

---

Review

# Challenges and Potential Opportunities of Mobile Phone Call Detail Records in Health Research: Review

---

Kerina Helen Jones, BSc (Hons), PhD; Helen Daniels, BSc (Hons), PhD; Sharon Heys, BSc (Economics); David Vincent Ford, MBA

Data Science Building, School of Medicine, Swansea University, Swansea, United Kingdom

---

**Corresponding Author:**

Kerina Helen Jones, BSc (Hons), PhD  
Data Science Building  
School of Medicine  
Swansea University  
Singleton Park  
Swansea, SA28PP  
United Kingdom  
Phone: 44 1792602764  
Email: [k.h.jones@swansea.ac.uk](mailto:k.h.jones@swansea.ac.uk)

## Abstract

---

**Background:** Call detail records (CDRs) are collected by mobile network operators in the course of providing their service. CDRs are increasingly being used in research along with other forms of big data and represent an emerging data type with potential for public good. Many jurisdictions have infrastructures for health data research that could benefit from the integration of CDRs with health data.

**Objective:** The objective of this study was to review how CDRs have been used in health research and to identify challenges and potential opportunities for their wider use in conjunction with health data.

**Methods:** A literature review was conducted using structured search terms making use of major search engines. Initially, 4066 items were identified. Following screening, 46 full text articles were included in the qualitative synthesis. Information extracted included research topic area, population of study, datasets used, information governance and ethical considerations, study findings, and data limitations.

**Results:** The majority of published studies were focused on low-income and middle-income countries. Making use of the location element in CDRs, studies often modeled the transmission of infectious diseases or estimated population movement following natural disasters with a view to implementing interventions. CDRs were used in anonymized or aggregated form, and the process of gaining regulatory approvals varied with data provider and by jurisdiction. None included public views on the use of CDRs in health research.

**Conclusions:** Despite various challenges and limitations, anonymized mobile phone CDRs have been used successfully in health research. The use of aggregated data is a safeguard but also a further limitation. Greater opportunities could be gained if validated anonymized CDRs were integrated with routine health records at an individual level, provided that permissions and safeguards could be put in place. Further work is needed, including gaining public views, to develop an ethically founded framework for the use of CDRs in health research.

(*JMIR Mhealth Uhealth* 2018;6(7):e161) doi:[10.2196/mhealth.9974](https://doi.org/10.2196/mhealth.9974)

---

**KEYWORDS**

call detail records; mobile phone data; health research

## Introduction

---

**Background**

There are already over 5 billion unique mobile device subscribers globally, and the number of mobile connections

exceeds the world population at over 8 billion [1]. Mobile phone penetration is constantly rising and is predicted to exceed 5 billion users by 2019 [2]. As mobile phones are now an integral part of modern life, their potential to be used as a means of improving health care is increasingly promising. Call Detail Records (CDRs) are collected by mobile network operators

(MNOs) in the course of providing their service. Each time a mobile phone user connects to a mobile network, either by voice call or text message, a record is generated that includes the starting time of the call (or message), its duration, the caller and receiver phone numbers, and their locations [3]. Locations are estimated from the positions of activated cell towers and can be made more precise via tower triangulation and Wi-Fi connections [3]. Unlike interaction with mobile phone apps, CDRs result from passive data collection requiring no additional effort by the end user. MNOs receive billions of CDRs globally; they are necessary for billing, monitoring data usage, and for understanding and targeting customers according to their mobile phone use [3].

Due to the lack of landline infrastructures, mobile phones are the preferred method of communication in low- to middle-income countries (LMICs) [4], and they are playing a crucial part in these countries' socioeconomic developments [5]. Using mathematical modeling techniques, researchers are able to use CDRs to estimate the location of different populations and how this changes over time. Information on migratory patterns within and between countries can offer valuable information to policy makers in areas such as agriculture, transportation, poverty, conflict prevention, and disaster response and humanitarian aid [6]. In the last 5 years, CDR data have been used to improve health and health care, for example, via the generation of epidemiological models that can infer the spatial spread of infectious diseases from human mobility patterns [7].

Unlike for some of these LMIC settings, many developed countries have created data-intensive health research infrastructures, integrating multiple sources of anonymized routine health and administrative data for secondary uses. Main models vary between centralized repositories [8,9] and federated systems with distributed data nodes [10,11]. Typically, in a repository model, data are centralized for integration, whereas in a federated model data remain distributed among their original sources, with various permutations on these models and on how a user accesses data. Data provision to researchers can entail, subject to information governance regimes, external release of linked data to researchers [12], access via a data-safe haven and release of results [8], or by using privacy-preserving distributed data mining that computes distributed data without revealing sensitive information [13,14]. Address-based grid reference location data used with routine data present additional opportunities by enabling health geography studies but also present particular disclosure risks that must be mitigated so the data can be used safely [15]. CDRs represent an alternative type of spatial data that could add to, or augment, routine data collections to enable new opportunities for health research.

## Objective

The objective of this study was two-fold: (1) to review the ways in which CDRs (particularly the location elements) have been used in health research to identify the challenges encountered and benefits gained, and (2) to use this information to explore the issues that would need to be addressed to enable wider use of CDRs for health research, including their integration with routine health and administrative data.

## Methods

A literature review was conducted using structured search terms making use of major search engines. Predetermined eligibility criteria were set and adhered to in order to avoid the introduction of bias and to preclude the selection of studies on the basis of whether they favored a particular conclusion. To be included in the review, studies must have been published in the English language and in either peer-reviewed journals or conference proceedings. Studies must have used CDR data to answer a research question. Therefore, methodological papers, for example, outlining different mathematical approaches for analyzing mobile phone datasets, were excluded. Studies using data derived from mobile phone apps were also excluded. Research on any study population and health-related condition was included.

A search strategy was devised based on these inclusion criteria aided by identifying keywords from seminal works in this field [3]. The search terms were as follows: (mobile phone location data) OR (mobile phone call data) OR (mobile phone data) OR (cell phone data) OR (cell phone call data) OR (cell phone location data) OR (call detail records). These keywords were chosen in order to conduct a sensitive, rather than a specific search to ensure a higher probability of including all relevant articles. The strategy was customized according to the stipulations of each database for building search strings. Searches covered studies published up until January 2017 with no restrictions on the earliest date of publication. The following databases were searched from February 7-25, 2017: the Cochrane Central Register of Controlled Trials (CENTRAL) [16], Google Scholar [17], PubMed [18], Scopus [19], Web of Science [20], and WorldPop [21]. It was not intended as an exhaustive review, but we took a pragmatic approach to identify a comprehensive range of health-related studies to identify benefits and challenges with reasonable confidence.

All search results were imported into an online reference generator [22] and duplicate references were removed. The title and abstracts of these results were screened against the inclusion criteria as outlined above to identify potentially eligible studies for which the full texts were reviewed. Two reviewers independently performed the search, and disagreements between the reviewers were resolved by consensus. The reference lists of the eligible articles were searched to identify additional studies.

The following data were extracted for each eligible study: author, year published, title of article, country of population studied, research topic area, how datasets were used, findings, information governance, and limitations of the data.

