Ethical Considerations in the Conduct of Unregulated mHealth Research: Expert Perspectives

Catherine M. Hammack-Aviran, Kathleen M. Brelsford, and Laura M. Beskow

Introduction

Mobile applications and devices intended to monitor and promote health ("mHealth apps") are becoming ubiquitous, and the amount and scope of data they collect are constantly expanding.[1] These data can be valuable for an array of health-related research ("mHealth research"), including research conducted outside traditional academic settings.[2] However, ethical uncertainties arise when mHealth data are collected and/or used in research that is beyond the scope of federal regulations intended to protect human research participants ("unregulated research").[3] Confronting these challenges is essential to ensuring that end users (individuals who ultimately use an mHealth app/device)[4] are protected against the kinds of risks and harms that such regulations address, while supporting the conduct of potentially beneficial research.

To help inform these issues and contribute to the development of ethical policy and practice in mHealth research, we conducted in-depth qualitative interviews with experts from key stakeholder groups. We explored their perspectives on two hypothetical scenarios involving unregulated research using health, behavioral, and other data originally collected by commercial mHealth apps for non-research purposes.

Catherine M. Hammack-Aviran, M.A., J.D., *is an Associate in Health Policy in the Center for Biomedical Ethics and Society at Vanderbilt University Medical Center (Nashville, TN).* **Kathleen M. Brelsford, M.P.H., Ph.D.,** *is a Research Assistant Professor in the Center for Biomedical Ethics and Society at Vanderbilt University Medical Center (Nashville, TN).* **Laura M. Beskow, M.P.H., Ph.D.,** *is a Professor and the Ann Geddes Stahlman Chair in Medical Ethics in the Center for Biomedical Ethics & Society at Vanderbilt University Medical Center (Nashville, TN).*

Methods

Participants

We conducted in-depth qualitative interviews with experts from four key stakeholder groups:

- Patient and research participant advocates ("Advocate")
- Researchers who use mHealth technologies in their studies, including independent researchers and citizen and community scientists ("Researcher")
- Regulatory and policy professionals ("Regulatory")
- Mobile app and device developers ("Developer")

We identified potential participants based on leadership positions in prominent organizations, institutions, and studies; authorship of influential papers; and nominated expert sampling.[5] We used stratified purposive sampling to interview at least six experts per group, the minimum expected to reach saturation.[6]

Instrument Development

Based on our knowledge of the issues and in consultation with the larger research team, we developed a semi-structured interview guide centered around hypothetical scenarios (**Box 1**) involving two commercial mHealth apps collecting health, behavioral, and other data that may be shared for various purposes including research:

- "MoleStar," an app designed to support people diagnosed with or at high risk for melanoma
- An app designed to predict, detect, and prevent relapse during recovery from substance abuse ("Substance Abuse app")

The Journal of Law, Medicine & Ethics, 48 S1 (2020): 9-36. © 2020 The Author(s)
DOI: 10.1177/1073110520917027

Box 1

Hypothetical Scenarios

Scenario 1: MoleStar App *

Health Apps, Inc., has developed a comprehensive smartphone app called "MoleStar" for people at risk for melanoma or who have been diagnosed with melanoma. MoleStar includes this functionality:

- *Educational information* about melanoma, such as basic information about prevention, diagnosis, and treatment
- *Image capture tools* to help users monitor changes in moles over time
- *Disease management*, such as tools to track appointment times and locations, lab values, and medications. It also has tools to log daily physical and psychosocial wellbeing
- *Social support/networking*, such as tools to share information and communicate with others through social networks, and forums to communicate with other MoleStar users. It also enables optional participation in Health Apps. Inc, surveys about melanoma and MoleStar.

The primary purpose of MoleStar is to help people at high risk of melanoma to monitor moles between dermatologist visits, and also to provide support for people who have been diagnosed with melanoma.

All information captured by the app is automatically transmitted to Health Apps, Inc. and stored in a database to support further product development. Health Apps, Inc. also sells the data it collects to third parties for marketing, research, and/or product development purposes. Information about these uses is available via a link displayed at download; prospective users must click "I agree" in order to continue. The data use policies can also be accessed from MoleStar's "About" page.

Michael Lee is a computer engineer whose spouse recently died from melanoma. Eager to help others, he contacts Health Apps, Inc. to purchase the images captured via MoleStar along with the disease management information. His goal is to create a machine learning algorithm that can identify via the images early melanoma as well as moles at high risk of becoming cancerous.

Mr. Lee fills out an online form with basic information about himself and his intended use of the data. He signs an agreement saying he will use the data only for that purpose, and that he will not give or sell the data to others. After paying the data access fee, Mr. Lee's request is granted. Health Apps, Inc., provides him access to the images and data after removing direct identifiers (such as name, address, phone number) and replacing them with a code (which Health Apps, Inc., can link back to identifiers).

Scenario 2: Substance Abuse App **

Imagine now that Health Apps, Inc., made a *different app* that has a number of features intended to predict, detect, and prevent relapse in recovery from substance abuse, including:

- *Connecting with others* for support, *e.g.,* through discussion groups with other app users, video chats with counselors
- *GPS tracking* to detect when an individual is near a high-risk location (such as a liquor store). When near a high-risk location, the app causes the phone to ring and a number of recommended coping strategies display.
- A "*panic button*," which sends a text message to support prompting a response for assistance

Like the MoleStar app, Health Apps, Inc., captures and stores the data transmitted by the Substance Abuse app, and sells it to third parties for marketing, research, and/or product development purposes.

* Based on J.L. Bender et al., "A Lot of Action, But Not in the Right Direction: Systematic Review and Content Analysis of Smartphone Applications for the Prevention, Detection, and Management of Cancer," *Journal of Medical Internet Research* 15, no. 12 (2013): e287; G. Nasi, M. Cucciniello, and C. Guerrazzi, "The Performance of mHealth in Cancer Supportive Care: A Research Agenda," *Journal of Medical Internet Research* 17, no. 2 (2015): e9; B. Odeh et al., "Optimizing Cancer Care through Mobile Health," *Support Care Cancer* 23, no. 7 (2015): 2183-2188.

** Based on D. Gustafson, A-CHESS: Developing and Testing a Computer-Based Alcohol Use Disorder Recovery System, *available at* <https://center.chess.wisc.edu/research-projects/view/achess-developing-and-testing-a-computer-based-alcohol-use-disorder-recovery-system> (last visited September 23, 2019).

The Journal of Law, Medicine & Ethics, 48 S1 (2020): 9-36. © 2020 The Author(s)

Following pilot testing and refinement, the final guide (**Appendix A**) comprised questions about benefits and risks of the apps, approaches to notification/permission for research use, data access procedures, new primary data collection, offering individual research results, and data sharing and dissemination. We also asked questions about expectations for independent oversight, responses to which are reported elsewhere in this issue.[7]

Data Collection and Analysis

Prospective interviewees were invited by email to participate. Prior to the interview, we provided a study information sheet and a description of the MoleStar scenario. Interviews were conducted by telephone between October 2017 and February 2018 by two research team members and, with permission, audio-recorded and professionally transcribed. Interviews lasted approximately one hour and participants were offered $100 compensation for their time. The Vanderbilt University Institutional Review Board deemed this research exempt.

We uploaded transcribed interviews into NVivo 12 and used standard iterative processes to code and analyze the data.[8] See **Appendix B** for additional methodologic details. Narrative segments presented below are exemplary of frequently mentioned ideas unless stated otherwise; see **Appendix C** for additional examples.

Results

We interviewed 41 experts (**Table 1**) representing a range of demographic characteristics and holding diverse views on the basic level of risk associated with using MoleStar (**Table 2**).

Views on Notification/Permission Models for MoleStar

GENERAL NOTIFICATION

We asked about General Notification (i.e., a brief, broad disclosure) as a way to let people know that data collected by MoleStar could be shared and used for research. Interviewees discussed disadvantages and advantages for both end users and researchers. Over half noted that such notifications are often overlooked or ignored:

> When you put notices of any sort in connection with a cell phone app, people just click through them. People don't read them. People don't understand them. It's not an effective way of giving people notice. (09_Regulatory)

Many felt that, even when users read these notifications, they contain insufficient detail about what users are agreeing to or the risks involved, particularly regarding who may be conducting research and on what topics:

> People would have a top-level idea, 'it's going to be research,' but they may not understand all of what that entails. They may think, 'This is the company I'm giving it to and they're in control of it,' rather than, 'It's going to be sold and resold and resold to multiple other companies who I have no relationship with. I'm going to lose track and control of it.' … You don't know what types of research. You may assume that it's going to be research on melanoma, and not research on something totally unrelated… So, you may be giving consent for something that you didn't understand. (27_Regulatory)

Still, some emphasized the value in letting people know that their data may be shared for research, even if this disclosure is limited or may be disregarded:

> It's kind of like the general notification on a cigarette. 'Cigarettes are harmful to your health.' People don't really pay that much attention to it, but it's important that it is included. (11_Advocate)

Interviewees also identified low burden on end users as an advantage, describing General Notification as standard and easy to navigate:

> People are very familiar with the 'click here to consent' concept … that model of 'here's something, read it, click here to go read more'—that's become very familiar to all of us. (07_Regulatory)

Some remarked that General Notification is also efficient for researchers because, in addition to not constraining future uses, it likely increases the amount of data available:

> Most people will not think very much about the particular harms and risks. So the main strength is that they'll get a really lot of people agreeing to share their data. And many of those people—even if they knew all the [information], thought about it in more detail—would still probably agree. So, it's a strength in terms of getting your research numbers up in a very efficient way. (13_Advocate)

After discussing advantages and disadvantages, slightly less than half of interviewees indicated General Notification was acceptable for MoleStar and an

The Journal of Law, Medicine & Ethics, 48 S1 (2020): 9-36. © 2020 The Author(s)

Table 1
Participant Characteristics (n = 41)

	n	(%)
Category *		
Patient/participant advocate	10	(24)
Researcher	13	(32)
Regulatory/policy professional	9	(22)
Mobile app/device developer	9	(22)
Academic Degrees ^		
Bachelors	7	(17)
Masters	13	(32)
JD	5	(12)
PhD	16	(39)
MD	9	(22)
Geographic Region (U.S.)		
Midwest	2	(5)
Northeast	5	(12)
South	18	(44)
West	16	(39)
Gender		
Male	20	(49)
Female	20	(49)
Non-binary	1	(2)
Race		
White	30	(73)
Black, African American	3	(7)
Asian	5	(12)
> 1 Race	1	(2)
Not reported	2	(5)
Hispanic		
No	38	(93)
Yes	2	(5)
Not reported	1	(2)

* Many of our interviewees have multiple areas of expertise and could have been recognized as belonging to two or more categories of stakeholder groups; this table reflects the primary perspective for which we identified them as experts.
^ Reflects >1 degree per interviewee, as applicable

equal proportion said it was *not* acceptable (**Table 2**). The remainder said acceptability depended on content and/or presentation, i.e., the disclosure must be prominent, comprehensible, and sufficiently detailed for end users to understand what they are agreeing to.

MOST APPROPRIATE APPROACH TO NOTIFICATION/ PERMISSION

Interviewees expressed a wide range of opinions when asked about the *most appropriate* approach to notifying people that their data could be shared for research (**Table 3**). Some supported the use of General Notification, though often suggested the addition of subsequent reminders or optional supplemental information. Others advocated for a more active approach that would require some increased attention or additional action in response to the notification. Some suggested offering a broad yes/no choice, making research participation optional (i.e., not a barrier to using the app for its commercial purpose). Others went further, stating that providing multiple yes/no choices for various categories of data and researchers would be optimal. A few felt end users should be contacted and offered a choice about each specific research use.

