Original Paper

Understanding the Security and Privacy Concerns About the Use of Identifiable Health Data in the Context of the COVID-19 Pandemic: Survey Study of Public Attitudes Toward COVID-19 and Data-Sharing

Charlotte Summers[1], BSc; Frances Griffiths[2], PhD; Jonathan Cave[2,3,4], PhD; Arjun Panesar[1], MEng

[1]DDM Health, Coventry, United Kingdom

[2]Warwick Medical School, University of Warwick, Coventry, United Kingdom

[3]Department of Economics, University of Warwick, Coventry, United Kingdom

[4]Data Ethics Group, The Alan Turing Institute, London, United Kingdom

Corresponding Author:
Charlotte Summers, BSc
DDM Health
Technology House, Science Park
University of Warwick
Coventry, CV4 7EZ
United Kingdom
Phone: 44 7969091134
Email: charlotte@ddm.health

Abstract

Background: The COVID-19 pandemic increased the availability and use of population and individual health data to optimize tracking and analysis of the spread of the virus. Many health care services have had to rapidly digitalize in order to maintain the continuity of care provision. Data collection and dissemination have provided critical support for defending against the spread of the virus since the beginning of the pandemic; however, little is known about public perceptions of and attitudes toward the use, privacy, and security of data.

Objective: The goal of this study is to better understand people's willingness to share data in the context of the COVID-19 pandemic.

Methods: A web-based survey was conducted on individuals' use of and attitudes toward health data for individuals aged 18 years and older, and in particular, with a reported diagnosis of a chronic health condition placing them at the highest risk of severe COVID-19.

Results: In total, 4764 individuals responded to this web-based survey, of whom 4674 (98.1%) reported a medical diagnosis of at least 1 health condition (3 per person on average), with type 2 diabetes (n=2974, 62.7%), hypertension (n=2147, 45.2%), and type 1 diabetes (n=1299, 27.4%) being most prominent in our sample. In general, more people are comfortable with sharing anonymized data than personally identifiable data. People reported feeling comfortable sharing data that were able to benefit others; 66% (3121 respondents) would share personal identifiable data if its primary purpose was deemed beneficial for the health of others. Almost two-thirds (n=3026; 63.9%) would consent to sharing personal, sensitive health data with government or health authority organizations. Conversely, over a quarter of respondents (n=1297, 27.8%) stated that they did not trust any organization to protect their data, and 54% (n=2528) of them reported concerns about the implications of sharing personal information. Almost two-thirds (n=3054, 65%) of respondents were concerned about the provisions of appropriate legislation that seeks to prevent data misuse and hold organizations accountable in the case of data misuse.

Conclusions: Although our survey focused mainly on the views of those living with chronic health conditions, the results indicate that data sensitivity is highly contextual. More people are more comfortable with sharing anonymized data rather than personally identifiable data. Willingness to share data also depended on the receiving body, highlighting trust as a key theme, in particular who may have access to shared personal health data and how they may be used in the future. The nascency of legal guidance in this area suggests a need for humanitarian guidelines for data responsibility during disaster relief operations such as pandemics and for involving the public in their development.

XSL•FO

RenderX

Introduction

The World Health Organization declared the COVID-19 outbreak a public health emergency on January 30, 2020; after 6 weeks, it was categorized as a pandemic [1]. Certain groups of people are particularly likely to have serious or severe symptoms of COVID-19 [2]. Preliminary data suggest that people with obesity are at an increased risk of severe COVID-19 [3]. Type 2 diabetes mellitus and hypertension are the most common comorbidities in patients with COVID-19 [4]. According to several reports, including those from the Centers for Disease Control and Prevention, patients with type 2 diabetes are at a greater risk of death than those without type 2 diabetes [5].

Digital health technologies are being used in the fight against COVID-19 [6]. Global health care systems have seen an influx in the incidence of the same novel condition, and the contagious nature of the condition has driven the shift to remote medicine. Many health care pathways have been rapidly digitalized with face-to-face services seeing a drop in usage [7]. This has increased the collection, sharing, and use of data in digital form. Technology is used for remote monitoring, general practitioner consultations, providing structured education, and tracking the spread of disease. As well as the technologies themselves, the data they generate are also useful [8].

Timely, secure, and reliable data access and sharing are critical to understanding COVID-19, controlling its spread, improving the effectiveness and acceptance of government policies, and fostering global cooperation in the race to develop and distribute effective therapies and vaccines. During the COVID-19 pandemic, data are being rapidly shared to understand the location of infections, confirmed cases, recoveries, and deaths. The main data points of interest for this are geolocation and biometric data, both of which are available from users' mobile devices. However, there are serious concerns regarding the objectivity and accuracy of these data, and their utility has been compromised by inconsistent collection and definitions. This, in turn, feeds back into individuals' trust in the collecting organizations and in the extent to which their shared data will actually be used to help others, and this needs to be matched by the trustworthiness of those organizations.