## Results

### Overview

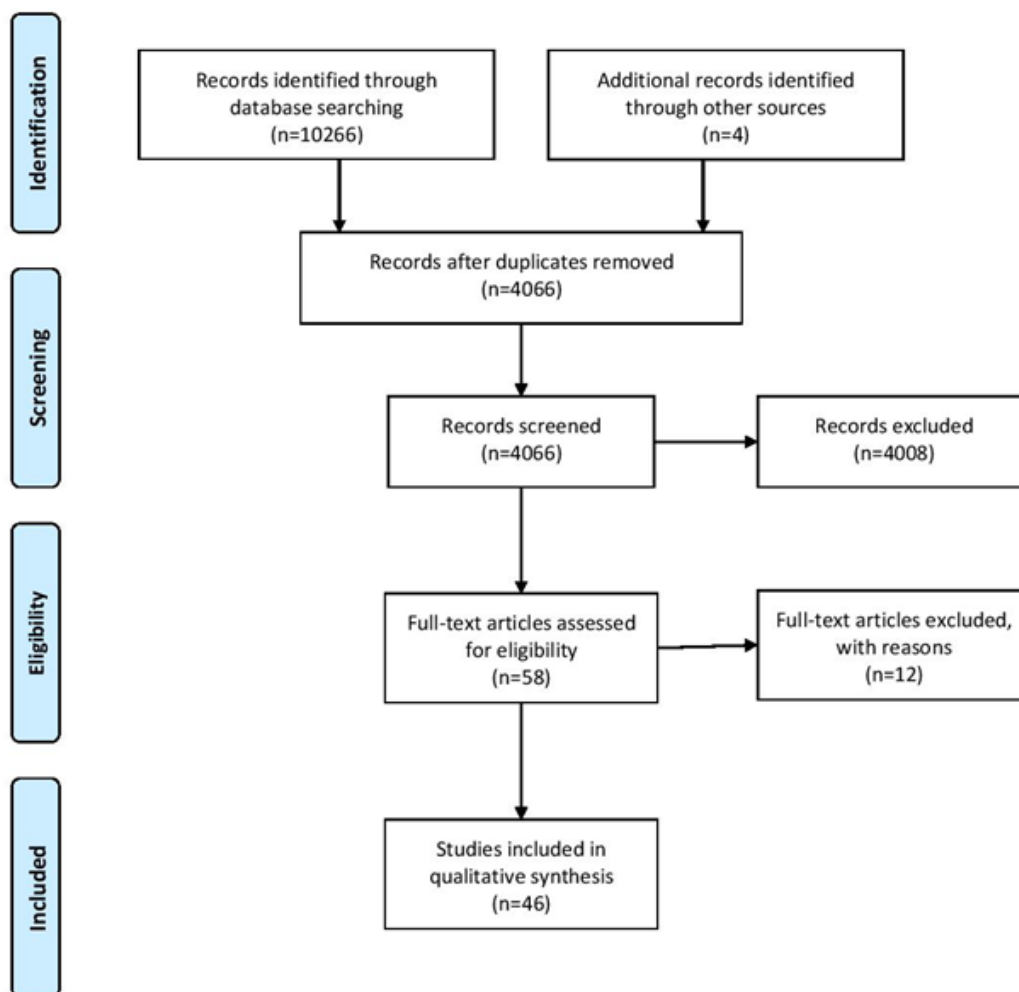
The search initially identified 4066 studies. Of these, 4008 were excluded on the basis of title and abstract alone as they did not meet the inclusion criteria. Most of the studies initially excluded tended to not be either research studies or health-related. Fifty-eight full-text articles were assessed for eligibility, and a

further 12 were subsequently rejected for not meeting the inclusion criteria. Of these, two papers were excluded as they were not research studies, but rather described anonymization processes; four were excluded as they did not use CDR data; five were excluded as they were geographical studies but not on health-related research; and one study was excluded as it was a methodology paper. Figure 1 shows the study process flowchart according to Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines [23]. As a result of the screening process, 46 full-text articles were included in the qualitative synthesis. Multimedia Appendix 1 presents the included studies [24-69]. Studies using CDR data for health research are summarized. CDR data were used alone or in conjunction with additional datasets for a variety of health-related purposes. In each study, CDR data were used in mathematical modeling to predict or identify population movement or to construct and predict social networks. A narrative description, based on the information provided in the publications, is given here to provide further details and draw out notable issues.

### Research Topic Area and Findings

As is evident from Multimedia Appendix 1, the majority of studies (n=42) focused on LMICs, with the other four using data from Belgium (n=2), Austria, and Italy. By using mathematical modeling, CDR location data were used to predict the movement of a population. Often, CDR data were used to develop models to address pressing concerns on infectious disease transfer. These included (numbers of studies shown where >1) malaria (n=11), HIV (n=3), cholera (n=3), influenza (n=3) dengue fever, Ebola virus, schistosomiasis, Rubella, meningitis, and tuberculosis. Other studies made use of CDR data to model population movement after disasters, to design and target public health interventions by identifying the location of at risk populations, to model hospital catchments, to model effects of air quality, and to explore options to arresting infectious diseases at outbreak. Where additional datasets were used, these were overlaid and compared with existing CDR data. For example, virological data were used to understand how cases of influenza were spreading. These were then used to verify mathematical models generated using CDR location data by mapping how populations move [27]. Although this is a method of integrating data, none of the included studies actually linked CDRs on an individual level to other datasets.

Figure 1. Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) diagram.



## Datasets and Data Governance

Thirteen of the articles were on studies from the 2013 Orange Data for Development Challenge on Mobile Phone Data focused on Ivory Coast [70]. Orange is a large, international mobile network operator, who allowed international research laboratories access to anonymized CDRs generated by 5 million of their Ivory Coast customers dating between December 2011 and April 2012. Researchers were tasked to use the data in a way that could potentially contribute to the socioeconomic development of the country. CDRs were anonymized by Orange Ivory Coast and processed by Orange Labs in Paris. In addition, the geographical locations of the mobile phone masts were blurred to protect the commercial interests of Orange. These datasets were then released to researchers [70].

Four datasets of varying granularity were provided by Orange:

1. The number and duration of calls between a pair of antennae aggregated by each hour. These data were provided for the whole of the observation period.
2. High spatial resolution data (which published antenna identifiers) of individual movement trajectories. To reduce the possibility of identification, these data were supplied on a random sample of 50,000 individuals for a 2-week time period only.
3. Data for the entire observation period but with reduced spatial resolution to mitigate the risk of identification. Spatial resolution was reduced by using the subprefectures of the mobile phone antenna location rather than specific antenna identifiers.
4. Social network subgraphs using call data generated by 5000 randomly selected individuals. These were divided into 2-week time periods for the entire duration of observation [70].

A year later, Orange issued a second challenge, this time using data from Senegal [71]. Twelve of the studies included in this review used the data provided by this challenge. The data were pseudonymized locally in Dakar by Sonatel in the first instance. Orange Labs in Paris then undertook several layers of anonymization on the data. As with the Ivory Coast data, the true geolocations of mobile phone masts were masked. Again, to reduce the risk of identification, datasets of high granularity were restricted in time span, and coarser aggregated data covered longer periods of time [72]. An internal ethics workgroup reviewed the governance of each application to receive access to the data. An additional safeguard was also put in place for this particular challenge. An external ethical review panel was set up that consisted of 14 international members who provided the Orange team with advice on information governance, particularly to review risks in publishing findings. As well as privacy concerns, the external review panel considered political concerns and issues of civil unrest (eg, regarding the Ebola epidemic).

Of the studies not involved with either of the Orange Data for Development challenges, all studies apart from three [38,39,64] stated that the mobile phone CDRs they used in their research were anonymized. Few studies described the anonymization process as this was usually undertaken by the MNO beforehand. However, we have assumed, at the very least, that personal

identifiers were removed from the data. MNO anonymization was not carried out in in one study [32] where researchers were provided with a raw dataset and undertook the anonymization process themselves. High- and low-volume users were excluded from analysis in this study to protect privacy. A second dataset from a different provider was used by the researchers in this same study. This dataset had been anonymized by the MNO: identifiable information was replaced by a hashed ID and encryption keys were exclusively managed by the MNO.