Some interviewees did not describe a particular approach, but instead made other suggestions that could be applied to any approach, such as a tracking system to make transparent who has accessed end users' data:

> In an ideal world I would include that a company, when they share and sell the data, would need to have a site that users could access to see with whom their data has been shared. (13_Advocate)

Examples of other suggestions included the use of educational modules, quizzes, and ongoing two-way communication, leveraging the interactive nature of mHealth apps:

> Apps are designed to keep our attention, to maintain our engagement. [The most appropriate approach would be] to design consents and notices that are like that as well—real-time, updated, frequently communicating with you and letting you know not only how your data is going to be used and how it will be protected privacy and security wise… I think a consent-information type notice should happen regularly [and] keep you engaged in understanding the continued use of this data. (20_Regulatory)

The Journal of Law, Medicine & Ethics, 48 S1 (2020): 9-36. © 2020 The Author(s)

Views on Data Access Procedures for MoleStar

CODED DATA AND DATA USE AGREEMENT

In discussing MoleStar's data access procedures, interviewees again identified several advantages and disadvantages for end users, app developers and companies, and unregulated researchers. For end users, many commented favorably on the privacy protection afforded by replacing direct identifiers with a code in data shared with researchers, but anticipated that techniques to re-identify data would continue to evolve. A few noted that photos uploaded to the app may be identifiable…

> What if the melanoma is on somebody's face? What if the melanoma is on a part of the body that has really recognizable features that could link somebody to that image? We know that image recognition has gotten pretty sophisticated and the image, even without the name, address, and phone number, could be linked to individuals. (39_Researcher)

…or that an app may be "*accidentally collecting more data than it should*" (38_Developer), undermining efforts to conceal identity:

> De-identification is really hard to do, almost impossible if there's enough content. [For example], does their app scrub the latitude and longitude coordinates that are put into each

Table 2

Participant MoleStar Ratings and Opinions (n = 41)

	n	(%)
How worried should people be about using the MoleStar app?		
Not worried	15	(37)
Neutral	10	(24)
Worried	14	(34)
Other	2	(5)
Is MoleStar's use of General Notification acceptable?		
Yes	18	(44)
No	18	(44)
Depends / Other	5	(12)
Are MoleStar's Data Access Procedures acceptable?		
Yes	22	(54)
No	8	(20)
Depends / Other	11	(27)
Should MoleStar inform end users of Mr. Lee's additional survey questions?		
Yes	28	(68)
No	6	(15)
Depends / Other	7	(17)
Should individual results of Mr. Lee's research be offered to end users?		
Yes	11	(27)
No	18	(44)
Depends / Other	12	(29)

Table reflects responses to direct interview questions.

The Journal of Law, Medicine & Ethics, 48 S1 (2020): 9-36. © 2020 The Author(s)

https://doi.org/10.1177/1073110520917027 Published online by Cambridge University Press

Table 3

Most Appropriate Approaches for MoleStar: Illustrative Quotes

Disclosure, Permission	
General Notification	To say, 'Just a reminder, any information or conversations that you have on here, that data could be used for research,' so that is clear to people who might not understand what 'data' means [or] what counts as 'data.' (14_Researcher)
Active Notification	Building it into the experience ... rather than an easily dismissed notification or terms of use, that 'I just want to get into the app' stuff that people tend to ignore or bypass just so they can get to where they're going. Finding a way to make that a part of the experience of the app would be the best way to make sure people are reading and seeing and understanding what's being done with their data. (05_Developer)
Broad Permission	The ideal way would be after you got into the app, and set up your account, and opted in to using the service … to say, 'you have the option to share your data for research… The data is de-identified. There is minimal risk to you and we hope to achieve X, Y, and Z. Would you like to donate your data to research? Yes or no.' They would be able to clearly say no but still get the benefits of using whatever the app is. (16_Researcher)
Categorical Permission	You want to give people some options. It's not all or nothing. By giving them options, you're also making them think, and making them be thoughtful about what they pick. (30_Developer)
Specific Permission	The most appropriate way is to inform the patient every time their data moves to the researcher or moves for a purpose and give them a chance to opt out or opt in each time. It may not be the most ideal for the company, but it's much more ideal for the patient. (08_Advocate)
Data Access Procedures	
A. Technical security measures	
De-identification, general	I'm wondering if there's some way to strengthen that to include some indirect identifiers. I'm not sure, but I think that just that first level of [removing] direct identifiers is probably not adequate. But if somehow, like a second level could be masked, that would be much more acceptable. (13_Advocate)
De-identification, images	The fact that it's coded imagery is good. There's still a risk of the images including identification so I would hope they also had a process where there was a machine running algorithm or a human to glance through the pictures and make sure there wasn't anything identifiable in addition to coding all the data. (16_Researcher)
Storage, transmission, access	Instead of giving access to the data, you can give access to query the data. Obviously this is more work because you have to set up a system that allows people to run queries against the data… But you can do things like limit how much of the data they can see. (24_Developer)
B. Data use agreements	
Vetting applicants	It's not just verifying [an applicant's] identity, but verifying some of his credentials to make sure that his accepting this agreement is really understood and that he has actually the capacity to fulfill the confidentiality of what he's signing, of the data that's he's accessing. That would be a minimum. (39_Researcher)
Monitoring, reporting	Other parts of it would be making sure there was a constant communication going on so the researcher wasn't just taking the data and selling it or doing other inappropriate actions with it. But that there was some accountability for what the researcher is looking for and how the researcher's policies and practices are aligning with what they said in the data use agreement. (20_Regulatory)
Enforcement	There's got to be some meat to who's maintaining privacy and security and how they're doing it and what happens if for any reason we're not able to maintain your privacy and security. (10_Advocate)
Penalties, remedies	One of the biggest issues is going to be specifying and building in legally enforceable penalties for misuse and violation of the data use agreement. (30_Developer)
C. General	
Trust, transparency	If there's was just more of a process built into the company that will be finding unidentified risks as they emerge and address them… I'm talking about something that's a little bit more proactive that's built into the DNA of the company, that there's someone accountable. An ethics officer, if you will … someone who's accountable for the consumer experience and if they hear about a threat, that they're working with engineers to address the threat. (21_Researcher)

The Journal of Law, Medicine & Ethics, 48 S1 (2020): 9-36. © 2020 The Author(s)

> photo? If not, then the photos themselves, even though they were scrubbed, can still give away the location of the person … and the developer may not even know. (24_Developer)

With respect to data use agreements, advantages included the ability to set clear boundaries and expectations for what may and may not be done with the data. Interviewees commented that a formal agreement between an app company and an unregulated researcher may be helpful in clarifying appropriate uses of data in a manner that is "*a bit more of a legal process as opposed to just a handshake or a hand-off in the hallway*" (03_Regulatory). Still, many observed that the effectiveness of such agreements varies depending on the terms:

> Sometimes they're really good and really protective. Sometimes they're very minimal and more about protecting proprietary interest than requiring privacy and security protection. So, it would really depend on what was in that agreement. (20_Regulatory)

Interviewees also noted weaknesses associated with having to rely on the parties to understand, uphold, monitor, and enforce such agreements:

> The limitations are that there's no vetting of who Mr. Lee is. He could be anyone. Just because he put some information in an online form and pays a fee and clicks 'Yes, I will only use it for algorithms.' There's no checking who Mr. Lee is. Whether he's actually a real person, whether he's the person he says he is and what his background is for even understanding what he's signing, how committed he is to fulfilling this agreement… If this is just a person out there with no background in this, he might not even realize the implications of what he's signing and … what the risks are for the people that donated the data. That is a huge risk. (39_Researcher)

Thus, interviewees recognized that MoleStar's data access procedures offered valuable protections, though not a panacea:

> It is important to have access procedures generally, so that the data aren't just wide open that anyone could use. Having a data use agreement that a researcher has to abide by is important. You can imagine I'm skeptical about total anonymization, because that's very, very hard to do, and somebody who really wants to re-identify people probably can. But, making it harder is valuable, because you're reducing the risk that someone is going to re-identify and do something wrong. If the data use agreement also punishes that sort of thing legally, that can be helpful as well. (30_Developer)

Regarding the effect of MoleStar's data access procedures on app developers/companies and unregulated researchers, some interviewees were concerned that removing identifiers from data and imposing overly-restrictive data use agreements may constrain potentially beneficial research:

> You can't link records over time, or different people, or whatever it happens to be. The researchers get a fixed set of data and that's gotta make the research harder… The weakness is that as we learn more and more about health outcomes, identifiers are really important to integrate the information that's needed to ultimately develop effective approaches. (09_Regulatory)

A few were concerned that high standards for removing identifiers may over-burden app developers/companies:

> It's not trivial cost in staff time to be able to actually make the data available, to clean the data, to code the data, so that they are anonymized, to maintain and enforce the [agreements], possibly to train the researchers on how use and interpret the data. It's a real operation, and, for a small start-up company, it's probably not feasible. (30_Developer)

After discussing their views of MoleStar's data access procedures, just over half of interviewees said they were acceptable (**Table 2**), and about one-fifth found them unacceptable; the remainder believed that acceptability depended on other factors, such as the specific terms of the agreement and technical security measures.

MOST APPROPRIATE APPROACH TO DATA ACCESS PROCEDURES

When asked about the *most appropriate* data access procedures, interviewees identified several key attributes (**Table 3**). Many emphasized the importance of technical security measures, including the use of multiple methods to remove identifiers, special attention to latent identifiers (particularly in images), and alter-

The Journal of Law, Medicine & Ethics, 48 S1 (2020): 9-36. © 2020 The Author(s)

native approaches to storing and transmitting data aimed at decreasing and detecting unintended access and/or use.

Interviewees also suggested additional terms beyond what was described for the MoleStar data use agreement, such as requirements to follow standards for regulated research, report any protocol deviations to the company, and dispose of the data in a particular way (e.g., return, destroy) upon completion of the project.

Several described stewardship functions — such as vetting applicants, monitoring recipients, and enforcing the data use agreement — as essential. More generally, interviewees emphasized the importance of transparency and building and maintaining trust with end users:

> Google, Amazon, they all have my data. I have no idea what they're doing with it. I just give it to them every day I'm on the computer, or when I'm using my phone, but I have no idea… So I think there's a lack of trust in general, and I think the way to close that trust is to be as transparent as possible. (12_Researcher)

Views on Developments in Research Using MoleStar Data

NEW PRIMARY DATA COLLECTION

We also asked experts to consider what, if anything, end users should be told if Mr. Lee (an unregulated researcher) requested that additional questions, specific to his research purpose, be added to the periodic surveys that are part of the basic functionality of MoleStar. The majority believed that users *should* receive some information about such questions (**Table 2**), primarily echoing the themes of transparency, trust, and choice. Some further mentioned specific details of *what* end users should be told. Most said users should be informed of the particular research purpose, with some also suggesting disclosure of researcher's identity and/or qualifications, which they emphasized as particularly important in unregulated research:

> I feel like I'd wanna know, because my choice to participate in research that I consider more rigorous versus less rigorous — to me, it would matter. I wouldn't waste my time with something that's just some random person playing around… I don't wanna be anti-open science, but I'm struggling with how I feel about just anybody having access to data. (14_Researcher)

Some interviewees described *how* these questions should be presented. The majority proposed the MoleStar app should, at a minimum, emphasize that these questions are optional and participation is voluntary. Some went further, suggesting users must give express permission for primary data collection for new unregulated research:

> People would have to be notified that these are research questions, and now you're getting into the point where you probably do need informed consent. I think you have to start from the drawing board again at this point. You can imagine, if I told my IRB, 'I'm just adding new questions to the study, it's going to be great, don't worry about it,' I think they would turn pitchforks on me. (31_Researcher)

The few interviewees who believed users did *not* need to be alerted to these questions argued that end users were previously notified that any data — including survey responses — may be shared and used for research:

> Any survey questions, whether original or new, could be used for research… It's not practical to think you're going to pick out two questions and say, 'Oh, these two questions are for research and the others are just for the company — and, well, they might be used for research, too, at some point if somebody really wants to pay for it, or it might be used for marketing' and so on. (37_Regulatory)

Some interviewees answered, "it depends," saying that whether end users should be informed of new research-specific questions could vary based on the nature and content of the initial notification regarding potential research use, the new research purpose and the sensitivity of the data to be collected, and the potential for interfering with the app's primary purpose (e.g., notification fatigue).