During an unprecedented time, some digital responses to the crisis have precipitated novel data governance and privacy challenges [9]. Governments are taking extraordinary measures to track, trace, and contain the spread of COVID-19 by transitioning to digital technologies and advanced analytics to collect, process, and share data for effective frontline responses. Government-mandated apps are bringing the fight against COVID-19 onto users' devices and have generally adopted pragmatic and contextualized approaches, but they have prompted concerns about security and privacy and the control

and use of data beyond the pandemic [10]. There is a trade-off between effectiveness and privacy, centralized and decentralized implementations, and the links to trace and isolate policies.

While the exceptional measures implemented in some countries may prove effective in limiting the spread of the virus, some have provoked controversy in terms of privacy and other fundamental rights, particularly when they lack transparency and public consultation [11]. In South Korea, the specificity of publicly available anonymized data raised privacy concerns when some researchers found that data trails were so detailed that individuals could be identified [12]. In Italy, the Department of Prevention released specific guidelines on the application of the European Union's and national data privacy rules in the context of the COVID-19 pandemic [13]. Similarly, the United Kingdom's Information Commissioner's Office, an independent authority set up to uphold information rights in public interest, confirmed that there would be no regulatory action taken against organizations that fail to meet the data protection standards if noncompliance results from the COVID-19 pandemic [14,15]. In China, new arrivals to the country are tested for COVID-19, instructed to download a government-mandated app, and wear a wristband that is linked to the app to monitor movement with a technology similar to that used in Singapore [16,17]. The United Kingdom's Track and Trace app was the center of a debate on centralization of data [18]. On May 5, 2020, the Government revealed its first attempt at a contact-tracing app, but 6 weeks later admitted that the app was flawed and it would switch to a more privacy-preserving model devised by Apple and Google [19,20]. Transparency is a key theme. One of the most common misconceptions about the United Kingdom's Track and Trace app was that it could allow users to specifically identify and map COVID-19 cases among their contacts and in their vicinity [21].

This study seeks to understand the opinions of British people with long-term health conditions on the themes of data privacy and security, data ethics, and data misuse and to assess the possible trade-offs in data utilization to manage a crisis such as the COVID-19 pandemic [22]. It is important to understand the concerns of people with long-term health conditions such as type 2 diabetes and hypertension as these conditions have been shown to be key risk factors in the progression and prognosis of COVID‐19 [23,24].

Methods

Study Design and Setting

A web-based survey study was conducted with a mixed methods design conforming to the checklist for reporting results of internet electronic surveys [25]. An email invitation to participate, which included a weblink to the survey, was sent to 11,213 people who had consented to be contacted for research opportunities.

Quantitative information (closed and multiple-choice questions) was collected on four topics: (1) demographic characteristics, (2) COVID-19 symptoms and clinical diagnoses, (3) sharing and privacy of pre– and post–COVID-19 health data, and (4) COVID-19 lockdown behaviors. Responses from the final topic are not included in this analysis.

The survey contained 31 questions: 26 closed questions, 1 open question, and 4 demographic questions. Questions on sharing and privacy of pre– and post–COVID-19 health data were answered on a 5-point Likert scale with responses ranging from *strongly disagree* to *strongly agree* or from *not concerned at all* to *very concerned*.

Participants

People aged ≥18 years who had joined the Diabetes.co.uk community were surveyed. The survey commenced with 1 screening question: "Do you consent to take part in the study?" Respondents who consented went on to complete the survey.

Procedure

Data collection occurred between July 6 and August 31, 2020. The survey was administered through the Jisc Online Surveys software and comprised closed, open, and multiple-choice questions. The survey was designed to elicit individual responses to questions about retrospective data use and privacy prior to the COVID-19 pandemic and prospective use during the COVID-19 pandemic.

It is intended to have multiple windows of data collection for several reasons: people's recollections of pre–COVID-19 attitudes may be unreliable, and changes in the course, apparent seriousness, and confidence in scientific understanding of the pandemic will have evolved.

The type and wording of each question was composed by the research team. The order of questions was not randomized. The survey followed a predetermined logic where contingent questions were included or automatically skipped on the basis of responses. Qualitative data were collected with 1 open question exploring what respondents would like to see happening: "What would you like to see happen to improve the COVID-19 situation?" (question 30).

Analysis

We exported all data from Jisc and conducted data analysis using SPSS (version 22; IBM Corp). We conducted descriptive data analyses of sample distributions and characteristics. Pearson *r* correlation coefficients were used to determine the relation between prior data-sharing behavior and attitudes toward data-sharing activity in the context of the COVID-19 pandemic. The data from the open question were read through and then categorized into themes.

Ethical Considerations

Ethics approval was obtained from the Human Research Ethics Committee of the University of Warwick (BSREC 144/19-20). Web-based informed consent was required before the survey could be accessed.