Wilson et al [67] stated that they followed the Groupe Spéciale Mobile Association (GSMA) privacy guidelines that advise that any analyses on mobile phone records should be done using de-identified data and that individual level data should not leave MNO servers [73]. All analyses conducted by this team were performed by connecting remotely to a Linux server with only aggregated data transferred outside the operator.

Wesolowski et al [66] detailed how they complied with the laws and regulations of Pakistan and the MNO Telenor, by using only data aggregated to tehsil level (an administrative unit of Pakistan). This was processed on their behalf by Telenor employees. The following measures were implemented to preserve the privacy of Telenor Pakistan's customers: (1) the CDR/mobility data were processed on a back-up and recovery server made available by Telenor Pakistan, with only Telenor employees having access to the detailed CDR/mobility data, (2) given the server arrangements, no detailed CDR/mobility data were taken out of Pakistan or left the premises of Telenor Pakistan, and (3) the processing of the detailed CDR/mobility data resulted in aggregations of the data on a tower-level granularity that was accessed only by Telenor employees.

Eight other studies mentioned explicitly that they used aggregated data in their analyses. There was no evidence to suggest that studies had a higher tendency to use aggregation if additional datasets had been used, but all additional datasets containing individual level data were de-identified. Three studies explicitly stated that they sought ethical approval from their own institution before beginning their research [28,36,57].

## Data Limitations

Several researchers commented on the limitations of the data, some of which are common to many data types and some specific to phone data. Gavric et al [41] reported that using aggregated data due to privacy concerns decreased the precision of their analyses. Differential mobile phone ownership due to financial means and socioeconomic status was noted as a potential source of bias whereby different sectors of a population may be over- or underrepresented [32,37,45,60]. For countries within Africa in particular, phone sharing is common, thus creating another potential source of bias since the network tracks the subscriber identification module (SIM), not the person. This would limit the value of studies focused on particular groups rather than general population movement [60].

## Discussion

### Principal Findings: Opportunities and Challenges

Emerging data types, such as mobile phone CDR data present valuable new opportunities for health research, particularly

when used in combination with additional datasets. This study reports on a structured literature review on the use of CDR data for health research, showing that the immense volumes of CDR data collected and held globally by MNOs can be used for public health benefit. Of particular interest is the location element that can be used to track phone user movement at various levels of granularity. Most of the studies included in this review used CDR data to create mathematical models based on population movement to predict the spread of disease epidemics. Where additional datasets were used, often these verified the validity of these models.

Using CDRs in health research has a number of benefits. These data are routinely and passively collected via mobile phones without any effort needed by the end user. It could be argued that creating these big datasets in this way is far easier and more effective than recruiting and consenting participants into a research study individually. Also, CDRs can be generated from basic mobile phones and do not require the use of smartphone. Therefore, their use in research does not preclude those from low socioeconomic groups in the way that GPS data would, for instance. An individual's home and work setting can be derived from CDR data, which is particularly valuable for countries where no integrated infrastructures exist for population census [74]. However, there are limitations that need to be taken into account in evaluating the use of CDRs in health research and in considering opportunities for their wider use. Data availability, formats, and levels of aggregation vary with MNO, and this will influence the type of analysis that can be done. For studies that require CDRs to be collected by multiple operators, researchers face the issue of having to join anonymized datasets: a difficult, although not insurmountable task [75]. There is an unknown level of discrepancy between phone owner and main user when phones are shared or bought for another person. This can call into question the validity of study findings focusing on particular demographics, since the extent of group representation is not known. Differential ownership of mobile phones among different sectors, particularly in LMICs where much of CDR research takes place, also calls into question the representativeness of the data and thus the findings that ensue [76]. However, it has also been observed that despite biases, there are few, if any, data sources that can provide such rich spatial and temporal movement data, particularly for much-needed research in LMICs [63,66].

A number of common patterns in data governance emerged from the review. Datasets were provided to researchers at different levels of spatial granularity and over variously restricted time periods to mitigate disclosure risks. These varied by study, or by programs of study. In some cases, data were subjected to several layers of anonymization and the true geolocations of mobile phone masts were masked. The use of anonymized (or strongly pseudonymized) data was the norm, with few studies outside this model and with many additionally using aggregation and the suppression of rare or extreme records. Reports of formal ethical review varied, but proposals were routinely submitted to an internal ethics workgroup, and in some cases to an independent external group for wider considerations such as political implications, societal benefits, and risk versus utility.

Concerns have been published on the ethics of using mobile phone data in research and the potential threats to privacy. Although MNOs have legal and organizational policies and researchers have jurisdictional health research governance to abide by, there is an absence of a clear, holistic, ethical, and regulatory framework to guide research using CDRs [3]. Most published research reports use anonymized data and many go further to protect against re-identification via aggregation. Nevertheless, examples of how anonymization and aggregation do not guarantee privacy in location data are abundant in the literature [77-79]. Furthermore, breaches in group privacy do not rely on the re-identification of individuals. It is considered that individuals who belong to certain groups on the basis of their gender, sexual orientation, ethnicity, or political preferences could become visible in CDR data [80]. There is clearly a need for an ethically founded framework for the use of CDRs in research for public good.

Considering that CDR data are collected about members of the general public, such a framework should take into account public viewpoints on the use of CDRs for research. Public engagement is, of course, a common feature in health research and in the use of large-scale, anonymized person-based data [81,82]. But it is acknowledged that similar work is both lacking and needful in the case of CDR data [3]. Although there have been surveys of public views on other aspects of mobile phone usage [83], there is no known literature on public perceptions of using CDRs for health research. Research using CDRs is still in its relative infancy, which means it is an ideal time to engage with the public so that their views can be taken into account in developing an ethically founded framework. This can be compared to similar work that has been done on genomic data sharing, another controversial area of health data science, where strong public views have necessitated the development of clear data sharing policies (eg, the Global Alliance for Genomics and Health [84]) [85-87].

### Wider Use and Future Work

From the review conducted in this study, plus seminal studies and reports [3,88], it is clear that CDR data can be valuable in health-related research. But as the majority of studies have focused on LMICs, the question arises of whether this can be translated into wider use in more developed countries, in particular where there are data-intensive infrastructures for health research that already have (or can gain) access to more traditional geolocation data, in the form of verified address-based grid references. Basically, it is a question of whether CDR data is valuable for health research per se, or only in particular settings. In support of the former position, some of the studies reviewed focused on Belgium, Austria, and Italy. Furthermore, Orange is not the only mobile network operator whose CDR datasets are being applied to health-related issues, as Telefonica have demonstrated via their Smart Steps initiative in various countries [89]. It may be that CDR data will prove to be a valuable resource for the public health sector, where, for example, the location element of the data could shed light on the way health promotion campaigns impact hospital attendances. Used in conjunction with additional datasets (eg, air quality monitoring databases, infection/virus outbreak data, emergency department attendance records), CDR location data

may bring about new opportunities for health research. However, during the course of this study, a number of issues have been revealed that impact on opportunities for the wider use of CDRs for health research and that can be learned from in moving forward. These are considered here in relation to their integration with other relevant datasets for use in data-intensive infrastructures, but they are also relevant to smaller-scale health research endeavors.