OFFERING INDIVIDUAL RESEARCH RESULTS

We also asked interviewees to imagine that Mr. Lee believes his algorithm is highly accurate in identifying moles at risk of becoming cancerous and wants to provide his research results to individuals in the dataset whose images show such lesions. About one-fourth of experts believed that such results *should* be offered to end users (**Table 2**). Among them, most discussed notions of fair exchange and decency:

> Mr. Lee's research would not have been successful without the people who provided their data… That's like a reciprocity, not to mention just being a good human being. (05_Developer)

The Journal of Law, Medicine & Ethics, 48 S1 (2020): 9-36. © 2020 The Author(s)

Other common themes included a general right to information about oneself, as well as the prospect of providing direct health benefit:

> It may not be a guarantee that this will happen, but one of the key issues in a disease such as this ... is early identification and spurring people to action. Melanoma is probably one of the cancers that kills a lot of people, I would imagine because they aren't aware of it and don't act early enough... So, if he has an algorithm that's more than 50% accurate, it's imperative that he let the individuals be aware. (11_Advocate)

Several interviewees made suggestions for how to offer results, emphasizing the importance of explaining the uncertainty of the results and protecting end users' privacy.

About forty percent of experts believed that individual research results generated in unregulated research should *not* be offered. Most expressed strong reservations about the likelihood that unregulated research would be conducted with sufficient rigor, validation, expertise, or skills. Some were primarily concerned that the algorithm and results had not been verified:

> Having Mr. Lee, who is not a card-carrying researcher, if you will, make diagnoses and share them back would just scare people, and there may not be any basis for his conclusions unless there's some kind of further review of the quality of his work and the conclusions that he's drawn. (09_Regulatory)

Others questioned unregulated researchers' qualifications to interpret and communicate health information, noting that "[Mr. Lee] *is not a physician, so he's not qualified to offer clinical diagnoses, treatment, or prognoses*" (41_Researcher).

Some interviewees highlighted the lack of upfront consent; as one interviewee stated, offering individual research results should not happen "*unless the user explicitly opted in when using the app—and even then, the risks of giving them insights when [the algorithm is] not validated at scale is potentially very damaging*" (04_Developer).

A few characterized offering results as an unwarranted invasion of privacy, given that MoleStar users were likely already under the care of a medical professional:

> It's just not Health Apps's business or Mr. Lee's business to invade people's privacy and tell them that they could be dying, when we're already pretty sure they're seeking medical attention and they're already monitoring their health on this particular issue. I don't think that's appropriate at all. (26_Regulatory)

About one third of interviewees believed that whether individual research results should be offered would depend "*on what the people were told, and what they said they did or didn't want to know, and if any of those things came into the consent process*" (27_Regulatory), the unregulated researcher's qualifications, the reliability and validity of the algorithm/findings, whether the results were independently verified, and additional logistical considerations such as "*if information is going to be relayed back,* [by] *who and how will that be done?*" (03_Regulatory).

Views on Substance Abuse App

When discussing the Substance Abuse app, many interviewees described ways it differed from MoleStar. Some perceived these particular end users as "*a vulnerable population*" (25_Researcher) and expressed overarching concern about marketing use of the data: *"I don't know what's to stop them from selling the information to liquor companies so they can send discount liquor ads to all these people cause they know they'll be good customers"* (09_Regulatory). More generally, interviewees characterized the data as more sensitive:

> The risk of somebody finding out that you have potentially cancerous mole is not going to impact your job prospects or your relationships. Somebody finding out that you might have a substance abuse problem could have some pretty serious social side effects. (05_Developer)

They were particularly concerned about GPS information...

> That location data and GPS data is incredibly sensitive information about people. Once you have that, you can pretty much paint a very detailed portrait of a person, where they go throughout their day. (20_Regulatory)

...and the associated potential for legal jeopardy:

> Suppose we're having a custody fight over kids, or some kind of divorce, or some kind of family issue, information about where you've been ... you spend hours a day hanging around a liquor store, or in an outdoor drug market, what have you, that information can be used against you. It also can be used by police... This is a source

The Journal of Law, Medicine & Ethics, 48 S1 (2020): 9-36. © 2020 The Author(s)

> of information that could be used by a variety of people if they knew it was there. (09_Regulatory)

Even so, over half of interviewees believed the most appropriate approach to notification/permission would be the *same as or substantially similar to* what they said would be most appropriate for MoleStar, although some suggested providing additional details would be ideal:

> Considering that substance use and abuse is very fraught with stigma and also legal ramifications, this has to be a much more detailed description of the user's rights and responsibilities in terms of understanding what the company plans to do with their information. (36_Advocate)

A few advocated for a *different* approach to notifications/permissions from the one they described as most appropriate for MoleStar. A common theme was the increased sensitivity and identifiability of the data:

> I think this one is really problematic because people get fired from jobs, they get massively stigmatized, they can get arrested. There's all kinds of problems users can get into if they are identified either personally or by their location. There's a huge amount of trouble that this app can get people into… It could potentially do some great good, but the amount of harm that it could do is so substantial that people would have to be really aware of what they're getting into. (13_Advocate)

Another was concern about undue influence and end users' capacity to understand and agree to terms of use:

> You're talking about someone who's probably in really bad shape and they're looking for all the help they can get in the world… That person who's dealing with substance abuse, they don't care about your research. They care about staying clean and sober… Research is the last thing on their mind. (19_Advocate)

Responses were generally similar regarding data access procedures. The majority of interviewees advocated for the same approach as they had for MoleStar: "*I think the same general principles apply — certainly this is more invasive, but I think the same protections apply regardless*" (31_Researcher). Some suggested additional or heightened technical data security measures and more stringent data use agreement terms:

> The company needs to be much better educated on the sensitivity and identifiability of data and data security management, especially the GPS data. I think they have an obligation to hire experts and consultants to make sure they do a good job understanding which data they're sharing with whom and whether the data … has enough information redacted… If they're giving people access to sensitive data, they may want stronger contractual agreements. Sometimes that involves only wanting to share the data with people who have organizations to back them up, for example, people who are in academic institutions. It sucks that we have that divide between the citizen scientists and the institutions, but part of the reason it does exist is because those institutions are there to try to enforce more precautions and vetting and resources to enable good practices. (38_Developer)

Only a few believed the most appropriate approach to data access procedures for the Substance Abuse app would be substantially *different* from what they described as most appropriate for MoleStar, citing potential interest by parties other than researchers:

> It may put people at greater risk if there are no legal protections other than a data use agreement between the data provider and data user. That by itself can't protect you against all the legal jeopardy that's out there. (09_Regulatory)

Views on Sharing Data from Unregulated Research

A large majority of experts said unregulated researchers *should* seek to make their data and/or aggregate results available to others. Over half emphasized the importance of dissemination for advancing science:

> Why were the researchers interested in this in the first place, if it's not to draw conclusions, expand the scientific body of knowledge, generate new information? Why were they doing it if it's not to disseminate the results? (03_Regulatory)

Some focused especially on the importance of efficiency, referencing the value to the field of "*having access to all the research, not just selected research*" (09_Regulatory), to avoid unnecessary duplication and build on what others have learned.

Many interviewees cited ethical considerations, describing dissemination of data and aggregate results as "*the right thing to do*" (08_Advocate) and "*a basic ethical and civic responsibility that people have*

The Journal of Law, Medicine & Ethics, 48 S1 (2020): 9-36. © 2020 The Author(s)

to share what they learn when they're learning it based on engaging the public" (32_Advocate).

Some experts who believed that unregulated researchers should seek to share their findings nonetheless noted competing considerations that may act as disincentives:

> They may want to hide their results — they may want to profit from their results if they find something that's exploitable, they may want to hide their failures, they may want to hide their incompetence. There are reasons why people would not want to publish research. (09_Regulatory)

Regarding publication in particular, views expressed ranged from having a dedicated "*citizen science section, so everyone knows exactly what we're looking at*" (18_Researcher) to the expectation that all researchers should meet the same standards:

> If you're doing science, then that would be the goal: to have it subject to the scrutiny that science is subject to. I don't think we can reinvent science for a group of people who wanna circumvent standards. If you have something that you've found that's awesome, then yes, I think the only way that this can be considered legitimate is to submit it as a paper and publish it. (14_Researcher)

Only a few interviewees did *not* believe that unregulated researchers should seek to make their data and results available to others. These experts doubted that unregulated research would be appropriately designed, conducted, or validated; thus, they foresaw risks arising from sharing data and/or results from research lacking ethical or scientific rigor:

> We've seen it historically: unregulated or unfounded research claiming things really makes societal impact in a negative way. We find out five to ten years later that it was all bogus. It is tremendously important that in this day in age of fake news, and how fast things spread, regulation on these types of things are more and more critical. Not validated? Should not be published. (02_Developer)

Another theme was concern about stifling innovation:

> Mr. Lee, if he knew he would have to make his data available, is probably not going to invest in this because he wants to make money off it. He apparently wants to save the world from melanoma, but he has to make a living. If he's investing in this, he probably doesn't want to give his data away for that reason… On a voluntary basis, do I think it's a good idea? Sure. Do I think the government should come down and tell them they have to do it? I have a harder time with that. (27_Regulatory)

About one-fifth believed that unregulated researchers' ethical obligations regarding dissemination would depend on contextual factors such as competing obligations to stakeholders, limited resources, and privacy considerations.

Discussion

Although data collected by mHealth apps and devices may be valuable for research,[9] the use of mHealth data in research that is not subject to federal regulations for the protection of human research participants raises pressing ethical, legal, and social questions. Answering these questions is essential to ensuring that end users are protected against the kinds of risks and harms that federal regulations are intended to address, without overly restricting the conduct of potentially beneficial research.

The role of empirical data is to inform the development of ethical policy and practice. Qualitative, descriptive studies such as ours do not provide definitive answers, but rather illuminate critical issues from multiple perspectives. The interview results reported here suggest several key points to consider in the development and implementation of ethical approaches to unregulated mHealth research.

First, there are several possible approaches to informing end users about the potential for research use of their data, ranging from models that simply notify them, to those that allow for a broad yes/no choice, to those that provide for more detailed choices or even full informed consent for each specific research use. Consistent with prior recommendations,[10] experts in our study identified a range of factors that should be considered in selecting the most appropriate approach, including the level of risk involved (e.g., identifiability, sensitivity of data), burden on end users, practicability for research, and the effectiveness of the approach in informing end users.

Regardless of the approach selected, promoting and maintaining transparency and trust are essential to protecting end users as well as the research enterprise. Potential strategies include designing processes to actively call end users' attention to information about research use, developing easy-to-read disclosures that focus on key details most important to end users, and

The Journal of Law, Medicine & Ethics, 48 S1 (2020): 9-36. © 2020 The Author(s)

making additional information about research use available elsewhere for those who might be interested.

Second, when an unregulated researcher seeks to use data gathered originally by an mHealth app for non-research purposes, a variety of data access procedures could be used to help protect end users, such as data use agreements and careful removal of identifiers from datasets, as well as requiring independent oversight of the proposed research.[11] Designing and implementing the most appropriate combination of procedures should account for the ability of all parties to understand, uphold, monitor, and enforce the terms of data use agreements, and the associated need to meaningfully vet potential researchers' capabilities and resources.

Third, if an unregulated researcher were to request new data collection through an app whose primary purpose is not research, app developers/companies and researchers must decide what information (if any) should be disclosed to end users about this activity (e.g., the particular research purpose, the researcher's identity and/or qualifications). Important considerations include what end users were initially told about potential research uses of their data and the sensitivity of the new data to be collected.

Fourth, when determining whether or not — and, if so, how — individual results from unregulated mHealth research might be offered, vital considerations include what end users knew about and/or gave permission for when they downloaded the app, the validity and reliability of the results (particularly given the unregulated context), researchers' ability and expertise to accurately assess and communicate any health-related implications of the results, and end users' potential claims to information about themselves.

Finally, stakeholders should consider unregulated researchers' responsibilities with respect to sharing their results (e.g., making data available through centralized databases, publications). To make a positive contribution to generalizable knowledge, careful attention must be given to the rigor, expertise, and validity brought to the conduct and interpretation of the research.