Results

Survey Respondents

Of 11,213 people emailed, 10,705 clicked through to the survey; in total, 4764 gave their consent and began the survey. As indicated in Table 1, all of them completed the survey and were included in the analysis. All respondents were located in the United Kingdom. In total, 2287 (48.0%) respondents were male and 3083 (64.8%) were aged between 55 and 74 years. A total of 115 (2.8%) respondents reported having been clinically diagnosed with COVID-19. The majority of patients (n=4674, 98.1%) reported a prior clinical diagnosis of at least one health condition (on average 3 per person). There was a high prevalence of individuals living with type 2 diabetes (n=2974, 62.7%), hypertension (n=2147, 45.2%), type 1 diabetes (n=1299, 27.4), obesity (n=892, 18.8%), and depression (n=871, 18.3%). Respondent demographics are shown in Table 1.

Table 1. Respondent demographics (N=4764).

Characteristics	Respondents, n (%)
Gender	
Male	2287 (48.0)
Female	2435 (51.1)
Prefer not to say	42 (0.9)
Age (years)	
18 to 24	23 (0.5)
25 to 34	104 (2.2)
35 to 44	298 (6.3)
45 to 54	839 (17.6)
55 to 64	1550 (32.6)
65 to 74	1533 (32.2)
75 or older	410 (8.6)
Prefer not to say	7 (0.1)
Health conditions[a]	
Type 2 diabetes	2974 (62.7)
Hypertension	2147 (45.2)
Type 1 diabetes	1299 (27.4)
Arthritis	1002 (21.1)
Obesity	892 (18.8)
Depression	871 (18.3)
Employment[b]	
Full-time employment	1205 (24.9)
Part-time employment	571 (11.8)
Retired	2298 (47.5)
Student	39 (0.8)
Unemployment	464 (9.6)
Furloughed	188 (3.9)
Volunteering in my community (National Health Service, key services)	69 (1.4)
Ethnicity	
Indian or Pakistani	69 (1.4)
Black, British African, or Caribbean	46 (1.0)
Middle Eastern	5 (0.1)
Mixed groups	27 (0.6)
White	4,434 (93.1)
Other	44 (0.9)
Chinese, Japanese, or East Asian	7 (0.1)
Prefer not to say	132 (2.8)

[a]Frequently occurring health conditions selected.

[b]Respondents selected multiple statuses; for example, full-time employed and furloughed.

XSL•FO

RenderX

COVID-19 Symptoms and Clinical Diagnosis

Of the 4764 respondents who completed the survey, 494 stated they had had symptoms of COVID-19. The most common symptoms were the following: a continuous cough (n=467, 94%), fever (n=325, 65.7%), difficulty breathing (n=384, 77.7%), and loss of taste (n=324, 65.5%). In total, 384 (77.7%) respondents reported another symptom, predominantly fatigue (16.4%).

Of those reporting symptoms, 111 (22.5%) reported a clinical diagnosis of COVID-19. Of these respondents, 73 (63.5%) reported that their symptoms were severe or very severe. In total, 26 (22.6%) respondents reported that their symptoms were not severe at all. All respondents who reported a clinical diagnosis of COVID-19 reported at least one symptom, including loss of smell or taste (63.1%), fever (62.2%), difficulty breathing (61.3%), or continuous cough (53.2%).

A total of 131 (2.8%) respondents reported that a household member had been tested and was clinically diagnosed with COVID-19.

Sharing and Privacy of Pre–COVID-19 Health Data

Prior to the COVID-19 pandemic, almost half of the respondents (n=2313, 49.2%) agreed or strongly agreed that they often consented to anonymized sharing of their private health data, while only 608 (13%) respondents often consented to sharing of private health data without anonymization. Two-thirds of respondents (n=3113, 66.7%) disagreed or strongly disagreed with sharing their private health data without anonymization. Similarly, 3121 (66.3%) respondents would share their data if it keeps other people healthy; 3026 (63.9%) respondents agreed or strongly agreed to sharing private health data with the government or health authority; 1911 (40.7%) respondents agreed or strongly agreed to share their private health data with services that provide health services to the National Health Service (NHS) such as the Low Carb Program and PushDoctor. Only 232 (5%) participants agreed or strongly agreed to share private health data with social media platforms. Over a quarter of respondents (n=1297, 27.8%) agreed or strongly agreed that they did not trust *any* organization to protect their private health data. Just under a quarter of respondents (n=1094, 23.5%) agreed or strongly agreed that they were not concerned by the implications of sharing private health data. General health data–sharing responses are shown in Table 2. Respondents who reported that they felt "neutral" in response to the statements were excluded.

Table 2. General health data sharing responses.

Question	Disagree or strongly disagree, n (%)	Agree or strongly agree, n (%)
I often consent to share my private health data with any organisation as long as it is anonymised	1273 (27.1)	2313 (49.2)
I often consent to share my private health data to any organisation without anonymisation	3113 (66.7)	599 (13)
I am not concerned about the implications of sharing my private health data	2528 (54.2)	1094 (23.5)
I don't trust any organisation to protect my private health data	1541 (33.8)	1297 (27.8)
I'm happy to share my private health data if it helps keep other people healthy	526 (11.2)	3121 (66.3)

Sharing and Privacy of Post–COVID-19 Health Data

Over half (n=3026, 63.9%) agreed or strongly agreed to share their private data with the government or health authority if asked; 1911 (40.7%) respondents would happily consent to share their private data with services that provide health services to the NHS such as the Low Carb Program and PushDoctor, if asked. Only 232 (5%) participants agreed or strongly agreed that they would consent to sharing private data with social media if asked.