An essential and primary issue is data availability since CDRs reside with the network operator. This alone could be a show-stopper for various reasons, including that GSMA guidance states that individual level CDR data should not leave MNO servers [73,90]. Assuming data could be accessible, there would need to be suitable mutually beneficial agreements between the MNO and research program owner, with the prerequisite of a shared vision for data use [88]. Following modeling exercises in conjunction with Vodafone, the UK Office of National Statistics is planning to use CDRs to monitor commuter travel and collect census data [91]. There will, of course, be a financial cost for the use of the data, the extent of which is not known. But to be viable, this will need to be lower than current costs and/or provide valuable new information. Cost is a second major operator-related issue that may preclude the wider use of CDRs in health research unless suitable collaborative arrangements can be made.

Limitations inherent in the data (outlined earlier) need to be quantified and addressed for greater confidence in research findings. Encouragingly, there are reports that propose how bias in phone data can be addressed [74], but other issues remain. It would be useful to see a series of studies that have assessed the validity of CDRs in health research compared to other forms of geolocation data, for example, with Bluetooth data [92], photographic data [93], and flight data [94]. Such efforts would strengthen the evidence base on whether investment in the wider use of CDR data is warranted.

As with the use of any person-based data type, suitable physical, technical, and procedural controls need to be agreed between the data provider (MNO) and data user, as part of a proportionate data governance regime. This is true, though the stipulations may differ, whatever infrastructural and data access models are in place. The fact that anonymized and even aggregated data can pose identity disclosure risks is beyond doubt. Disclosure risks in the use of CDRs have not been studied to the same extent, and further work would be beneficial [15]. Ideally, this should be done in collaboration with an amenable network operator and be based on integrating CDRs with health records. However, this could be a challenge for reasons already described, but modeling a variety of data use scenarios using metadata (if available) could be a useful compromise and still yield meaningful information. This, along with the identified need for the input of public views on the use of CDR data for health research, would also inform the much needed consistent, holistic, ethically founded framework.

Public engagement forms an important part of an ethical framework for data use, beyond strict legal compliance requirements. Consultation with the public to gain their views

on the use of CDRs for health research would gauge knowledge and expectations and, as with other emerging data types, would inform the socially acceptable use of the data. This is not too much to ask, since after all, CDRs originate with individuals and are based on their activities. It would be interesting to know more about the public's actual awareness of data collected by mobile operators and the ways data are used in-house and by external agencies.

A last consideration is of the ultimate feasibility of integrating CDRs with health record data in terms of whether there is an appetite for it among the participating stakeholders. This is not a question that can be answered easily without seeing how the many challenges could be addressed so that the risk versus utility of such a development could be evaluated. A definite appetite and strong drivers to move initiatives forward will be needed for success.

### Strengths and Limitations

This study adds to the discussion on the suitability of CDRs for health research and raises issues to be addressed if the wider use of CDRs is to become a reality. It is the only known study to carry out a review of publications using CDRs in health research in order to learn from their practices and identify challenges. It also considers the use of CDRs in relation to data-intensive infrastructures and sets out problems to be solved to enable informed decisions on whether investment in seeking CDR integration is warranted. In accordance with other authors, this study recommends the development of an ethically founded framework for the use of CDRs in health research, but furthermore, it recommends that public views on the use of CDR data should be integral.

However, there are some limitations to consider. This study focused on articles using CDRs for health-related topics published in peer-reviewed journals and conference proceedings; it does not claim to be an exhaustive review. The use of health monitoring apps is outside the scope of the study, like Ginger.io [95] and the work of Sandy Pentland who, for example, used behavioral information from voice recordings and texting to detect signs of posttraumatic stress disorder in returning soldiers [96]. It also drew on reports published by MNOs and other organizations. However, operators will also have proprietary information not shared publicly, which may contain planned developments to address pertinent issues and take forward the use of CDRs for public benefit.

### Conclusion

All things considered, there are possibilities for the wider use of CDR data in health research but there are also major challenges to be addressed. Some important points have been discussed here, but this is not an exhaustive list of issues. Questions remain around the suitability of CDRs for wider use in health research and particularly as part of data-intensive infrastructures for population-scale studies. A concerted effort will be needed to create solutions to determine if mobile phone CDRs are a worthwhile data type to pursue and invest in to augment currently used geolocation data.

## Acknowledgments

The authors acknowledge funding for the study from the Natural Environmental Research Council.

## Conflicts of Interest

None declared.

## Multimedia Appendix 1

Summary of included studies.

[[PDF File \(Adobe PDF File\), 80KB - mhealth\\_v6i7e161\\_app1.pdf](#)]

## References

1. GSMA Intelligence. 2018. URL: <https://www.gsmaintelligence.com/> [accessed 2018-07-11] [WebCite Cache ID 70pfRxy8b]
2. Statista. 2018. URL: <https://www.statista.com/statistics/470018/mobile-phone-user-penetration-worldwide/> [WebCite Cache ID 70pnBIHWe]
3. Vinck P, Letouzé E, Kammourieh L. DataPop Alliance. 2015. The law, politics and ethics of cell phone data analytics URL: [http://datapopalliance.org/wpcontent/uploads/2015/04/WPS\\_LawPoliticsEthicsCellPhoneDataAnalytics.pdf](http://datapopalliance.org/wpcontent/uploads/2015/04/WPS_LawPoliticsEthicsCellPhoneDataAnalytics.pdf) [accessed 2018-07-11] [WebCite Cache ID 70pnJbx6T]
4. How mobile phones are changing the developing world.: Consumer Technology Association; 2017. URL: <https://www.cta.tech/News/Blog/Articles/2015/July/How-Mobile-Phones-Are-Changing-the-Developing-World.aspx> [WebCite Cache ID 70pnXKpg5]
5. Arie S. Can mobile phones transform healthcare in low and middle income countries? *BMJ* 2015 Apr 22;350(apr22):h1975-h1975 [FREE Full text] [doi: [10.1136/bmj.h1975](https://doi.org/10.1136/bmj.h1975)]
6. Mobile policy handbook: an insider's guide to the issues.: GSMA; 2017. URL: [https://www.gsma.com/publicpolicy/wpcontent/uploads/2016/08/Mobile\\_Policy\\_Handbook\\_2017\\_EN.pdf](https://www.gsma.com/publicpolicy/wpcontent/uploads/2016/08/Mobile_Policy_Handbook_2017_EN.pdf) [accessed 2018-07-11] [WebCite Cache ID 70pnhtw80]
7. Mobile phone network data for development.: Global Pulse; 2013. URL: [http://www.unglobalpulse.org/sites/default/files/Mobile%20Data%20for%20Development%20Primer\\_Oct2013.pdf](http://www.unglobalpulse.org/sites/default/files/Mobile%20Data%20for%20Development%20Primer_Oct2013.pdf) [accessed 2018-07-11] [WebCite Cache ID 70pnkjhGI]
8. SAIL Databank. URL: <https://saildatabank.com/> [WebCite Cache ID 70pnseCAi]
9. Population Data BC. 2018. URL: <https://www.popdata.bc.ca/> [WebCite Cache ID 70poL1fC1]
10. CNODES. URL: <https://www.cnodes.ca/> [accessed 2018-07-11] [WebCite Cache ID 70pwBoYVs]
11. SHRINE. URL: <https://catalyst.harvard.edu/services/shrine/> [WebCite Cache ID 70pwIE2Ht]
12. Data Linkage Western Australia. URL: <https://www.datalinkage-wa.org.au/> [accessed 2018-07-11] [WebCite Cache ID 70pwQvUK2]
13. Aggarwal CC, Yu PS. A General Survey of Privacy-Preserving Data Mining Models and Algorithms. In: Aggarwal CC, Yu PS, editors. *Privacy-Preserving Data Mining. Advances in Database Systems*. New York, NY: Springer; 2008:11.
14. Hailemichael MA, Yigzaw KY, Bellika JG. Emnet: A tool for privacy-preserving statistical computing on distributed health data. In: 13th Scandinavian Conference on Health Informatics. 2015 Presented at: Scandinavian Conference on Health Informatics; June 15-17, 2015; Tromsø, Norway.
15. Carter A, Liddle J, Hall W, Chenery H. Mobile Phones in Research and Treatment: Ethical Guidelines and Future Directions. *JMIR Mhealth Uhealth* 2015 Oct 16;3(4):e95 [FREE Full text] [doi: [10.2196/mhealth.4538](https://doi.org/10.2196/mhealth.4538)] [Medline: [26474545](https://pubmed.ncbi.nlm.nih.gov/26474545/)]
16. The Cochrane Central Register of Controlled Trials (CENTRAL). URL: <http://www.cochranelibrary.com/about/central-landing-page.html> [accessed 2018-07-11] [WebCite Cache ID 70pwW4tm0]
17. Google Scholar. URL: <https://scholar.google.co.uk/> [accessed 2018-07-11] [WebCite Cache ID 70pwYm3hZ]
18. PubMed. URL: <https://www.ncbi.nlm.nih.gov/pubmed/> [accessed 2018-07-11] [WebCite Cache ID 70pockTas]
19. Scopus. URL: <https://www.elsevier.com/solutions/scopus> [accessed 2018-07-11] [WebCite Cache ID 70pwbEWGL]
20. Web of Science. URL: <http://apps.webofknowledge.com/> [WebCite Cache ID 70pwmVyxe]
21. Worldpop. URL: [http://worldpop.org.uk/about\\_our\\_work/publications%20/](http://worldpop.org.uk/about_our_work/publications%20/) [accessed 2018-07-11] [WebCite Cache ID 70pwtYFdk]
22. Cite this for me. URL: <http://www.citethisforme.com/walkthrough> [accessed 2018-07-11] [WebCite Cache ID 70pwwQUZy]
23. Liberati A, Altman DG, Tetzlaff J, Mulrow C, Gotzsche PC, Ioannidis JPA, et al. The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate healthcare interventions: explanation and elaboration. *BMJ* 2009 Jul 21;339(jul21 1):b2700-b2700 [FREE Full text] [doi: [10.1136/bmj.b2700](https://doi.org/10.1136/bmj.b2700)]
24. Agarwal R, Gauthier V, Becker M. Information dissemination using human mobility in realistic environment (E-Inspire). In: D4D Challenge 2013 special session. 2013 Presented at: NetMob 2013; May 1-3, 2013; Boston, MA URL: <https://perso.uclouvain.be/vincent.blondel/netmob/2013/D4D-book.pdf>

