. 5, 2017, pp. 22 965–22 991.

[19] EU Commission, "Art. 29 data protection working party: Guidelines on consent under regulation 2016/679," 2017. [Online]. Available: http://ec.europa.eu/newsroom/just/document.cfm?doc_id=48849

[20] TrustLaw, "Patient privacy in a mobile world: A framework addresses privacy law issues in mobile health," TrustLaw Connect, a Thomson Reuters Foudantion Service, Tech. Rep., June 2013.