Almost half of respondents (n=2228, 47.1%) were concerned or very concerned about who would have access to their personal health data in the context of the COVID-19 pandemic and 2310 (49.1%) respondents were concerned or very concerned about how their personal health data may be used in the future. Almost two-thirds of respondents (n=3054, 65%) were concerned or very concerned around the legislation of data misuse.

Just over a third of respondents (n=1563, 33.4%) would consent to share their private data with any organization if it was providing essential COVID-19 support services such as the supermarkets, pharmacies, and banks. Responses toward the use of post–COVID-19 patient data is shown in Table 3, along with the sentiment toward the use of patient data in the context of the COVID-19 pandemic, and Table 4 shows the sentiment toward future use or misuse of data collected and used under the provisions of the COVID-19 pandemic.

Table 3. Responses toward the use of post–COVID-19 patient data.

Question	Disagree or strongly disagree	Agree or strongly agree
I would happily consent to share my private health data with the government or health authority	728 (15.4)	3026 (63.9)
I would happily consent to share my private health data with social media e.g Twitter, Facebook, Google	4023 (85.9)	232 (5)
I would happily consent to share my private health data with services that provide health services to the NHS[a] such as Low Carb Program, PushDoctor, Babylon Health	1351 (18.8)	1911 (40.7)

[a]NHS: National Health Service.

Table 4. Sentiment toward future use or misuse of data collected and used under the provisions of the COVID-19 pandemic.

Question	Not concerned at all	Concerned or very concerned
In light of COVID-19, how concerned are you about who would have access to your personal health data?	1138 (14)	2228 (47.1)
How concerned are you about how your personal health data may be used in the future?	1162 (24.7)	2310 (49.1)
How concerned are you around the legislation of data misuse?	644 (13.6)	3054 (65)

Prior Willingness to Share Data

Correlations between retrospective data-sharing that happened in the context of generalized concerns and attitude changes associated with the course of the pandemic were determined. Changes were not linked to any specific studies, policies, or measures. There were strong correlations in the attitudes of people exhibiting high levels of concern about future uses of shared data and concerns about access (r_{4685}=0.816; P<.001). There was a strong correlation between people exhibiting attitudes of concern that firmer legislation for data misuse is needed and concerns about future repurposing and reuse of personal health data collected during the COVID-19 pandemic (r_{4663}=0.636; P<.001). The Pearson r correlation coefficient as a normalized measure of the strength of a possible linear correlation, lying between –1 and +1. The Pearson r correlation coefficient measures nonlinear correlations (eg, when extreme views are highly correlated but more moderate ones are more independent).

Respondents agreed to share their personal data with roughly the same parties prior to the COVID-19 pandemic and within the context of the COVID-19 pandemic; governments and health authorities (r_{4710}=0.762; P<.001), health service providers such as the Low Carb Program and PushDoctor (r_{4662}=0.783; P<.001), and social media platforms such as Twitter, Facebook, and Google (r_{4662}=0.736; P<.001).

COVID-19 News and Information

Of the 4764 respondents, 2666 (56.1%) were concerned that they may be receiving misinformation about COVID-19 from trusted sources, 1079 (22.7%) were not concerned (genuinely unconcerned and those who feel that they are in control of the consumption of news and information), and 1006 (21.2%) had never considered it.

In total, 4237 responded to the open-ended question of what they would like to see happening to improve the COVID-19 situation. The majority of respondents shared a single response:

1348 (31.8%) stated they would like to see a reliable vaccine and treatment, 884 (20.8%) stated they would like to see balanced information from the government, and 485 (11.4%) wanted to see stricter measures to prevent the transmission of COVID-19.

Discussion

Principal Findings

Our study provides insights into public perception and attitudes toward the use of identifiable health data in the context of the COVID-19 pandemic; in particular, the perspectives of those living with chronic, long-term health conditions, with an average of 4 health conditions reported per respondent.

Our study suggests that data sensitivity is highly contextual. A significant proportion of people felt that their own attitudes have shifted as a result of the COVID-19 pandemic. More people reported being comfortable with sharing private health data with any organization during rather than before the COVID-19 pandemic. In order, people appear to trust their data with the government, health organizations, and social media. There is significant distrust of private health data use by social media organizations (eg, Twitter, Facebook, and Google) even though social media is used as a channel for communication by people caught up in crises such as emergency relief operations after earthquakes, tsunamis, and typhoons; where it provides a trusted and highly salient source of information about what is happening and what to do [26,27]. This is surprising as although users worldwide report that privacy and use of personal data are important issues, most rarely make an effort actively to protect these data and often even give them away voluntarily on social media where even innocuous data can reveal sensitive health information when suitably processed [28,29]. People treat data revelation and sharing differently depending on the perceived sensitivity of the data, and the sensitivity attached to different types of data is neither stable nor uniform.