25. Arai A, Honorant A, Witayangkurn R, Shibasaki R. Neighbourhood structures in socio-demographic and HIV infection conditions: indication of the potential of mHealth for tackling HIV/AIDS in Ivory Coast. 2013 Presented at: NetMob 2013; May 1-3, 2013; Boston, MA URL: <https://perso.uclouvain.be/vincent.blondel/netmob/2013/D4D-book.pdf>
26. Azman A, Urquhart E, Zaitchik B, Lesser J. Using mobile phone data to supercharge epidemic models of cholera transmission in Africa: a case study of Cote d'Ivoire. 2013 Presented at: Boston, MA; May 1-3, 2013; NetMob 2013 URL: <https://perso.uclouvain.be/vincent.blondel/netmob/2013/D4D-book.pdf>
27. Baldo N, Closas P. Disease outbreak detection by mobile network monitoring: a case study with the D4D datasets. 2013 Presented at: NetMob 2013; May 1-3, 2013; Boston, MA URL: <https://perso.uclouvain.be/vincent.blondel/netmob/2013/D4D-book.pdf>
28. Bengtsson L, Lu X, Thorson A, Garfield R, von Schreeb J. Improved response to disasters and outbreaks by tracking population movements with mobile phone network data: a post-earthquake geospatial study in Haiti. *PLoS Med* 2011 Aug;8(8):e1001083 [FREE Full text] [doi: [10.1371/journal.pmed.1001083](https://doi.org/10.1371/journal.pmed.1001083)] [Medline: [21918643](https://pubmed.ncbi.nlm.nih.gov/21918643/)]
29. Bengtsson L, Gaudart J, Lu X, Moore S, Wetter E, Sallah K, et al. Using mobile phone data to predict the spatial spread of cholera. *Sci Rep* 2015 Mar 09;5:8923 [FREE Full text] [doi: [10.1038/srep08923](https://doi.org/10.1038/srep08923)] [Medline: [25747871](https://pubmed.ncbi.nlm.nih.gov/25747871/)]
30. Chunara R, Nsoesie EO. Large-scale measurements of network topology and disease spread: a pilot evaluation using mobile phone data in Cote d'Ivoire. 2013 Presented at: NetMob 2013; May 1-3, 2013; Boston, MA URL: <https://perso.uclouvain.be/vincent.blondel/netmob/2013/D4D-book.pdf>
31. Deka D, Vishnawath S. On models characterizing cellular social networks. 2013 Presented at: NetMob 2013; May 1-3, 2013; Boston, MA URL: <https://perso.uclouvain.be/vincent.blondel/netmob/2013/D4D-book.pdf>
32. De Monasterio J, Salles A, Lang C, Weinberg D, Minnoni M, Travizano M, et al. Analyzing the spread of chagas disease with mobile phone data. 2016 Presented at: IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM); Aug. 18-21, 2016; San Francisco, CA URL: <http://ieeexplore.ieee.org/document/7752298/> [doi: [10.1109/ASONAM.2016.7752298](https://doi.org/10.1109/ASONAM.2016.7752298)]
33. Dewulf B, Neutens T, Van Dyck D, de Bourdeaudhuij I, Int Panis L, Beckx C, et al. Dynamic assessment of inhaled air pollution using GPS and accelerometer data. *Journal of Transport & Health* 2016 Mar;3(1):114-123. [doi: [10.1016/j.jth.2015.10.004](https://doi.org/10.1016/j.jth.2015.10.004)]
34. Djotio NT, Nganmeni Z, Nouho N. Design and implementation of a tool for the Correlation between the rate of prevalence of a pathology and the flow of communication between diverse localities. 2013 Presented at: NetMob 2013; May 1-3, 2013; Boston, MA URL: <https://perso.uclouvain.be/vincent.blondel/netmob/2013/D4D-book.pdf>
35. Enns EA, Amuasi JH. Human mobility and communication patterns in Cote d'Ivoire: a network perspective for malaria control. 2013 Presented at: NetMob 2013; May 1-3, 2013; Boston, MA URL: <https://perso.uclouvain.be/vincent.blondel/netmob/2013/D4D-book.pdf>
36. Farrahi K, Emonet R, Cebrian M. Epidemic contact tracing via communication traces. *PLoS One* 2014;9(5):e95133 [FREE Full text] [doi: [10.1371/journal.pone.0095133](https://doi.org/10.1371/journal.pone.0095133)] [Medline: [24787614](https://pubmed.ncbi.nlm.nih.gov/24787614/)]
37. Finger F, Genolet T, Mari L, de Magny GC, Manga NM, Rinaldo A, et al. Mobile phone data highlights the role of mass gatherings in the spreading of cholera outbreaks. *Proc Natl Acad Sci U S A* 2016 Jun 07;113(23):6421-6426 [FREE Full text] [doi: [10.1073/pnas.1522305113](https://doi.org/10.1073/pnas.1522305113)] [Medline: [27217564](https://pubmed.ncbi.nlm.nih.gov/27217564/)]
38. Frias-Martinez V, Rubio A, Frias-Martinez E. Measuring the impact of epidemic alerts on human mobility. 2013. URL: <http://www.vanessafriasmartinez.org/uploads/epidemics.pdf> [accessed 2018-07-11] [WebCite Cache ID 70px93ht9]
39. Frias-Martinez E, Williamson G, Frias-Martinez V. Simulation of Epidemic Spread using Cell-Phone Call Data: H1N1 Case Study. 2014. URL: <http://www.vanessafriasmartinez.org/uploads/h1n1netmob13.pdf> [accessed 2018-07-11] [WebCite Cache ID 70pxF99nF]
40. Gariazzo C, Pelliccioni A, Bolignano A. A dynamic urban air pollution population exposure assessment study using model and population density data derived by mobile phone traffic. *Atmospheric Environment* 2016 Apr;131:289-300 [FREE Full text] [doi: [10.1016/j.atmosenv.2016.02.011](https://doi.org/10.1016/j.atmosenv.2016.02.011)]
41. Gavric K, Brdar S, Culibrk D, Crnojevic V. Linking the human mobility and connectivity patterns with spatial HIV distribution. 2013 Presented at: NetMob 2013; May 1-3, 2013; Boston, MA URL: <https://perso.uclouvain.be/vincent.blondel/netmob/2013/D4D-book.pdf>
42. Guzetta G, Staiano G, Ajell M, Kirschner D, Lepri B, Merler S. Impact and scale of human mobility on the spatial transmission dynamics of tuberculosis in Senegal. 2014. URL: [http://www.d4d.orange.com/fr/content/download/43453/406503/version/1/file/D4DChallengeSenegal\\_Book\\_of\\_Abstracts\\_Scientific\\_Papers.pdf](http://www.d4d.orange.com/fr/content/download/43453/406503/version/1/file/D4DChallengeSenegal_Book_of_Abstracts_Scientific_Papers.pdf)
43. Hamaoka Y. Detecting Anomalies and Supporting Community To Ensure Healthy Society. 2014. URL: [http://www.d4d.orange.com/fr/content/download/43453/406503/version/1/file/D4DChallengeSenegal\\_Book\\_of\\_Abstracts\\_Scientific\\_Papers.pdf](http://www.d4d.orange.com/fr/content/download/43453/406503/version/1/file/D4DChallengeSenegal_Book_of_Abstracts_Scientific_Papers.pdf)
44. Kafsi M, Kazemi E, Maystre L, Yartseva L, Grossglauser M, Thiran P. Mitigating epidemics through mobile micro-measures. 2013 Presented at: NetMob 2013; May 1-3, 2013; Boston, MA URL: <https://perso.uclouvain.be/vincent.blondel/netmob/2013/D4D-book.pdf>
45. Le Menach A, Tatem AJ, Cohen JM, Hay SI, Randell H, Patil AP, et al. Travel risk, malaria importation and malaria transmission in Zanzibar. *Sci Rep* 2011 Sep;1:93 [FREE Full text] [doi: [10.1038/srep00093](https://doi.org/10.1038/srep00093)] [Medline: [22355611](https://pubmed.ncbi.nlm.nih.gov/22355611/)]