The results of our study are subject to some limitations. Given the qualitative nature of our study and our interviewees' multiple areas of expertise, we did not attempt to analyze similarities or differences between stakeholder groups. The prevalence of and rationales for potentially differing perspectives between stakeholder groups may be an area for future research. We conducted these interviews in 2017-2018 with experts throughout the United States. While we believe many of our findings reflect fundamental ethical considerations, mHealth technologies, research using mHealth data, and privacy expectations among individuals and groups are rapidly and constantly evolving; ongoing attention to these issues over time is essential.

Acknowledgments

Research on this article was funded by the following grant: Addressing ELSI Issues in Unregulated Health Research Using Mobile Devices, No. 1R01CA20738-01A1, National Cancer Institute, National Human Genome Research Institute, and Office of Science Policy and Office of Behavioral and Social Sciences Research in the Office of the Director, National Institutes of Health, Mark A. Rothstein and John T. Wilbanks, Principal Investigators.

Note

The authors have no conflicts of interest to disclose.

References

1. *Research2Guidance*, mHealth App Economics 2017/2018: Current Status and Future Trends in Mobile Health, *available at* <https://research2guidance.com/product/mhealth-economics-2017-current-status-and-future-trends-in-mobile-health/> (last visited January 10, 2020); *Research2Guidance*, mHealth App Economics: Connectivity in Digital Health, *available at* <https://research2guidance.com/product/connectivity-in-digital-health/> (last visited January 10, 2020).
2. M.A. Rothstein, J.T. Wilbanks, and K.B. Brothers, "Citizen Science on Your Smartphone: An ELSI Research Agenda," *Journal of Law, Medicine & Ethics* 4, no. 3 (2015): 897-903; E.R. Dorsey et al., "The Use of Smartphones for Health Research," *Academic Medicine* 92, no. 2 (2017): 157-60.
3. M.A. Rothstein et al., "Unregulated Health Research Using Mobile Devices: Ethical Principles and Policy Recommendations," *Journal of Law, Medicine & Ethics* 48, no. 1, Suppl. (2020): 196-226.
4. D. Downing, M.A. Covington, and M.M. Covington, *Dictionary of Computer and Internet Terms*, 8th ed. (Hauppauge, NY: Barron's Educational Series, 2003): at 171.
5. E.E. Namey and R.T. Trotter, "Qualitative Research Methods," in G.S. Guest and E.E. Namey, eds., *Public Health Research Methods* (California: SAGE Publications, Inc., 2015): at 447.
6. G. Guest, A. Bunce, and L. Johnson, "How Many Interviews Are Enough? An Experiment with Data Saturation and Variability," *Field Methods* 18, no. 1 (2006): 59-82.
7. L.M. Beskow et al., "Expert Perspectives on Oversight for Unregulated mHealth Research: Empirical Data and Commentary," *Journal of Law, Medicine & Ethics* 48, no. 1, Suppl. (2020): 138-146.
8. K.M. MacQueen et al., "Codebook Development for Team-Based Qualitative Analysis," *Cultural Anthropology Methods* 10, no. 2 (1998): 31–6.
9. Rothstein et al., *supra* note 3; Precision Medicine Initiative (PMI) Working Group Report to the Advisory Committee to the Director, NIH, The Precision Medicine Initiative Cohort Program – Building a Research Foundation for 21st Century Medicine), *available at* <https://acd.od.nih.gov/documents/reports/DRAFT-PMI-WG-Report-9-11-2015-508.pdf> (last visited January 10, 2020); *All of Us Research Program*, Operational Protocol (March 28, 2018), *available at* <https://allofus.nih.gov/sites/default/files/aou_operational_protocol_v1.7_mar_2018.pdf> (last visited January 10, 2020); E.R. Dorsey et al., "The Use of Smartphones for Health Research," *Academic Medicine* 92, no. 2 (2017): 157-60.
10. E. Vayena and J. Tasioulas, "Adapting Standards: Ethical Oversight of Participant-Led Health Research," *PLoS Medicine* 10, no. 3 (2013): 1-5; S.J. Lynn et al., "Designing a Platform for Ethical Citizen Science: A Case Study of CitSci.org," *Citizen Science: Theory and Practice* 4, no. 1: (2019) 1-15.
11. Beskow et al., *supra* note 7.

The Journal of Law, Medicine & Ethics, 48 S1 (2020): 9-36. © 2020 The Author(s)

Appendix A. Interview Guide

- Did you have a chance to read over the study information sheet?
- Did you have a chance to read over the hypothetical scenario?
- Any questions about either?
- Is it okay to audio record the interview? ☐ Yes ☐ No

1. ☐ **App/device:** Before we get started, tell me about your experience or involvement in health-related research using mobile devices or apps.

   ☐ **Regulatory/policy:** Before we get started, tell me about your experience or involvement in health-related research using mobile devices or apps.

   ☐ **Researcher:** Before we get started, tell me about your experience or involvement in health-related research using mobile devices or apps.

   ☐ **Patient/participant Advocate:** Today we're going to talk about health-related research using mobile apps/devices—is that anything you've thought about for your advocacy organization, or you think your constituency group would be interested in?

[Scenario 1: MoleStar App]

2. Imagine that a close family member or friend is thinking about using the MoleStar app.

   2a. Considering the various features of the app, how would you describe to him/her the main benefits of using it?

   2b. Considering the various features of the app, how would you describe to him/her the main risks of using it?

   2c. Thinking about the benefits and risks you just told me about, how worried should people be about using this app?

   → *If participants ask for clarification regarding "worry":* Consider the benefits you've identified, along with the probability of the risks you identified occurring, and the magnitude of harms if the risks did occur.

   2d. Where would you put that on a scale from 1 to 5, where 1 is "not at all worried, by all means, use the app" and 5 is "very worried, think long and hard before using the app"?

   ---1----------2----------3----------4----------5--- Other: _____

In the hypothetical scenario, we described several different steps that were being taken to protect human subjects. One was general notification to inform app users that their data could be used for research. A second set of protections were about data access procedures, including a data use agreement between the company and the researcher, and the researcher getting only coded data. So, I'd like to ask you about each one of those separately.

3. Let's start with general notification of potential research use:

   3a. In your opinion, what are the strengths of general notification as a way to let people know their data could be used for research?

   3b. What do you think are the limitations of general notification as a way to let people know their data could be used for research?

The Journal of Law, Medicine & Ethics, 48 S1 (2020): 9-36. © 2020 The Author(s)

3c. How reassured do you think your family member/friend should be by general notification as a way to let people know their data could be used for research?

3d. Where would you put that on a scale from 1 to 5, where 1 is "not at all reassured" and 5 is "very reassured"?

---1----------2----------3----------4----------5--- Other: _____

3e. Okay, thank you. Now, for this next question, let's set aside strict regulatory requirements (or what IRBs typically expect or do), and just focus more generally about informing people that their data could be used for research.

Given the strengths and weaknesses you just described, do you think general notification is acceptable? In other words, regardless of whether you think this is best approach or not, is it an acceptable approach?

3f. Let's continue to set aside strict regulatory requirements (or what IRBs typically expect or do). What do you think would be the most appropriate approach to let people know that their data could be used for research?

3g. So, if we think about putting that into actual practice, tell me how you would see the pros and cons of that approach.

4. Now let's talk about the data access procedures, which included a data use agreement between the company and the researcher, and the researcher getting only coded data.

4a. In your opinion, what are the strengths of these procedures for protecting human subjects?

4b. What do you think are the limitations of these procedures for protecting human subjects?

4c. How reassured do you think your family member/friend should be by these data access procedures?

4d. Where would you put that on a scale from 1 to 5, where 1 is "not at all reassured" and 5 is "very reassured"?

---1----------2----------3----------4----------5--- Other: _____

4e. Okay, thank you. Now, again, let's set aside strict regulatory requirements (or what IRBs typically expect or do), and just focus more generally on data access procedures.

Given the strengths and weaknesses you just described, do you think the procedures described are acceptable? In other words, regardless of whether you think this is best approach or not, is it an acceptable approach?

4f. And continuing to set aside strict regulatory requirements (or what IRBs typically expect or do), what do you think would be the most appropriate approach to data access?

4g. So, if we think about putting that into actual practice, tell me how you would see the pros and cons of that approach?

Now let's talk about a few potential developments in Mr. Lee's research.

5. Let's say that Mr. Lee believes his algorithm would very likely work better with additional health information that is not currently captured by MoleStar. To help Mr. Lee test his theory, Health Apps, Inc. agrees to gather the information he needs by adding a few questions to one of its periodic surveys.

What, if anything, do you think users should be told about the purpose of these new questions?

→ *Probe*: Tell me more about the reasons why it's important or necessary that users be told that?

The Journal of Law, Medicine & Ethics, 48 S1 (2020): 9-36. © 2020 The Author(s)

6. As another possibility, let's say Mr. Lee *believes* that his algorithm *seems* to be highly accurate in identifying moles at risk of becoming cancerous. He becomes extremely concerned about individuals in the data set whose images show such lesions. He feels strongly that his research results should be conveyed to them so they can take action to avoid his wife's fate.

   In your opinion, should results be offered?

[Scenario 2: Substance Abuse App]

7a. What do you think would be the most appropriate approach to letting people who use this substance abuse recovery app know that their data could be used for research?

7b. What do you think would be the most appropriate approach to allowing researchers to access data from this substance abuse recovery app?

Appendix B. Consolidated Criteria for Reporting Qualitative Studies (COREQ)

Domain 1: Research Team and Reflexivity	
PERSONAL CHARACTERISTICS	
1. Interviewer/facilitator: Which author/s conducted the interview or focus group?	The interviews were conducted by Catherine M. Hammack-Aviran (author) and Kathleen M. Brelsford (author) under the leadership of the Principal Investigator Laura M. Beskow (author).
2. Credentials: What were the researcher's credentials? (*e.g. PhD, MD)*	Catherine M. Hammack-Aviran, MA, JD; Associate in Health Policy; female; law, bioethics Kathleen M. Brelsford, MPH, PhD; Research Assistant Professor, Health Policy; female; medical anthropology, methodology Laura M. Beskow, MPH, PhD; Principal Investigator; female; health policy, research ethics
3. Occupation: What was their occupation at the time of the study?	
4. Gender: Was the researcher male or female?	
5. Experience and training: What experience or training did the researcher have?	Each team member has at least ten years of research experience and extensive training in qualitative techniques, including the conduct of semi-structured interviews and qualitative coding and analysis.
RELATIONSHIP WITH PARTICIPANTS	
6. Relationship established: Was a relationship established prior to study commencement?	No relationship was established between an interviewee and interviewer prior to study commencement.
7. Participant knowledge of the interviewer: What did the participants know about the researcher? *(e.g., personal goals, reasons for doing the research)*	Prospective participants were provided with information about the funding source, the overall goals of the study, and the specific goals of the interview.
8. Interviewer characteristics: What characteristics were reported about the interviewer/facilitator? *(e.g., bias, assumptions, reasons and interests in the research topic)*	No interviewer characteristics were reported to interviewees.

The Journal of Law, Medicine & Ethics, 48 S1 (2020): 9-36. © 2020 The Author(s)
https://doi.org/10.1177/1073110520917027 Published online by Cambridge University Press

Domain 2: Study Design	
THEORETICAL FRAMEWORK	
9. Methodological orientation and Theory: What methodological orientation was stated to underpin the study? *(e.g., grounded theory, discourse analysis, ethnography, phenomenology, content analysis)*	We used an over-arching grounded theory research methodology. Within the overall framework, we employed an applied thematic analysis (including constant comparative analysis) to identify and refine meaningful categories.
PARTICIPANT SELECTION	
10. Sampling: How were participants selected? *(e.g., purposive, convenience, consecutive, snowball)*	Purposive and referral sampling, as described under *Methods: Participants.*
11. Method of approach: How were participants approached? *(e.g. face-to-face, telephone, mail, email)*	Prospective participants were invited by email. Purposive sampling was used to maximize demographic diversity.
12. Sample size: How many participants were in the study?	n = 41

13. Non-participation: How many people refused to participate or dropped out? Reasons?	Among the 71 individuals we invited to participate:
	• 9 declined
	  ○ 6 were unavailable (*e.g.*, too busy)
	  ○ 1 reported a potential conflict of interest
	  ○ 2 did not specify reason
	• 5 agreed but failed to respond to our attempts to schedule an interview
	• 14 did not respond
	• 2 scheduled an interview but did not participate (*i.e.*, "no-shows")
	No individual failed to complete an interview in progress (*i.e.*, no one dropped out), and no completed interviews were omitted from the dataset.