XSL•FO

RenderX

When examining the correlations between retrospective views of data-sharing behavior and comfort regarding data-sharing in the context of the COVID-19 pandemic, individuals were comfortable (or not) in sharing personal data with the same organizations prior to and during the COVID-19 pandemic, suggesting that COVID-19 has not drastically shifted people's willingness to share or withhold their personal data. This may be because attitudes have shifted both retrospectively and prospectively and also depends on whether people accurately remember and report their past views and actions. One of the strongest correlations observed in the analysis was between high levels of concern about the requirement for stronger legislation protecting individuals from data misuse and future repurposing and reuse. This highlights the need for improved communication, transparency, and potentially stronger regulation on how such data may be repurposed in the future, who will be accountable for inappropriate use of data, and a commitment to cease or reverse exceptional uses of data when the crisis is over. Individuals' data rights are protected by law in regulation such as General Data Protection Regulation 2018 in Europe and Health Insurance Portability and Accountability Act in the United States, which make clear the scope, purpose, and time limitations of data usage [30,31]. Concerns may therefore reflect ignorance of existing rules, doubts over enforcement, or a belief that current legislation does not go far enough (for instance, in the requirement of erasure after 3 years rather than a shorter time duration).

A key theme emerging from the literature that was confirmed in this study is the importance of trust [32-34]. Over a quarter of respondents stated they did not trust any organization to protect their data, over half reported concern about the implications of sharing personal information, and almost two-thirds were concerned about data misuse regulation not being strict enough. When asked during the pandemic (the United Kingdom's first wave), almost half of respondents were concerned about who would have access to their personal health data and a similar number were concerned about how their personal health data might be used in the future. This is consistent with prior research suggesting that public involvement in data policy is crucial to bolstering trust and provides support for legislation that is more enforceable [35]. Attitudes may have been perturbed by news stories relating to cybersecurity and privacy and by policy announcements (eg, around Huawei, the Online Harms Bill, etc) [36,37].

Although there are no directly comparable studies, the results from this study complement prior research on public perceptions about COVID‑19 and data-sharing. Data privacy and protection are important concepts [38]. Data policy tends to address human concerns about privacy by making rules about data protection; however, this can lead to category errors since data protection can undermine privacy.

Willingness to share anonymized personal health information varies depending on the degree to which the receiving body is trusted and the uses to which the data will be put [39,40]. The more commercial the objectives of the receiving institution appear, the less respondents are willing to share their personal health information. This in turn suggests that anonymization's disadvantages (in terms of confirming data and correlating

shared with other data) might be offset by better (wider, deeper, and more accurate) sampling leading to greater validity of results. Further evidence comes from the interaction (or correlation) between these attitudinal responses and other characteristics, meaning that nonanonymized collection might lead to biased results.

Virus tracking apps are used at scale by governments; however, concerns about transparency, privacy, and morality remain [41,42]. There has been substantial research into the challenges involved in the digital response to the COVID-19 pandemic and proposed methodologies for the ethical design and use of digital public health tools [43,44]. Clear and effective data ethics is both a moral and a practical obligation. The nascency of legal guidance in this area combining ethics, law, and humanitarian impulses suggests the requirement for humanitarian guidelines for data responsibility during global crises such as pandemics. Therefore, rather than recalibrating the expectations of people with regard to their own privacy, the requirements for the use of data should be broader and more comprehensive as ethically collected big data could prove to be extremely useful in the prediction, monitoring, and mitigation of pandemics such as COVID-19 [45].

Strengths and Limitations

Despite the importance of the findings reported here, it is important to note that this study had several limitations. Conducting this study via a web-based survey carries a risk of response bias, simply because the respondents are likely to be more technology-savvy than the general population. However, the population studied (those with chronic health conditions) is of interest as these participants have a degree of awareness and the ability to self-manage their condition that is not (yet) typical of the population at large, and this sheds light on how policies that raise awareness may lead to greater effectiveness in terms of uptake of technical solutions and effectiveness of public health advice and other policies.

Participants were asked to rate retrospectively their perceptions of data sharing prior to the COVID-19 pandemic; these ratings may be inaccurate owing to faulty memories and response bias. There is some ambiguity between what people (now) thought they would have done had they been asked and how they responded to actual requests for consent to data-sharing. In particular, one could disagree with a statement like "I often consent to share my private health data to any organisation without anonymisation" simply because one rarely recalls being asked to share (even without a principled objection to such sharing should the occasion arise). In addition, the phraseology of the questionnaire refers to data that have either been anonymized or are identifiable. This dichotomous representation leaves out pseudonymized data. The participants were not educated about the concepts of anonymization prior to answering the questions potentially allowing ambiguity of the terminology to cause a strong bias in the response behavior.

In addition, the sample is concentrated on people with diabetes and those with other diagnosed health conditions, rather than the general population. This is a strength as well as a weakness as it focuses on a population with particular circumstances and perspectives and one that may be more representative of a

post–COVID-19 population that has been sensitized to a continuing health concern than the current population. This in turn means that a comparison of these findings with a similar survey of the general population can shed light on the potential impact of awareness-raising policies.