46. Leidig JP, Kitsumi Y, O'Hearn KA, Sauer CM, Scripps J, Wolffe G. Applying mobile datasets in computational public health research. 2013 Presented at: NetMob 2013; May 1-3, 2013; Boston, MA URL: <https://perso.uclouvain.be/vincent.blondel/netmob/2013/D4D-book.pdf>
47. Leidig JP, Theisen C, Vogel N, Graham DH, Scripps J, Wolffe G. Modeling Ebola Virus Diffusion in Senegal using Mobile Phone Datasets and Agent-based Simulation. 2014. URL: [http://www.d4d.orange.com/fr/content/download/43453/406503/version/1/file/D4DChallengeSenegal\\_Book\\_of\\_Abstracts\\_Scientific\\_Papers.pdf](http://www.d4d.orange.com/fr/content/download/43453/406503/version/1/file/D4DChallengeSenegal_Book_of_Abstracts_Scientific_Papers.pdf)
48. Lima A, De Domenico M, Pejovic V, Musolesi M. Exploiting cellular data for disease containment and information campaigns strategies in country-wide epidemics. 2013 Presented at: NetMob 2013; May 1-3, 2013; Boston, MA URL: <https://perso.uclouvain.be/vincent.blondel/netmob/2013/D4D-book.pdf>
49. Lima A, Pejovic V, Rossi L, Musolesi M, Gonzalez M. Progmosis: Evaluating Risky Individual Behaviour During Epidemics Using Mobile Network Data. 2014. URL: [http://www.d4d.orange.com/fr/content/download/43453/406503/version/1/file/D4DChallengeSenegal\\_Book\\_of\\_Abstracts\\_Scientific\\_Papers.pdf](http://www.d4d.orange.com/fr/content/download/43453/406503/version/1/file/D4DChallengeSenegal_Book_of_Abstracts_Scientific_Papers.pdf)
50. Mari L, Casagrandi R, Ciddio M, Sokolow SH, De Leo G, Gatto M. Uncovering the impact of human mobility on schistosomiasis via mobile phone data. 2014. URL: [http://www.d4d.orange.com/fr/content/download/43453/406503/version/1/file/D4DChallengeSenegal\\_Book\\_of\\_Abstracts\\_Scientific\\_Papers.pdf](http://www.d4d.orange.com/fr/content/download/43453/406503/version/1/file/D4DChallengeSenegal_Book_of_Abstracts_Scientific_Papers.pdf)
51. Matamalas JT, Domenico MD, Arenas A. Weakening the Incidence of Transmittable Diseases by Taking Advantage of Mobile Phone Activity. 2014. URL: [http://www.d4d.orange.com/fr/content/download/43453/406503/version/1/file/D4DChallengeSenegal\\_Book\\_of\\_Abstracts\\_Scientific\\_Papers.pdf](http://www.d4d.orange.com/fr/content/download/43453/406503/version/1/file/D4DChallengeSenegal_Book_of_Abstracts_Scientific_Papers.pdf)
52. Milushelva S. Quantifying the Effect of Movement Associated with Holidays on Malaria Prevalence Using Cell Phone Data. 2014. URL: [http://www.d4d.orange.com/fr/content/download/43453/406503/version/1/file/D4DChallengeSenegal\\_Book\\_of\\_Abstracts\\_Scientific\\_Papers.pdf](http://www.d4d.orange.com/fr/content/download/43453/406503/version/1/file/D4DChallengeSenegal_Book_of_Abstracts_Scientific_Papers.pdf)
53. Mutafungwa E. Mobile Data as Public Health Decision Enabler: A Case Study of Cardiac and Neurological Emergencies. 2014. URL: [http://www.d4d.orange.com/fr/content/download/43453/406503/version/1/file/D4DChallengeSenegal\\_Book\\_of\\_Abstracts\\_Scientific\\_Papers.pdf](http://www.d4d.orange.com/fr/content/download/43453/406503/version/1/file/D4DChallengeSenegal_Book_of_Abstracts_Scientific_Papers.pdf)
54. Perez-Saez J, Finger F, Mari L, Rinaldo A, Bertuzzo E. Human mobility and the spreading of waterborne diseases. 2014. URL: [http://www.d4d.orange.com/fr/content/download/43453/406503/version/1/file/D4DChallengeSenegal\\_Book\\_of\\_Abstracts\\_Scientific\\_Papers.pdf](http://www.d4d.orange.com/fr/content/download/43453/406503/version/1/file/D4DChallengeSenegal_Book_of_Abstracts_Scientific_Papers.pdf)
55. Resch B, Arif A, Krings G, Vankeerberghen G, Buekenhout M. Deriving hospital catchment areas from mobile phone data. 2016. URL: [http://www.berndresch.com/download/work/publications/resch-et-al\\_catchment-areas-mobile-phones\\_giscience\\_2016.pdf](http://www.berndresch.com/download/work/publications/resch-et-al_catchment-areas-mobile-phones_giscience_2016.pdf) [WebCite Cache ID 70ps5AJD2]
56. Saravanan M, Karthikeyan P, Aarthi A, Kiruthika M, Suganya S. Exploring community structure to understand disease spread and control using mobile call detail records. 2013 Presented at: NetMob 2013; May 1-3, 2013; Boston, MA URL: <https://perso.uclouvain.be/vincent.blondel/netmob/2013/D4D-book.pdf>
57. Tatem A, Huang Z, Narib C, Kumar U, Kandula D, Pindolia D, et al. Integrating rapid risk mapping and mobile phone call record data for strategic malaria elimination planning. *Malar J* 2014 Feb 10;13:52 [FREE Full text] [doi: [10.1186/1475-2875-13-52](https://doi.org/10.1186/1475-2875-13-52)] [Medline: [24512144](https://pubmed.ncbi.nlm.nih.gov/24512144/)]
58. Tatem AJ, Qiu Y, Smith DL, Sabot O, Ali AS, Moonen B. The use of mobile phone data for the estimation of the travel patterns and imported Plasmodium falciparum rates among Zanzibar residents. *Malar J* 2009 Dec 10;8:287 [FREE Full text] [doi: [10.1186/1475-2875-8-287](https://doi.org/10.1186/1475-2875-8-287)] [Medline: [20003266](https://pubmed.ncbi.nlm.nih.gov/20003266/)]
59. Tompkins AM, McCreech N. Developing an agent based migration model for Senegal for malaria transmission. 2014. URL: [http://www.d4d.orange.com/fr/content/download/43453/406503/version/1/file/D4DChallengeSenegal\\_Book\\_of\\_Abstracts\\_Scientific\\_Papers.pdf](http://www.d4d.orange.com/fr/content/download/43453/406503/version/1/file/D4DChallengeSenegal_Book_of_Abstracts_Scientific_Papers.pdf)
60. Tompkins AM, McCreech N. Migration statistics relevant for malaria transmission in Senegal derived from mobile phone data and used in an agent-based migration model. *Geospat Health* 2016 Mar 31;11(1 Suppl):408 [FREE Full text] [doi: [10.4081/gh.2016.408](https://doi.org/10.4081/gh.2016.408)] [Medline: [27063741](https://pubmed.ncbi.nlm.nih.gov/27063741/)]
61. Vogel N. Mining Enormous Mobile Datasets to Improve Mitigation Strategies for Limiting the Spread of Infectious Disease. 2014 Presented at: Big Data Conference 2014; 2014; Michigan URL: [http://scholarworks.gvsu.edu/bigdata\\_conference2014/12/](http://scholarworks.gvsu.edu/bigdata_conference2014/12/)
62. Wesolowski A, Buckee C. Are gravity models appropriate for estimating the spatial spread of malaria? 2013 Presented at: NetMob 2013; May 1-3, 2013; Boston, MA URL: <https://perso.uclouvain.be/vincent.blondel/netmob/2013/D4D-book.pdf>
63. Wesolowski A, Metcalf C, Eagle N, Kombich J, Grenfell BT, Bjørnstad ON, et al. Quantifying seasonal population fluxes driving rubella transmission dynamics using mobile phone data. *Proc Natl Acad Sci U S A* 2015 Sep 01;112(35):11114-11119 [FREE Full text] [doi: [10.1073/pnas.1423542112](https://doi.org/10.1073/pnas.1423542112)] [Medline: [26283349](https://pubmed.ncbi.nlm.nih.gov/26283349/)]
64. Wesolowski A, Eagle N, Tatem AJ, Smith DL, Noor AM, Snow RW, et al. Quantifying the impact of human mobility on malaria. *Science* 2012 Oct 12;338(6104):267-270 [FREE Full text] [doi: [10.1126/science.1223467](https://doi.org/10.1126/science.1223467)] [Medline: [23066082](https://pubmed.ncbi.nlm.nih.gov/23066082/)]
65. Wesolowski A, Stresman G, Eagle N, Stevenson J, Owaga C, Marube E, et al. Quantifying travel behavior for infectious disease research: a comparison of data from surveys and mobile phones. *Sci Rep* 2014 Jul 14;4:5678 [FREE Full text] [doi: [10.1038/srep05678](https://doi.org/10.1038/srep05678)] [Medline: [25022440](https://pubmed.ncbi.nlm.nih.gov/25022440/)]