SETTING	
14. Setting of data collection: Where was the data collected? *(e.g., home, clinic, workplace)*	Interviews were conducted by telephone.
15. Presence of non-participants: Was anyone else present besides the participants and researchers?	No
16. Description of sample: What are the important characteristics of the sample? *(e.g., demographic data, date)*	The sample is described in detail under *Methods: Participants* and under *Results: Participant Characteristics* (Table 1).
DATA COLLECTION	
17. Interview guide: Were questions, prompts, guides provided by the authors? Was it pilot tested?	The interview questions and prompts associated with the data reported here are provided under *Methods: Instrument Development*; the entire interview guide is available upon request. The interview guide, including the hypothetical scenarios, was pilot tested with eight experts representing all four stakeholder groups.
18. Repeat interviews: Were repeat interviews carried out? If yes, how many?	No interviews were repeated.
19. Audio/visual recording: Did the research use audio or visual recording to collect the data?	With participants' permission, interviews were digitally audio-recorded. One participant declined to be audio-recorded and in two cases recording quality was poor. In these three cases, we relied on detailed notes for use during the analysis.
20. Field notes: Were field notes made during and/or after the interview or focus group?	Yes; interviewers took handwritten notes directly onto interview materials designated for each participant throughout the interview, as well as additional post-interview contextual notes when relevant.

The Journal of Law, Medicine & Ethics, 48 S1 (2020): 9-36. © 2020 The Author(s)

21. Duration: What was the duration of the interviews or focus group?	On average, each interview lasted approximately one hour.
22. Data saturation: Was data saturation discussed?	After coding 34 transcripts, no additional themes were identified to add to the codebook, suggesting saturation.
23. Transcripts returned: Were transcripts returned to participants for comment and/or correction?	No
Domain 3: Analysis and Findings	
DATA ANALYSIS	
24. Number of data coders: How many data coders coded the data?	Three
25. Description of the coding tree: Did authors provide a description of the coding tree?	The research team members (LMB, CMH, and KMB) met to develop an initial structural codebook. Research support staff then applied structural codes to all 41 transcripts; two team members (CMH and KMB) each independently reviewed approximately half of structural code applications to identify and resolve discrepancies. To develop content codes, two coders (CMH and KMB) each reviewed half of the structural code reports to generate an initial list of content codes. They then met with a third team member (LMB) to review codes and develop an initial content codebook. Each coder then completed approximately half of the content coding in an iterative fashion, with regular meetings to discuss any necessary additions or revisions to the codebook. Once all coding was completed, the coders then each reviewed the other's code applications to assess intercoder agreement and identify and resolve discrepancies.
26. Derivation of themes: Were themes identified in advance or derived from the data?	Themes were derived from the data.
27. Software: What software, if applicable, was used to manage the data?	NVivo 12
28. Participant checking: Did participants provide feedback on the findings?	No
REPORTING	
29. Quotations presented: Were participant quotations presented to illustrate the themes / findings? Was each quotation identified? *(e.g., participant number)*	Participant quotations are presented and each quote is identified by participant number.
30. Data and findings consistent: Was there consistency between the data presented and the findings?	Our manuscript integrates extensive use of direct quotes to provide evidence for each conclusion drawn.
31. Clarity of major themes: Were major themes clearly presented in the findings?	Major themes are clearly identified within distinct headings and subheadings.
32. Clarity of minor themes: Is there a description of diverse cases or discussion of minor themes?	There is substantial discussion of themes within each subheading, including diverse cases and minority opinions.

The Journal of Law, Medicine & Ethics, 48 S1 (2020): 9-36. © 2020 The Author(s)

Appendix C. Additional Selected Quotes

1. Level of concern about MoleStar app

A. Not worried	
Low risk, risk-benefit ratio	The risk is someone may get anonymous user data of yours—do you care? That is the question I would ask them. My recommendation with that is, it's not a big deal. *[Rating: 1/5]* (02_Developer) I would say that there's a minimal risk that the information that they put into the app would be shared. But I wouldn't really classify that as a big risk as compared to the benefit. *[Rating: 1/5]* (05_Developer)
No more risky than other apps, daily life	If someone's already using a lot of other phone apps, their day to day exposure for general risks of app usage is already there. This is not going to change that much. *[Rating: 1.5/5]* (38_Developer) I don't think they should be worried at all relative to other things they have to worry about. *[Rating: 1/5]* (15_Regulatory) I'm sure that what many people worry about, 'those health apps can link back to the identifiers and what if people know that I've got melanoma,' and 'is it a problem that I'm admitting to that by joining this app situation?' I don't know, I just think our information is all over the internet by now, so I don't see this as a huge privacy risk, unless you really don't want, for instance, your employer to know that you've got melanoma. *[Rating: 2/5]* (18_Researcher)
App user can control risk	I would say it's a very low risk situation and I would not be particularly worried. Because … depending on what you want to put out there, it's really up to you. You could put your mole pictures up, but most of your moles, even *you* wouldn't recognize if you saw them in another forum, so I'm not sure that it would be a huge deal if someone else did. There's not a high market for mole pictures that I know of… The other thing is that it depends on how much you want to connect the social media aspect of it and networking. I think if you're mindful of the content that you place out there, it's a really low risk situation. *[Rating: 2/5]* (08_Advocate)
B. Neutral	
Depends	The risk or the worry that somebody should have about this should be balanced. I think it's very personal. The issue with privacy in general is just so contextual. But if you ... feel that the benefits of tracking a mole or having this kind of accountability tool outweigh any kind of concern that you have about the data coming back to haunt you, then I would say not to worry that much. But I think that also depends on where you sit and the kind of person that you are, the kind of factors that go on in your life. *[Rating: 3/5]* (20_Regulatory)
C. Worried	
Data security; privacy and confidentiality	I guess the main risk would be all of the privacy issues—that their data, no matter how secure anyone says it is, any data that's virtual can be hacked. And that if we're going to network with other people, even if there's some kind of honor system, they still can't be sure those people are not going to share what they know. *[Rating: 4/5]* (13_Advocate) If there truly is nothing that says they won't share it with people who shouldn't have access to it, and they're not transparent about that—they won't ask your permission or anything like that—I would say, given that we're still at a time when politicians are fighting about whether everybody should get coverage or not, and irrespective of preexisting conditions, I would say they should probably be nervous. *[Rating: 4/5]* (30_Developer) What does the app do to actually physically protect people's data? Is it encrypted on the device? If you lose your device, will someone see it? Is it encrypted in transit? Is it encrypted how it's stored? What are the physical security risks and what are the protections this app has done to keep people's data safe? *[Rating: 5/5]* (31_Researcher) I think the risks are that you're wasting your time, that you are perhaps unwittingly contributing for the development of somebody's business that you have no stake in. It's possible that you're being exploited for financial gain. It's also possible that you would have an illusion of learning something that turns out to be false. You would be misled, potentially, by the app. I think wasting your time and being exploited and confused are all potential risks. Of course, there's also another risk—which is not specific to this app but I think it's worth mentioning because it comes up a lot in the world of apps—which is the app itself may be insecure and may contain a virus or privacy violating features that are concealed. The app may actually be malicious. *[Rating: 5/5]* (32_Advocate)
Sensitivity of data	I don't see any information here about how their photos would be de-identified, unless they do that themselves, or you have an employee that's … making their own personal judgment about what to blur out. And especially pictures of diseased tissue—very personal, very private. People don't typically want to share close-up shots of their weaknesses with the rest of the world, especially if they're not fully aware of what they agreed to when they downloaded it, that it will be sold. I just think that's a really touchy subject, right there in particular. And then also are the appointments and medications also being logged and being sold and associated with the pictures… That's all very rich, very personal data. *[Rating: 5/5]* (26_Regulatory)
Unanticipated data collection	The image capture: I have concerns about someone sharing information about themselves that they are not aware that they're sharing. If you think about that, a lot of people don't recognize that images are stamped. When you take a digital image, they're stamped with a lot of information about the camera that was used to take it, the device that was used to take it, where it was taken, GPS coordinates are on it. All of that sort of information is encoded, and you can post it online and someone can find you if you don't know that you have to turn that information off. I imagine a lot of users don't know that. They'll upload a picture, and completely be unaware that they can be giving out a lot of information about themselves just because they're uploading a mole that they're concerned about. *[Rating: 3/5]* (36_Advocate)

The Journal of Law, Medicine & Ethics, 48 S1 (2020): 9-36. © 2020 The Author(s)

Negative effects on heath, healthcare	I think that's a risk, because patients aren't trained to interpret. It potentially keeps them out of a physician's office who may have provided the interpretation and diagnosis. That's definitely a risk. *[Rating: 5/5]* (10_Advocate) There's a risk that it's going to create panic in individuals because any small change might be something that they're overly monitoring. It can create a drain on the system… And then people over-purchase those scanning tools, they overindulge in medical interventions because they think they've got a problem here, they've got a problem there. That has complications. It's expensive for the patient, it's expensive for insurance. There's just an over-utilization of medical services. *[Rating: 4/5]* (21_Researcher)
Providing incomplete or inaccurate information	The people creating these apps ... they may be researchers but the quality of the content and what they're developing and whether it's theoretically based or evaluated, is actually another major problem… The concern I would have is, this sounds great, but where's the proof in the pudding? And how well has it been evaluated? Who's created it? What's the quality of the data? Is it accurate information? Those would be some of the things for me personally, as a developer and a user, that I would want to know… And it doesn't really matter what the platform is, but we get caught up in thinking that it needs to be this cool tool and we forget about what the information is. *[Rating: 3/5]* (12_Researcher) It's sorta like garbage in, garbage out, right? If the tool is not very effective at doing what it's supposed to be doing, then the information that it provides is not gonna be very useful or actionable. If you have a tool that is supposed to accurately monitor changes over time, and it hasn't really been tested or well validated, then I don't think that tool's gonna be very useful... What we know from mobile health apps is that a lot of them are in sort of the pilot phase of development, especially apps that are geared towards diagnostics. That's a very unstudied area. *[Rating: 5/5]* (17_Researcher)
Trust-worthiness of app company; vetting process for third party access	The description of how Mr. Lee was able to get access to data tells me it wasn't rigorously vetted—who could ask for and get access to the data... the process for vetting requests for access to all the data … can amount to a leak because a bad actor could ask for it and presumably get access. It's not a leak in the traditional sense of stolen, explicitly. But it is under false pretenses. It could be leaked through the permissible licensing and approval process. *[Rating: 4/5]* (24_Developer) Especially since they're selling it to this guy who's creating a machine learning algorithm—he's probably not fully aware of the bias through this app, and so his machine learning algorithm is going to learn the exact same bias, which is that everybody who clicked on this app either is an exhibitionist in one way or another or in huge favor of science ... or they just didn't read it, but I don't think the people who would be willing to download and agree to this and use this app are necessarily an unbiased portion of the population. *[Rating: 5/5]* (26_Regulatory)