Another strength of the study is the high number of respondents who completed the questionnaire. In total, 4764 people participated in the study. This provides a unique insight into the views of a population deemed as being at the highest risk of severe disease and mortality related to COVID-19 [46]. There was a skew in the representation of the demographic distribution of individuals in the nationwide population of people living with chronic health conditions, since White people were overrepresented in our sample (n=4434, 93.4%) but not overrepresented among those with diabetes more generally. While our survey focused on those with diabetes, the results provide novel insight into concepts crucial for societal trust in data use and sharing initiatives.

While the study design did not allow us to ascertain whether technology use itself was correlated with higher acceptance of data-sharing, such an analysis is possible and will be an important topic for future research.

The study's findings suggest potential targets for further study and possible considerations for policy makers. There are two main implications: storing and processing data in pseudonymized form and emphasizing the use of synthetic data (generated from models estimated from real data but not involving any actual or identifiable human beings).

Understanding attitudes toward data sensitivities and trust can contribute to developing policies, improving transparency, and increasing the trust, speed, focus, and effectiveness of epidemic responses. Future practice should emphasize transparent data-sharing and privacy initiatives, while research should evaluate whether this does indeed lead to greater levels of trust and engagement. Encouraging ethical and relevant data-sharing can provide significant epidemic intelligence and support public health emergency relief operations [47].

Conclusions

Data sensitivity is highly contextual. More people are comfortable with sharing anonymized data than personally identifiable data. Willingness to share data also varied depending on the receiving body, highlighting trust as a key theme, who may have access to shared personal health data and how it may be used in the future. The nascency of legal guidance in this area suggests the requirement for humanitarian guidelines for data responsibility during disaster relief operations such as pandemics, and the requirement to involve the public in their development.

Conflicts of Interest

CS and AP are founding employees of DDM, who operate the Diabetes.co.uk Forum community.

References

1. Chan M, Yeo SE, Chong Y, Lee Y. Stepping Forward: Urologists' Efforts During the COVID-19 Outbreak in Singapore. Eur Urol 2020 Jul;78(1):e38-e39 [FREE Full text] [doi: 10.1016/j.eururo.2020.03.004] [Medline: 32192816]

2. Bornstein SR, Dalan R, Hopkins D, Mingrone G, Boehm BO. Endocrine and metabolic link to coronavirus infection. Nat Rev Endocrinol 2020 Jun 02;16(6):297-298 [FREE Full text] [doi: 10.1038/s41574-020-0353-9] [Medline: 32242089]

3. Stefan N, Birkenfeld AL, Schulze MB, Ludwig DS. Obesity and impaired metabolic health in patients with COVID-19. Nat Rev Endocrinol 2020 Jul;16(7):341-342 [FREE Full text] [doi: 10.1038/s41574-020-0364-6] [Medline: 32327737]

4. Sun Y, Guan X, Jia L, Xing N, Cheng L, Liu B, et al. Independent and combined effects of hypertension and diabetes on clinical outcomes in patients with COVID-19: A retrospective cohort study of Huoshen Mountain Hospital and Guanggu Fangcang Shelter Hospital. J Clin Hypertens (Greenwich) 2021 Feb 25;23(2):218-231 [FREE Full text] [doi: 10.1111/jch.14146] [Medline: 33369066]

5. Cole SA, Laviada-Molina HA, Serres-Perales JM, Rodriguez-Ayala E, Bastarrachea RA. The COVID-19 Pandemic during the Time of the Diabetes Pandemic: Likely Fraternal Twins? Pathogens 2020 May 19;9(5):389 [FREE Full text] [doi: 10.3390/pathogens9050389] [Medline: 32438687]

6. Scott BK, Miller GT, Fonda SJ, Yeaw RE, Gaudaen JC, Pavliscsak HH, et al. Advanced Digital Health Technologies for COVID-19 and Future Emergencies. Telemed J E Health 2020 Oct 01;26(10):1226-1233. [doi: 10.1089/tmj.2020.0140] [Medline: 32456560]

7. Virtual Solutions for Managing Cancer Care In a Pandemic Era: Lessons from COVID-19. A Rapid Evidence Review. NHS Peninsula Cancer Alliance. 2020. URL: https://peninsulacanceralliance.nhs.uk/wp-content/uploads/2020/10/Virtual-solutions-for-managing-cancer-care_FINAL-01Oct2020.pdf [accessed 2022-06-01]

8. Panesar A, Panesar H. Artificial intelligence and machine learning in global healthcare. In: Handbook of Global Health. Berlin: Springer; 2020:1-39.