66. Wesolowski A, Qureshi T, Boni MF, Sundsøy PR, Johansson MA, Rasheed SB, et al. Impact of human mobility on the emergence of dengue epidemics in Pakistan. *Proc Natl Acad Sci U S A* 2015 Sep 22;112(38):11887-11892 [[FREE Full text](#)] [doi: [10.1073/pnas.1504964112](https://doi.org/10.1073/pnas.1504964112)] [Medline: [26351662](https://pubmed.ncbi.nlm.nih.gov/26351662/)]
67. Wilson R, Zu E, Albert M, Power D, Tudge S, Gonzalez M, et al. Rapid and Near Real-Time Assessments of Population Displacement Using Mobile Phone Data Following Disasters: The 2015 Nepal Earthquake. *PLoS Curr* 2016 Feb 24;8 [[FREE Full text](#)] [doi: [10.1371/currents.dis.d073fbcce328e4c39087bc086d694b5c](https://doi.org/10.1371/currents.dis.d073fbcce328e4c39087bc086d694b5c)] [Medline: [26981327](https://pubmed.ncbi.nlm.nih.gov/26981327/)]
68. Wu H, Gosh S, Ramakrishnan N. Forecasting Influenza in Senegal with Call Detail Records. 2014. URL: [https://people.cs.vt.edu/~prithwi/resources/papers/hao\\_netmob14\\_d4dSenegal.pdf](https://people.cs.vt.edu/~prithwi/resources/papers/hao_netmob14_d4dSenegal.pdf) [accessed 2018-07-11] [[WebCite Cache ID 70q02R1cr](#)]
69. Yu Y. Individual-based modelling of contact networks of epidemic diffusion using real-life data. 2014. URL: [http://www.d4d.orange.com/fr/content/download/43453/406503/version/1/file/D4DChallengeSenegal\\_Book\\_of\\_Abstracts\\_Scientific\\_Papers.pdf](http://www.d4d.orange.com/fr/content/download/43453/406503/version/1/file/D4DChallengeSenegal_Book_of_Abstracts_Scientific_Papers.pdf)
70. Blondel V, Esch M, Chan C, Clerot F, Deville P, Huens E, et al. The D4D challenge on mobile phone data. 2013. URL: <https://arxiv.org/pdf/1210.0137.pdf> [accessed 2018-07-11] [[WebCite Cache ID 70ptEb6vq](#)]
71. Orange. Challenge 4 development. 2014. URL: <http://www.d4d.orange.com/en/Accueil> [[WebCite Cache ID 70q253A17](#)]
72. De Montjoye Y, Smoreda Z, Trinquart R, Ziemlicki C, Blondel V. D4D-Senegal: The Second mobile phone data for development challenge. 2014. URL: <https://arxiv.org/pdf/1407.4885.pdf> [accessed 2018-07-11] [[WebCite Cache ID 70q001oig](#)]
73. Mobile privacy principles - public policy.: GSMA; 2012. URL: <http://www.gsma.com/publicpolicy/mobile-and-privacy/mobile-privacy-principles> [accessed 2018-07-11] [[WebCite Cache ID 70ptHLfef](#)]
74. Pestre G, Letouze E, Zagheni E. Assessing biases in call detail records for development estimates, , No. AFD Research Paper Series 2016;32.
75. Wang K, Fung B, Dong G. Integrating private databases for data analysis. In: Kantor P, Roberts F, Wang FY, Muresan G, Zeng D, Chen H, et al, editors. *Intelligence and Security Informatics*. Berlin: Springer; 2005:171-182.
76. Wesolowski A, Buckee C, Engø-Monsen KC, Metcalf C. Connecting Mobility to Infectious Diseases: The Promise and Limits of Mobile Phone Data. *J Infect Dis* 2016 Nov 14;214(suppl 4):S414-S420. [doi: [10.1093/infdis/jiw273](https://doi.org/10.1093/infdis/jiw273)]
77. Zang H, Bolot J. Anonymization of location data does not work. In: *MobiCom '11*. 2011 Presented at: 17th Annual International Conference on Mobile Computing and Networking; Sept. 19-23, 2011; Las Vegas, NV URL: <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.651.44&rep=rep1&type=pdf>
78. Song Y, Dahlmeier D, Bressan S. Not So Unique in the Crowd: a Simple and Effective Algorithm for Anonymizing Location Data. 2014. URL: <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.663.4441&rep=rep1&type=pdf>
79. Narayanan A, Felten E. No silver bullet: De-identification still doesn't work. 2014. URL: <http://randomwalker.info/publications/no-silver-bullet-de-identification.pdf> [accessed 2018-07-11] [[WebCite Cache ID 70q0Yjq6M](#)]
80. Taylor L, Floridi L, Sloot B. *Group Privacy. New Challenges of Data Technologies*. Berlin: Springer; 2016.
81. Cameron D, Pope S, Clemence M. Dialogue on data: Exploring the public's views on using administrative data for research purposes. 2014. URL: <https://esrc.ukri.org/files/public-engagement/public-dialogues/dialogue-on-data-exploring-the-public-s-views-on-using-linked-administrative-data-for-research-purposes/> [accessed 2018-07-11] [[WebCite Cache ID 70q0mxHMf](#)]
82. Wellcome Trust. The One-Way Mirror: Public attitudes to commercial access to health data. 2016. URL: <https://wellcome.ac.uk/sites/default/files/public-attitudes-to-commercial-access-to-health-data-wellcome-mar16.pdf> [accessed 2018-07-11] [[WebCite Cache ID 70q0qgGJt](#)]
83. GSMA. Mobile Privacy: Consumer research insights and considerations for policymakers. 2014. URL: [https://www.gsma.com/publicpolicy/wp-content/uploads/2014/02/MOBILE\\_PRIVACY\\_Consumer\\_research\\_insights\\_and\\_considerations\\_for\\_policymakers-Final.pdf](https://www.gsma.com/publicpolicy/wp-content/uploads/2014/02/MOBILE_PRIVACY_Consumer_research_insights_and_considerations_for_policymakers-Final.pdf) [accessed 2018-07-11] [[WebCite Cache ID 70q0tYII0](#)]
84. GA4GH. Global Alliance for Genomics and Health. 2014. URL: <https://www.ga4gh.org/> [accessed 2018-07-11] [[WebCite Cache ID 70q0vxc9q](#)]
85. Shabani M, Bezuidenhout L, Borry P. Attitudes of research participants and the general public towards genomic data sharing: a systematic literature review. *Expert Rev Mol Diagn* 2014 Nov;14(8):1053-1065. [doi: [10.1586/14737159.2014.961917](https://doi.org/10.1586/14737159.2014.961917)] [Medline: [25260013](https://pubmed.ncbi.nlm.nih.gov/25260013/)]
86. Majumder M, Cook-Deegan R, McGuire AL. Beyond Our Borders? Public Resistance to Global Genomic Data Sharing. *PLoS Biol* 2016 Nov;14(11):e2000206 [[FREE Full text](#)] [doi: [10.1371/journal.pbio.2000206](https://doi.org/10.1371/journal.pbio.2000206)] [Medline: [27806054](https://pubmed.ncbi.nlm.nih.gov/27806054/)]
87. Haeusermann T, Greshake B, Blasimme A, Irdam D, Richards M, Vayena E. Open sharing of genomic data: Who does it and why? *PLoS One* 2017;12(5):e0177158 [[FREE Full text](#)] [doi: [10.1371/journal.pone.0177158](https://doi.org/10.1371/journal.pone.0177158)] [Medline: [28486511](https://pubmed.ncbi.nlm.nih.gov/28486511/)]
88. GSMA. The state of mobile data for social good report. 2017. URL: [http://unglobalpulse.org/sites/default/files/MobileDataforSocialGoodReport\\_29June.pdf](http://unglobalpulse.org/sites/default/files/MobileDataforSocialGoodReport_29June.pdf) [accessed 2018-07-11] [[WebCite Cache ID 70q0zWy1B](#)]
89. Telefonica. Telefonica Digital's Smart Steps. 2018. URL: <http://datacollaboratives.org/cases/telefonica-digitals-smart-steps.html> [accessed 2018-07-11] [[WebCite Cache ID 70q17L2L0](#)]
90. Office for National Statistics. ONS methodology working paper series No. 8. 2017. Statistical uses for mobile phone data: literature review URL: <https://www.ons.gov.uk/methodology/methodologicalpublications/generalmethodology/>