2. MoleStar's use of General Notification

A. Advantages	
Value in notification	I think that there's a great deal of strength in notifying people, but I think that it's not adequate, especially in the health context. (20_Regulatory) There's good evidence that many of these apps just don't put anything out there—you can download them and have no idea of what they're doing with your data, if anything. So, the fact [the General Notification] exists is a plus. The fact that it does outright say it sells data is a plus. They don't hint around it like other apps. It just doesn't seem to offer people that many protections. (31_Researcher)
Familiar, low burden	It's very common for us to have that in almost all computer related, internet things, so I would expect nothing different. I guess the strength is that it's a one-time notification and you're done. (08_Advocate)
Efficient for researchers, app companies	The app developers would probably like [General Notification] because it's harder to say no … that means you won't use the app. Some researchers might think, 'Oh, this is good. They'll slip in the willingness to share for research with the general agreement to the terms.' So, they might like that. Other researchers might be more troubled by that. (30_Developer)
B. Disadvantages	
Overlooked, ignored	[App users are] coming in for the gratification of getting access to the products and services. They're not thinking, 'If in two years I've been through cancer, they have … all of my pictures of that questionable thing on my lower back and my entire life story of how it started.' That's not what someone is thinking of the second they sign up … they're thinking 'I'd like to try that app.' (06_Advocate) I think it's pretty inadequate on an ethical basis. Primarily because I think that most people ... if it's just going to be embedded in the general terms and conditions, kind of the disclaimer part of an app, almost no one reads that. And if they want the app they're just going to agree without even reading it. So, I don't think that the message may even be received by most people. Even if it's received, if they do kind of skim through or even read the terms and conditions, then I think most people, or many people at least, won't really think through the implications for them. (13_Advocate)
Insufficient detail	Well, who's doing the research? Are you giving the data to some Silicon Valley man who wants to get rich quick or are you giving it to somebody else whose really gonna try to do something good? What's the end of it? 'Research' is so big and broad. (01_Advocate)
	It says basically that it's going to be used for research purposes and … that's not enough information. I would want to know, what kind of research are you talking about and how is my data going to be protected? How is it going to be de-identified in a way that I'll feel comfortable? (20_Regulatory)

The Journal of Law, Medicine & Ethics, 48 S1 (2020): 9-36. © 2020 The Author(s)

C. Acceptability	
Acceptable: Standard	It is accepted, so therefore it is acceptable. It is the standard… It's acceptable because it's what is used. (06_Advocate) I guess it's acceptable to most people. Yeah, I guess it's acceptable—I think that's how things are done in general. (41_Researcher)
Acceptable: Efficient for research	I think that there's a harm to research if we put friction on it, in the sense that we require asking permission, going back to people to asking permission, or we harm research by making it sound scary for their data to be used in research, and then it ends up not being used at all... Why remove the potential benefit of a greater good? (38_Developer)
Not acceptable: Not enough information	What research means to people is very nebulous. The public doesn't know what is research and what is not research. When you say generically 'research,' they may have a particular kind of organization or kind of purpose in mind. I would think research was, 'it's somebody at a university doing studies in the public good for knowledge of the entire world for these wonderful humanitarian purposes,' right? But there's also research [done by] government organizations, political organizations, for-profit companies. Anybody could have nefarious purposes and anybody could have good humanitarian purposes… I think research needs to be defined and who the parties are that do research need to be defined. (39_Researcher)
Not acceptable: Passive	There's simply an 'I Agree' button and people click through those all the time. We do that every time we download an app and it's thousands of lines of small text. That's certainly not sufficient… It's pretty passive and there's probably a lot of misunderstanding. (21_Researcher)
Not acceptable: Sensitive information	Sensitive information and vulnerable participants: risk, risk, risk. It's not cool. (26_Regulatory)

3. Most appropriate approach to disclosure/permission in MoleStar

Model	Quote
General notification (enhanced)	I think the only thing I would add under that approach is ... I think it's good if there's a link to a little bit of material that if they want to, you know ... if they want to they can just read a little and have some better understanding of the kind of things that are involved in that. That's really all I think is necessary though. (25_Researcher)
Active notification	If I had to say the most appropriate, I would say it would probably be something more interactive and informational than just relying on passive notification, whether that meant somebody answering questions or a filled-in miniature quiz to fill in the blank about what I was really agreeing to. Something that was more interactive than passive notification, but I would not think it rises all the way to requiring active informed consent, as we might think of it in the research context for a study like this. (03_Regulatory)
Broad permission	A screen where you can say 'agree' or 'don't agree' and it's not buried in a long text; like a one paragraph screen that is only related to that information—I think that's a great solution. It can't be one huge end-user-like agreement that you scroll through and somewhere it mentions that. (28_Developer)
Categorical, tiered permission	[The most appropriate approach would be] you can agree to researchers at an institution or researchers at a company, and that you might get four categories and you have to click them all, or if you only click one, then you go into a different bucket of where we can share it. (18_Researcher) When people are signing the e-consent, we would have a session talking about how is research defined in this new space … and, 'these are your options, these are different tiers of research.' … I think that would be the way to go. (34_Researcher)
Specific permission, consent	Ideally, you would be informed upfront that your data may be used subsequently, but that the user would have an opportunity to opt-in to that later use. You would get a notification when someone else requested access to the data, you would be able to see who is requesting access and for what purpose—what types of studies they want to do, or if it's for a marketing purpose. Then the user should have the option to opt in or opt out at that point. (04_Developer)
Other suggestions	[The most appropriate approach would include] notifying when somebody has access to the data and/or creating some type of centralized repository so you could check what's going on … a live, online website that allows you to log in and see who has accessed [your data], kind of like your credit report. (12_Researcher) There should be a set of questions designed to try and test their understanding… If you had explanatory videos in that person's language that are easy for them to watch, then if you had a way to ask them the questions that can give you some confidence that they actually understood the stuff that was explained to them, that would be the ideal way to do it. (24_Developer)

The Journal of Law, Medicine & Ethics, 48 S1 (2020): 9-36. © 2020 The Author(s)

	It could either be things like having education modules that people can watch... [Apps] can also interact with people too. They can send push notifications, and they can ask people, can we use your data in this purpose for this study? The beauty of this technology is information flow can go both ways, and we can certainly survey people in real time, or send them questionnaires. I don't think we've seen two-way pull of information. We've seen such a push to take people's information, but never check back in with them and see how they're doing or what they're feeling. (31_Researcher) I think any company that wants to maintain the trust of its users would disclose things like what is research, who is a qualified researcher, how do research requests get reviewed, is there a public list of them, what happens to the data, do I get the results back? … And reaching out to diverse populations, I think, would be a first step. The more graphical and the more friendly, the better. (40_Researcher)

4. MoleStar's data access procedures

A. Advantages	
Coded data: Privacy	The strength is the anonymity that the human subjects are given in that process… Because it's coded, and the person using the data on the other end for the research can't see the specific human being that that data is attached to. (05_Developer) Coding and anonymize the data, I think that's really important. Because, I think in general, people are quite willing to share anonymized data in a way that's gonna be helpful to learn about health... And to be quite honest, it's very important not only for the consumer, it's also important for the researcher… It protects … identity risks for the individual; it protects the researchers from liability. It's good all around. (25_Researcher)
DUA: Boundaries, expectations	Well, there is a ... what is it ... not road block; a speed bump, right? So, that there is a process, and Mr. Lee as a third party has to take an initiative to go through this process of seeking this data. The assumption is that someone would review this and they could say yes or no. They don't just simply, because you fill the form out, give them the data… Is it ideal? No. But is it perhaps limiting who gets access to data? At least you can track who is getting the data… Is this adequate? I don't know. If I was in the role of a researcher I would think this would be adequate. If it's my data, I would definitely would want more security. (12_Researcher) I guess not just leaving the data on the internet for anyone to take is also a step for protecting the users. Just open source library on Wikipedia. 'Look at all these pictures of moles!' They're not doing that; they're charging for it and gating it. (26_Regulatory) I think the strength is that there's at least an attestation that the requester is saying that they will only use it for creating this algorithm. That's not a real great strength, but it's at least one level of protection for the participant. (39_Researcher)
B. Disadvantages	
Identifiability: Re-identification	I don't think this is a very protective approach. I think we've all been victims of data breach at some point and this information is incredibly valuable to all kinds of entities, including the government, hackers, and others. And so, when you have an access approach that puts data out there with removing just a couple identifiers and using some kind of pseudo-anonymization with a code, that's really not difficult for a hacker to get both the code and to figure out who the people are and to sell that information. (20_Regulatory) You can just assume that if you remove a name and birthdate that that's all you need to do, but it may not be enough depending on the type of data that you collect… If you actually have hospital data or other health data that results in a pattern that allows you to recreate a person's medical history, then you can maybe re-identify that person, so that's a concern, however unlikely. (28_Developer) I know that [Mr. Lee] only is given a code, and they've removed direct identifiers, but when you have smartphone data, depending on what else is in this app, it's pretty easy to re-identify data in this day and age… If this was collecting what phones people have, geolocation information, IP addresses, and the photos it was taking, and there's things in the background in the photos—I think there's probably lot more identifying information that may not be so hard to put back together and figure out who these people are… It's unclear if removing the direct identifiers really is offering a protection that the company and Mr. Lee and the [end user] are expecting. (31_Researcher)
Identifiability: Unanticipated data collection	If there's real concerns with this app, it would be in whether it's accidentally collecting more data than it should, which is very easy to do with mobile apps, with apps in general. With mobile apps, they have to ask for permissions to access data on the phone. There's been incidents where apps accidentally ask for more permission than they need, or they start recording data that they didn't need to record or didn't intend to record, and that's all stuff that's not going to be covered in what the company has promised doing, but it's a general issue for apps in general. Then it really boils down to how much you trust the entities that have created the app to have done a good job in those respects. (38_Developer)

The Journal of Law, Medicine & Ethics, 48 S1 (2020): 9-36. © 2020 The Author(s)

DUA: Understand, uphold, monitor, enforce	People with enough data in today's environment can crossmatch databases and the like if they're really inclined, and the ability to re-identify even coded information is increasing … But having said that, you have to be motivated, and you have to be willing to violate your agreement, so I think that's not a great risk… An agreement is only as good as the party's willingness to abide by the agreement, and that's a reality with *any* agreement. (03_Regulatory) It seems that Mr. Lee has signed an agreement: he'll use data for research for his machine learning algorithm only, and he won't give or sell it to others. But, it's hard to know how he's going to protect the data. How is he going to keep it safe and secure? Is he going to leave it sitting on his desktop for someone else to access, to take? (31_Researcher) It is worrisome that there's no oversight that whoever's buying the data uses it in a way that they say that they will use it, as well. (41_Researcher)
Constrain, limit research	I'm kind of disappointed by how limiting the agreement Mr. Lee filled out is. It says using the data only for that purpose and not using it for other purposes nor giving data or selling data to others. That's highly restrictive. I would prefer the data be shared with contractual obligations that minimize the restrictions and … specifying things that cannot be done rather than saying you can only use it for this [one] purpose. (38_Developer)
Burden developers, companies, researchers	I've never been on the researcher end of this, but I would say that there's probably potential, I would think, for the data use agreement to be so specific that the person conducting the research, using the data, can't dig as deep as they might need to. … If data use agreements are really specific, then every time Mr. Lee wants to perhaps use an additional data point or something that wasn't in the original data use agreement, there's a lot of overhead, which sort of puts barriers in place to conducting research, which is a limitation. (05_Developer)
C. Acceptability	
Acceptable	I want them to be better than they are, but yes, they are acceptable. (06_Advocate) I don't see any red flags there. (14_Researcher)
Not acceptable	I think that we should have some kind of regulatory structures put in place before I would say that they were acceptable. (13_Advocate) It's not just verifying [Mr. Lee's] identity, but verifying some of his credentials to make sure that his accepting this agreement is really understood and that he has actually the capacity to fulfill the confidentiality of what he's signing, of the data that's he's accessing. (39_ Researcher)
Depends	I think the tricky part is that the devil's in the details of the data use agreement, and the devil's in the details in terms of the coding. (30_Developer)