9. Ensuring data privacy as we battle COVID-19. OECD. 2020 Apr 14. URL: http://www.oecd.org/coronavirus/policy-responses/ensuring-data-privacy-as-we-battle-covid-19-36c2f31e/ [accessed 2022-06-01]

10. Hoffman AS, Jacobs B, van Gastel B, Schraffenberger H, Sharon T, Pas B. Towards a seamful ethics of Covid-19 contact tracing apps? Ethics Inf Technol 2020 Sep 28;23(S1):1-11 [FREE Full text] [doi: 10.1007/s10676-020-09559-7] [Medline: 33013191]

11. Henderson J. Patient privacy in the COVID-19 era: Data access, transparency, rights, regulation and the case for retaining the status quo. Health Inf Manag 2021;50(1-2):6-8. [doi: 10.1177/1833358320966689] [Medline: 33176491]

12. Ahn NY, Park JE, Lee DH, Hong PC. Balancing Personal Privacy and Public Safety During COVID-19: The Case of South Korea. IEEE Access 2020;8:171325-171333. [doi: 10.1109/access.2020.3025971]

13. Torri E, Sbrogiò LG, Rosa ED, Cinquetti S, Francia F, Ferro A. Italian Public Health Response to the COVID-19 Pandemic: Case Report from the Field, Insights and Challenges for the Department of Prevention. Int J Environ Res Public Health 2020 May 22;17(10):3666 [FREE Full text] [doi: 10.3390/ijerph17103666] [Medline: 32456072]

14. Data protection and Coronavirus-19 – relaxation of government measures. Information Commissioner's Office. URL: https://ico.org.uk/global/data-protection-and-coronavirus-information-hub/data-protection-and-coronavirus/ [accessed 2021-03-27]

15. The ICO's regulatory approach. Information Commissioner's Office. URL: https://ico.org.uk/media/about-the-ico/policies-and-procedures/2617613/ico-regulatory-approach-during-coronavirus.pdf [accessed 2021-03-27]

16. Abuhammad S, Khabour OF, Alzoubi KH. COVID-19 Contact-Tracing Technology: Acceptability and Ethical Issues of Use. Patient Prefer Adherence 2020;14:1639-1647 [FREE Full text] [doi: 10.2147/PPA.S276183] [Medline: 32982188]

17. Simko L, Calo R, Roesner F, Kohno T. COVID-19 Contact Tracing and Privacy: Studying Opinion and Preferences. arXiv. Preprint posted online on May 12, 2020. [FREE Full text]

18. Sweeney Y. Tracking the debate on COVID-19 surveillance tools. Nat Mach Intell 2020 Jun 16;2(6):301-304. [doi: 10.1038/s42256-020-0194-1]

19. Alanzi T. A Review of Mobile Applications Available in the App and Google Play Stores Used During the COVID-19 Outbreak. JMDH 2021 Jan;14:45-57. [doi: 10.2147/jmdh.s285014]

20. Wise J. Covid-19: UK drops its own contact tracing app to switch to Apple and Google model. BMJ 2020 Jun 19;369:m2472. [doi: 10.1136/bmj.m2472] [Medline: 32561511]

21. Williams SN, Armitage CJ, Tampe T, Dienes K. Public attitudes towards COVID-19 contact tracing apps: A UK-based focus group study. Health Expect 2021 Apr;24(2):377-385 [FREE Full text] [doi: 10.1111/hex.13179] [Medline: 33434404]

22. Cairney P. The UK government's COVID-19 policy: assessing evidence-informed policy analysis in real time. Br Polit 2020 Nov 01;16(1):90-116. [doi: 10.1057/s41293-020-00150-8]

23. Guo W, Li M, Dong Y, Zhou H, Zhang Z, Tian C, et al. Diabetes is a risk factor for the progression and prognosis of COVID-19. Diabetes Metab Res Rev 2020 Mar 31:e3319 [FREE Full text] [doi: 10.1002/dmrr.3319] [Medline: 32233013]

24. Lippi G, Wong J, Henry BM. Hypertension in patients with coronavirus disease 2019 (COVID-19): a pooled analysis. Pol Arch Intern Med 2020 Apr 30;130(4):304-309 [FREE Full text] [doi: 10.20452/pamw.15272] [Medline: 32231171]

25. Eysenbach G. Improving the quality of Web surveys: the Checklist for Reporting Results of Internet E-Surveys (CHERRIES). J Med Internet Res 2004 Sep 29;6(3):e34 [FREE Full text] [doi: 10.2196/jmir.6.3.e34] [Medline: 15471760]

26. Palen L, Hughes AL. Social Media in Disaster Communication. In: Handbook of Disaster Research. Cham: Springer; 2017:497-518.