5. Most appropriate approach to data access procedures in MoleStar

A. Technical security measures	
Reducing identifiability, general	So, there's anonymizing where you pull out identifiers like name, address, and phone number, but then there's also 'fuzzing' the data, where you remove even more specificity so it becomes more generic. You can turn date of birth into date range buckets that have like a 10-year spread or something like 18-25, 26-34, etc. And you can also do things like not share if there are too few instances of the data so that it wouldn't be aggregated in a way that allows [re-identification]. … These are the kinds of things you can do to really improve the anonymization and lack of specificity of the data to protect the participants. (24_Developer)
Reducing identifiability, images	One thing would be whether or not the images are ever available [to unregulated researchers]. Someone might have a really distinctive tattoo or something like that, maybe that's identifiable information. … [The most appropriate approach would be to] have some understanding of what's included in the image. So, when you take an image off of a cellphone … usually your location is embedded in it, for example. I guess you'd want some assurances about, are these the images or any of the kind of ancillary information also being taken into account in terms of inherent differentiability from the data, in addition to just having their name coded out. (25_Researcher)
Storage, transmission, access	Storing all data on the cloud and not on the phone is another safeguard you can do. I think that would be a good start. … I'd love to see the data being stored in a third-party service like Google Cloud or Amazon web services. If so, I'd like to see a business services agreement between them and the company… For a program like this, you're not going to have six cyber security experts on staff. So, … using an Amazon web service or a Google Health Cloud service is much more secure with much more diligence around it than having two or three people monitoring a server, in my opinion. (02_Developer) I just think the process of how [the app company] hands [the data] over, I think ideally perhaps there would be a data base where a researcher would get to log it and be able to access the data versus storing a local copy where theoretically the researcher says they won't share it, but it does have the ability to move on and get used in other projects without permission in the future… I think there would be a server that the data is stored in, researchers would have access to and do all their work on, so it could also be tracked to make sure they are not re-using the data and the images in other projects. (16_Researcher)

The Journal of Law, Medicine & Ethics, 48 S1 (2020): 9-36. © 2020 The Author(s)

	You could just contact a company and say, 'Hey, I'm really interested in seeing these pictures because I have a cool idea.' That's not the same as, 'Here is my fifteen page proposal about my idea and literature supporting it, and here is how I plan on going, phase one, phase two, phase three, and I really need your data for phase two of this project, and this I how I use it, this is how I would destroy it. Can I have some data?' Those are two different approaches. … I just think it's more important that people have the knowledge of research practices before they approach a company that requires my data. … It's not to exclude people who are not affiliated with academic or government research institutions. I don't think that research skill or capability is necessarily associated with these vocational affiliations. I think that it's more a matter of how well constructed and subset of information and he can use that and kind of ping it when he needs to do different analytics, or other kinds of statistical research, and look at it to build his algorithm. I think that's a more protective way to commence this kind of relationship. (20_Regulatory) Can you even detect that they violated [a DUA]? … One way that you can do this is they don't actually get to take data with them, but they can run queries against data. So, you hold onto the data and they only ever get a fraction of the data as the result then from the query they're running. But you're also using more traceability over what data they see. Another practice that you can do is you can essentially create little hidden markers in the data where you modify some data in ways that are traceable. So, if data showed up out in the wild that had this manipulated data you would know that this data was from you and that it was a breach of contract. (24_Developer)
B. Data use agreements	
Vetting applicants	The data use agreement would include some vetting of the recipient. It's not just anybody that walks in. Presumably, the custodian has an opportunity to vet the applicant to see if there's any warning flags go off or if it's a legitimate research use. They're controlling the flow of information, so I think that's a strength and an appropriate mechanism. (03_Regulatory) You could assume that Mr. Lee is well-intentioned and interested and abiding by all the other standards that medical research would ... but I suppose there's always going to be the Mr. Lees who aren't well-intentioned, not well-prepared, not abiding by standards that you would expect. … A dentist could start his or her own research study and basically do very bare minimum of what's required. … The difference is the dentist's licensing is on the line, so somebody has a stake in the quality, whereas Mr. Lee does not. (22_Regulatory) I think you want to figure out first, who is the person accessing the data? Where are they coming from? Have they, or their organization, received appropriate training for human subjects data? How are they keeping that data secure in terms of physical security, encryption, et cetera? What are their access rights, and what are their policies around keeping this data secure? Who's accessing the data? How are they going to delete it, or what are they going to do if they're done with it? How are they going to publish on it? Is Mr. Lee actually seeking to de-identify people? And publish who these people are? Is he seeking to do aggregate data analysis, personal data analysis? I think those are perhaps more the questions that need to be considered and vetted. (31_Researcher) You could just contact a company and say, 'Hey, I'm really interested in seeing these pictures because I have a cool idea.' That's not the same as, 'Here is my fifteen page proposal about my idea and literature supporting it, and here is how I plan on going, phase one, phase two, phase three, and I really need your data for phase two of this project, and this I how I use it, this is how I would destroy it. Can I have some data?' Those are two different approaches. … I just think it's more important that people have the knowledge of research practices before they approach a company that requires my data. … It's not to exclude people who are not affiliated with academic or government research institutions. I don't think that research skill or capability is necessarily associated with these vocational affiliations. I think that it's more a matter of how well constructed and thought through the proposed protocol is, and what they plan on doing with the results. (36_Advocate) So in a perfect world, I think that some research should be done on the purchaser of the data to make sure that they are who they say they are and then continued follow up with that purchaser to make sure that they're using the data in the manner that they had said that they would use it as well. I don't know about private individuals accessing data. I would want it to be from an institution, someone that's representing an institution, an academic institution or even an industry—maybe not private people, like individuals. … I guess I'd be less comfortable with it because first of all, you don't know what they're actual intention is for use of that data and how they will sort of get to the outcome that they want. In academic science, we go through a lot of procedures to make sure that the way that we analyze data and sort of release the findings from that data are really ... it's a lot of rigor. And I don't know that private citizens would hold themselves to the same standards or that they would be able to sort of match those standards that we think about in science. Also, while we'd like to believe that everyone has the best of intentions, some people don't. And I think sometimes data can be manipulated to point to a certain outcome that someone may want to see. It's harder to regulate that if there's no sort of academic or scientific oversight. And then the other thing is although you've signed an agreement saying that you won't sell the data to others, there's no way to guarantee that. (41_Researcher)
Monitoring, reporting, enforcement, remedies	I would want the agreement to have certain elements in it. I want it to have a purpose statement, I want it to say about redisclosure, I want it to say about reidentification, I want there to be some penalty or liquidated damage or something that binds the person, or makes them feel responsible. I would want some kind of not-too-burdensome reporting requirement if something goes wrong or something changes… You might say that the data has to be returned or destroyed after the purpose is carried out. (27_Regulatory) I'd want data subjects to be third party beneficiaries of the [DUA], which means, basically, that if I'm a data subject here and somebody does something untoward with the data, I can sue the researcher. (09_Regulatory)

The Journal of Law, Medicine & Ethics, 48 S1 (2020): 9-36. © 2020 The Author(s)

C. General	
Trust, transparency	In my perfect world, being able to subscribe and follow your data wherever it goes. … sort of like Twitter, I can follow a hashtag or I can follow a person and then I get updates whenever they do something. … I can choose what I wanna follow. If I know that my data's being used in a particular project, I can subscribe to that project and follow that project. … if there was a way that a third party organization could allow me to say I want to subscribe to this and follow this through its lifecycle, then that would be great. I don't wanna stumble upon research in some journal that someone posted a link to on Twitter in 10 years that talks about me and I didn't even know I was doing it. Ethically you've checked the box, but morally meh, that's my blood, sweat and tears—or this case, images—that you're using to inform research, and I'd like to know what you're learning. (01_Advocate) If a company has to put up on it's website—and it's accessible to users—who they're selling and sharing the information with, I think just that transparency might make them more careful about who they're sharing with. I think that would be important. (12_Researcher) The classic way to solve that is that you make it more and more about informing individuals rather than trying to make the system itself trustworthy, such that even if someone is not fully informed, they know because of the other people watching out for them that it is trustworthy, which then allows us to feasibly get more representative data and to be able to use the data and get broader data access. … I mean, some ways that I've heard people talk through this is that there is ultimately a central repository of data, or you can play around with the logic of block chains or other things such that you can build systematic ways of knowing, of giving data providence and using that as sort of a logical check... So, there's basically the human element … but then there's also the technological element of the tools that are increasingly being built to have an open infrastructure where everyone has the sort of real running ledger on where information is, even if they don't necessarily have access themselves. There's a check on this, which can then be the policing structure for those who might actually even be trustworthy at one time but end up falling down against their own better graces and doing something nefarious with the information. … I do really see the value of trying to make it easy for people to be able to get access to data, thinking about this in terms of societal benefit. The tension is, can we actually be trustworthy to the people who are giving us the data? (29_Researcher) Capture as many details as possible about what's happening with the information and the data in case people ask. That can help them to say, 'here are some instances of how people have used data and so on.' You can even tell people to put it on your website, … 'some of the ways we've used data that we think will be helpful to people.' … I would think just something to gather more data on what's happening and what takes place. (37_Regulatory) I'm not convinced that making a promise to people means that the entity will follow through on it, and I have very often seen researchers make promises about security and not follow through with it at all, sort of like a blisteringly, astonishingly insecure, 'what are you doing, do you understand that that's not okay?' way. They don't know data security, and they make huge promises encouraging entities to make sort of promises that's troubling to me, then they don't follow through. (38_Developer)

6. Should app users be told the purpose of new questions on periodic surveys?

A. Yes	
Yes: General	From the perspective of the user, I think it's important to develop that trust and transparency of why you're adding these questions and what is it going to be used for. As a researcher, we always want questions, and we neglect to think about who and why we're collecting that, we're just trying to get as much data as possible. There's a dichotomy and if there's a way of perhaps breaking down that difference, I wonder if you could get a better win for everybody. (12_Researcher)
Yes: What users should be told	They should be explicitly told why the information is being added and that it is for third party research—and that someone is paying money to access the data. (04_Developer) They should be told expressly what's being done, the identity of the researcher, and the purpose of the research, and possibly the qualifications of the researcher. If anyone comes along and is willing to pay the fee—and this is that I'm concerned about—if this a commercial for-profit company that will sell the data to anyone who waves a nickel in their face. The Mr. Lee you describe here sounds like a reasonable kind of guy, maybe I trust him, maybe I wouldn't, but at least he sounds like he has reasonable motivation. The next guy who comes along may just be doing marketing. (09_Regulatory) There'd need to be some transparency. You'd need to have some kind of write up about who's asking, what their intentions are, why they need it, and what their qualifications are. (10_Advocate) That, in my opinion, should be an update to the terms—'we're adding these and this is the reason why', just as we would do in a clinical trial, an amendment with an updated consent form. You have to go to the users first and re-consent based on this new information you're collecting. This goes beyond what the members agreed to in the first place … there's a specific purpose for this and the questions are a means to an end, and that end is different from what they originally agreed to. So, you must be transparent about what the new data would be and what it would be used for, because it's likely going to be different from what they signed up for in the first place. (23_Developer) My sense would be they should be told who's using these data and why in as simple language terms as possible. (29_Researcher)

The Journal of Law, Medicine & Ethics, 48 S1 (2020): 9-36. © 2020 The Author(s)

Yes: How users should be told – Active	We should consent them again because now we're changing the rules a little. So, we should ask people if they are willing to provide that additional information and tell them the truth. That would be an ethical way to approach it. (35_Advocate) As soon as they pop up, these questions—I think there's no need to alert them before they see these questions, but when they show up—it would be fair to have one or two sentences explaining why they are asking these questions and that they are voluntary. (28_Developer) You have to tell people these are research questions. You'd probably have to have people opt into it. (31_Researcher)
Yes: How users should be told – Passive	I think you could just basically put a reminder, you don't have to update your General Notification or anything like that. (02_Developer)
B. No	
No: General	Well, they weren't told any information about the previous questions in terms of research, so from a methodological perspective it would look really weird to have one question, 'The reason we're asking it is because...', people would look at it and be like, 'Wait, what?', cause then it would call attention to the fact that there is this secondary use taking place… I don't think it would be smart to call it out that this is why we're asking these questions. I think it would be something that the researcher and/or the company would ... have to talk about... I would like to see what kinds of research is taking place with the data that's generated from this app. Sharing that information with the customers, participants helps alleviate some of that distrust or lack of knowledge about really what's going on, so I think it would be periodic updates about being very transparent... I think it's just a matter of calling attention to it in other forums, not specifically here's a survey and add this one question and say, 'this particular researcher wants this for…', because I just think that would look weird. (07_Regulatory)
C. Depends	
Nature and content of the general notification	My understanding of the scenario is that people have already been informed that this might be used. Now you get back to what do they really understand? Did they bother to read the notification? If we accept that they actually do understand that this is being used for research, then whether there's one question or five questions, I don't know that that changes the calculus a whole lot. I don't know that it forces health apps to turn the world upside-down, to take a different approach back to the participants, to the users. What did people understand when they originally agreed. (03_Regulatory)
New research purpose and sensitivity of data to be collected	I don't know that it's necessary to inform people unless you are creating an entirely new set of data out of the answers to these questions. So, in some ways, I guess it depends on the sensitivity of what you're asking. If you're asking somebody how frequently they use this app and do you think it has been effective … I think those would be fine. They're related to a reasonable expectation of how the app functions and to the company's interest. But if it started to get outside of that and say, do you have members of your family who have had melanomas? Then you're getting into more sensitive information that raises the identifiability of the person more. In which case I would say there should be some kind of notice, maybe not necessarily identifying the researcher and what they're trying to do but some kind of general thing that says, "This a supplemental question for research purposes. Here's how they might impact you. You can agree to answer these or you can skip them." (20_Regulatory)
Interfering with app's primary purpose	There's a big disutility in foisting too much information every time something new is added on consumers who are trying to just use this to monitor their own health. (21_Researcher)