27. Muniz-Rodriguez K, Ofori SK, Bayliss LC, Schwind JS, Diallo K, Liu M, et al. Social Media Use in Emergency Response to Natural Disasters: A Systematic Review With a Public Health Perspective. Disaster Med Public Health Prep 2020 Feb 09;14(1):139-149. [doi: 10.1017/dmp.2020.3] [Medline: 32148219]

28. Gerber N, Gerber P, Volkamer M. Explaining the privacy paradox: A systematic review of literature investigating privacy attitude and behavior. Comput Secur 2018 Aug;77:226-261. [doi: 10.1016/j.cose.2018.04.002]

29. Raij A, Ghosh A, Kumar S, Srivastava M. Privacy risks emerging from the adoption of innocuous wearable sensors in the mobile environment. 2011 Presented at: CHI '11: CHI Conference on Human Factors in Computing Systems; May 7-12, 2011; Vancouver, BC. [doi: 10.1145/1978942.1978945]

30. Goddard M. The EU General Data Protection Regulation (GDPR): European Regulation that has a Global Impact. Int J Mark Res 2017 Nov 01;59(6):703-705. [doi: 10.2501/ijmr-2017-050]

31. Cohen IG, Mello MM. HIPAA and Protecting Health Information in the 21st Century. JAMA 2018 Jul 17;320(3):231-232. [doi: 10.1001/jama.2018.5630] [Medline: 29800120]

32. Bargain O, Aminjonov U. Trust and compliance to public health policies in times of COVID-19. J Public Econ 2020 Dec;192:104316 [FREE Full text] [doi: 10.1016/j.jpubeco.2020.104316] [Medline: 33162621]

33. Ienca M, Vayena E. On the responsible use of digital data to tackle the COVID-19 pandemic. Nat Med 2020 Apr 27;26(4):463-464 [FREE Full text] [doi: 10.1038/s41591-020-0832-5] [Medline: 32284619]

34. Bunker D. Who do you trust? The digital destruction of shared situational awareness and the COVID-19 infodemic. Int J Inf Manage 2020 Dec;55:102201 [FREE Full text] [doi: 10.1016/j.ijinfomgt.2020.102201] [Medline: 32836649]

35. Liabo K, Boddy K, Bortoli S, Irvine J, Boult H, Fredlund M, et al. Public involvement in health research: what does 'good' look like in practice? Res Involv Engagem 2020 Mar 31;6(1):11 [FREE Full text] [doi: 10.1186/s40900-020-0183-x] [Medline: 32266085]

36. Lysne O. Containment of Untrusted Modules. In: The Huawei and Snowden Questions. Cham: Springer; 2018:99-107.

37. Consultation outcome: Online Harms White Paper. Department for Digital, Culture, Media & Sport. Government of the United Kingdom. 2020 Dec 15. URL: https://www.gov.uk/government/consultations/online-harms-white-paper/online-harms-white-paper [accessed 2022-06-01]

38. Zwitter A, Gstrein OJ. Big data, privacy and COVID-19 – learning from humanitarian expertise in data protection. Int J Humanitarian Action 2020 May 18;5(1). [doi: 10.1186/s41018-020-00072-6]

39. Ghafur S, Van Dael J, Leis M, Darzi A, Sheikh A. Public perceptions on data sharing: key insights from the UK and the USA. Lancet Digit Health 2020 Sep;2(9):e444-e446. [doi: 10.1016/s2589-7500(20)30161-8]

40. Panesar A. Machine Learning and AI Ethics. In: Machine Learning and AI for Healthcare: Big Data for Improved Health Outcomes. Berkeley, CA: Apress; Dec 16, 2020:207-247.

41. Williams SN, Armitage CJ, Tampe T, Dienes K. Public attitudes towards COVID-19 contact tracing apps: A UK-based focus group study. Health Expect 2021 Apr;24(2):377-385 [FREE Full text] [doi: 10.1111/hex.13179] [Medline: 33434404]

42. Ahmed N, Michelin RA, Xue W, Ruj S, Malaney R, Kanhere SS, et al. A Survey of COVID-19 Contact Tracing Apps. IEEE Access 2020;8:134577-134601. [doi: 10.1109/access.2020.3010226]

43. Budd J, Miller BS, Manning EM, Lampos V, Zhuang M, Edelstein M, et al. Digital technologies in the public-health response to COVID-19. Nat Med 2020 Aug 07;26(8):1183-1192. [doi: 10.1038/s41591-020-1011-4] [Medline: 32770165]

44. Gasser U, Ienca M, Scheibner J, Sleigh J, Vayena E. Digital tools against COVID-19: taxonomy, ethical challenges, and navigation aid. Lancet Digit Health 2020 Aug;2(8):e425-e434. [doi: 10.1016/s2589-7500(20)30137-0]

45. Santosh KC. COVID-19 Prediction Models and Unexploited Data. J Med Syst 2020 Aug 13;44(9):170 [FREE Full text] [doi: 10.1007/s10916-020-01645-z] [Medline: 32794042]

46. Rajpal A, Rahimi L, Ismail-Beigi F. Factors leading to high morbidity and mortality of COVID-19 in patients with type 2 diabetes. J Diabetes 2020 Dec 02;12(12):895-908 [FREE Full text] [doi: 10.1111/1753-0407.13085] [Medline: 32671936]

47. Kostkova P. Disease surveillance data sharing for public health: the next ethical frontiers. Life Sci Soc Policy 2018 Jul 04;14(1):16 [FREE Full text] [doi: 10.1186/s40504-018-0078-x] [Medline: 29971516]

Abbreviations

NHS: National Health Service