7. Offering individual research results

A. Yes	
Right thing to do	I would say yes. I think he can have a conversation and be human about it. People would probably be extremely appreciative of getting that information… It's the right thing to do. (06_Advocate)
Right to own information	Just as a general principle, if I contribute to the research, I should have access to what you learn as a result of it. (01_Advocate)
Health benefit	Offering the results helps me understand where my data is going and how it benefits the bigger, wider community. It helps me to possibly learn something that I didn't know and then it could directly improve my care. (08_Advocate) If image data yields a new suspected risk or diagnosis, or something that is highly pertinent and actionable for the patient, then I think there's an obligation to try to notify the patient about that. (30_Developer)

The Journal of Law, Medicine & Ethics, 48 S1 (2020): 9-36. © 2020 The Author(s)

How to offer: Explain uncertainty	If the algorithm is incorrect, then perhaps you've got false positives. But I think that can be handled via how you message to the users. You don't tell them they have cancer, you tell them that an algorithm identified that they have a potentially dangerous mole that they should speak to their doctor about. (05_Developer) Figuring out all the language and what is legally permissible to say, but telling people they need to make sure to also talk to their doctor about a particular mole or a trend or something, I think is very important, to be able to return those results from the research. (16_Researcher) 'Researchers reviewing your data have information … that may be about an elevated risk that you may choose to want to discuss with them. Click here to authorize direct contact,' or something like that. (30_Developer) I think the right approach would be to say, 'We've identified a mole that is of concern and this could be melanoma and there's no way of knowing from a photograph. You should seek help.' That would be the right way to do it. (35_Advocate)
How to offer: Through the company	I'd be concerned if I were Health Apps about [Mr. Lee] talking to their customers directly. Maybe they want it sent … back through them so that the company that has the relationship with the consumer is the one continuing to communicate with them. (27_Regulatory)
B. No	
Concerns about rigor, validation, expertise, skills	I'm glad [Mr. Lee] believes that, but it feels very weird. I don't know that machine learning is there yet, first of all. Given that he doesn't have the expertise, or a team that's trained to do that ... I'm not convinced of the validity of his data to go around informing people. Actually, that's a dangerous road. Those conversations should involve healthcare professionals. (14_Researcher) I would be wary, from a company standpoint, of going back to our patients and saying … we have a wonky computer scientist that figured out a way that shows that you might be more at risk than you think you are. I don't know, that just sounds to me like it's ... opening up a can of worms. (17_Researcher) I don't think it's clear from that description whether or not his research has been verified or reviewed in any way. So it's not clear that it's accurate. I think that's not ethical to give people information either that they don't want or that's potentially not accurate. (20_Regulatory) Not at this stage, no. I think that can be dangerous. I'm saying 'at this stage' because he may proceed with getting this vetted and published and verified, which is a different story. But if he just believes he sees something in his data and wants to convey that back to users, I don't think that should happen. (28_Developer)
Need for prior consent	One way to do that would be to say that like a question, 'Would you like to know your risk?' I think you have to allow that individual to make that decision for themselves, and not rely on Mr. Lee to give that information. (12_Researcher) The problem here is that if you ask them after the fact… I think the thing to do might be is if you decide you wanna do this, you probably need to just send out a blanket request for consent to all participants. So, it can't be like just targeted to only the person, or small number of people that you think you have something to report back on. (25_Researcher) It comes way down to that consent form. Like, "Hey, someone, anyone, who is interested in science might contact you with the opinion that you are at high risk." You know? If it was put like that, plainly, on that page one. (26_Regulatory)
Privacy concerns	Yes, if that happened to me, I would raise the red flag and say, 'I appreciate you telling me, but all this bullshit about you telling me that the data is not identifiable went out the window.' (12_Researcher)

8. Most appropriate approaches for Substance Abuse app

A. Notification / permission	
Most appropriate approach is the *same or substantially similar* *	The appropriate way to let them know that their data could be used for research? I assume the same way, the notification that we discussed before. I don't have different feelings about this particular app than [MoleStar], because it sounds pretty similar. (14_Researcher) I don't make a distinction in terms of privacy risk or anything like that, between the substance abuse and the melanoma. Some people might. I don't see that as different, necessarily. (30_Developer)
Most appropriate approach is *different* *	I don't think there's any effective way to do that, to be perfectly honest. We go back over the same issues we had before about 'what is research?' and who decides all of that. This particular app puts people at such incredible risk of various things that I don't think there's any way to inform them and get a rational decision from them… The whole process of notifying people through apps and through notices on cell phones doesn't work very well at all, but this one—the complexities are so great that it's just completely impractical to expect this process to work for that and to tell people the nature of the risks and to explain to them how this is different than the risks they would take if they go to a substance abuse treatment facility. It's just all too complicated. It would take hours to do. I think it would be hard to convey to a lot of people if you sat them down in a room and spent half an hour explaining it all to them but at least that's conceivable. On a cell phone, there's no way you could get away with that. There's no way you could do that effectively. (09_Regulatory)

The Journal of Law, Medicine & Ethics, 48 S1 (2020): 9-36. © 2020 The Author(s)
https://doi.org/10.1177/1073110520917027 Published online by Cambridge University Press

	As a substance abuser, I'm in a vulnerable position. I need something, this is a good program. I'm probably going to sign up for it irrelevant of what it is, and just hope that ... I'm not compromised. I think the potential for coercion may be even higher here because of the need, and obviously also the cognitive capacity, as well. (12_Researcher) I think a lot of substance use disorders have a cognitive component to it. This is a population that may have more difficulty actually understanding … so this probably needs in-person informed consent. (31_Researcher)
B. Data access procedures	
Most appropriate approach is the *same or substantially similar* *	I guess one thing to think through is anonymizing given the nature of the data, because now you've got location data—but I don't know that that presents any unique challenge that I can think of right now. (14_Researcher) The same kind of things are applicable, but maybe you add more rigor to who you decide to give the data to… Because people are engaged in illegal behavior, there are higher risks for this type of data. So maybe you have to be rigorous about the standards that you use to determine who is going to get to do research on these data and what that data use agreement says. (27_Regulatory)
Most appropriate approach is *different* *	If [an app is] going to use the phone's GPS or accelerometer, those are third party technologies and those third parties also have access to that information beyond what the app has. People are giving their data to third parties that are not specifically covered under this agreement. It's potentially available to Google and to my carrier… Then it's available to Facebook and then every app that is linked to Facebook now has access to it. That's why there are so many more entities that have access now. There has to be some information [in the DUA] about the technology providers and their access to information about the individual. (39_Researcher)

* Relative to what interviewee said would be the most appropriate approach for MoleStar

9. Should unregulated researchers seek to make their data widely available?

A. Yes	
Scientific advancement	For the same reasons that you do with the NIH: because patients don't have time to waste while researchers do 30 different things in parallel when they could figure out how to collaborate and build on top of what they're learning. (01_Advocate) The piece we differentiate as federal funding with these requirements is a false reality. Yes, there's the very practical argument, these are tax payer dollars so they should be public—but I think any research that contributes to the public good, and to advancing knowledge, and scientific responsibility, why not share it? (07_Regulatory) In many instances, somebody may be conducting something similar, and if they have access to [others'] research, they may not have to reinvent the wheel and they can build on it instead of doing the same thing over and over again. It just makes the system that much more efficient. (11_Advocate) Absolutely. All research needs to be made available, whether good or bad. One of the biggest problems, one of the conundrums of this research world is that a lot of the negative research is not published by industry, by payers, by providers, by academic traditional researchers because it doesn't make them look good. We're not going to learn just from the good research. We are also going to learn from the bad research. (19_Advocate) Ideally, people would share their data because there's an opportunity to learn more from it and we're always coming up with new ways that we could potentially learn something from data. (24_Developer)
Ethical considerations	Clearly, if you're getting other additional sources of funding as well as data from individuals, there's an ethical responsibility to share what you've [learned] ... and at a large high level, at least give it back. (12_Researcher) Ideally the information could be shared. The challenge is doing it in a way that's safe for the participants and doesn't surprise them. (24_Developer) I think ethically, it's wrong for the computer scientist to not share, but I don't know if legally there's anything wrong with it. (34_Researcher)
Business, marketing, competitive considerations	Absolutely. And for no other reason than it's a marketing opportunity. There's a lot of smoke and mirrors in the mobile health forum now, and we're now being consistently asked to provide data or justification. When we're asked to do that, you need some type of peer reviewed honest broker. You don't have to put out your secret sauce, you don't have to put the algorithms out there, but I think that if we keep the cards too close to ourselves, you're limiting your opportunity to disseminate what we're trying to do. (12_Researcher) There could be a lot more gained by sharing what we know rather than keeping it hidden. However, politically speaking, that's just not going to happen because of the capitalistic undertones, the for-profit undertones, of medical research, particularly FDA regulated or just anonymous but industry sponsored and otherwise unregulated research. I don't like to consider myself not progressive in this area, but I am certainly not politically unprogressive. I think in terms of likelihood of anything like that actually happening, it's doubtful. (22_Regulatory)

The Journal of Law, Medicine & Ethics, 48 S1 (2020): 9-36. © 2020 The Author(s)

	Maybe they've learned something really valuable that's helped them capitalize on the billions of dollars they spent doing research to make a product that will be valuable… So at the time, maybe that's not the best thing to have them share that data immediately. But maybe eventually. On the other hand, adverse events should be reported mandatorily to a regulated agency so that people don't get harmed. (24_Developer) I think that when private people, like private users or when individuals purchase the data, and they find something that goes against their hypothesis, there's probably a greater chance that they would not release those findings or if something's damaging to the commercial enterprise, they're less likely to release that information. (41_Researcher)
Where to disseminate	You don't have to have a PhD to publish in journals, but that certainly would subject it to the scrutiny that science typically gets subjected to, which would be important, if there are people who are doing this, and doing it well, more power to 'em. Otherwise, I don't think it'll disseminate very well, or be taken seriously, in which case, it could be a lack of ... It's a waste of people's time. (14_Researcher)
B. No	
Validity	I'm having these visions of Nazis, like sewing gorilla heads onto human bodies, you know? Is it cool if they do that, as long as they share their results? No. They really shouldn't tell people they were doing that. Also, don't do it, you know? (26_Regulatory)
C. Depends	
	I think in general it's a good practice, but I don't think it should be mandatory for private entities, because it really depends on where that funding comes from and what the business model is--they have a right to figure out how best to get a return on investment. (25_Researcher) Earlier on, I said that we've had experience that it's very hard to actually make any of this work with small start-up companies, that are the ones that are generating the data. If we then pile on top of that a whole lot of expectations, obligations, and legal requirements, regulations, it's not going to get better. It's going to get much worse, and so that's why I'm trying to be thoughtful about what gets extended to people. (30_Developer) There's a risk you take in data sharing. I want to see data shared. I want to see lots of data shared, but I don't want to say that I think there's an obligation to do so free of support. I think it needs to be within an environment where the company is not paying a price or taking on a lot of risk to do something that other companies won't do, and then it puts itself at a disadvantage. Good but not at expense of company. (38_Developer)

The Journal of Law, Medicine & Ethics, 48 S1 (2020): 9-36. © 2020 The Author(s)