- [onworkingpaperseries/onmethodologyworkingpaperseriesno8statisticalusesformobilephonedataliteraturereview](#) [accessed 2018-07-11] [WebCite Cache ID 70q1KGNpy]
91. Office for National Statistics. Research Outputs: Using mobile phone data to estimate commuting flows. 2017. URL: <https://www.ons.gov.uk/census/censustransformationprogramme/administrativedatacensusproject/administrativedatacensusresearchoutputs/populationcharacteristics/researchoutputsusingmobilephonedatatoestimatecommutingflows> [accessed 2018-07-11] [WebCite Cache ID 70q1PiK06]
92. Farrahi K, Emonet K, Cebrian M. Predicting a Community's Flu Dynamics with Mobile Phone Data. 2015. URL: <https://hal.archives-ouvertes.fr/hal-01146198/document> [accessed 2018-07-11] [WebCite Cache ID 70q1Eikbe]
93. Barchiesi D, Preis T, Bishop S, Moat HS. Modelling human mobility patterns using photographic data shared online. R Soc Open Sci 2015 Aug;2(8):150046 [FREE Full text] [doi: [10.1098/rsos.150046](https://doi.org/10.1098/rsos.150046)] [Medline: [26361545](https://pubmed.ncbi.nlm.nih.gov/26361545/)]
94. Lemey P, Rambaut A, Bedford T, Faria N, Bielejec F, Baele G, et al. Unifying viral genetics and human transportation data to predict the global transmission dynamics of human influenza H3N2. PLoS Pathog 2014 Feb;10(2):e1003932 [FREE Full text] [doi: [10.1371/journal.ppat.1003932](https://doi.org/10.1371/journal.ppat.1003932)] [Medline: [24586153](https://pubmed.ncbi.nlm.nih.gov/24586153/)]
95. Ginger.io. 2018. URL: <http://www.ginger.io/> [WebCite Cache ID 70q1iDLgj]
96. Matheson R. Watch your tone. 2016. URL: <http://news.mit.edu/2016/startup-cogito-voice-analytics-call-centers-ptsd-0120> [WebCite Cache ID 70q1n4TOF]

## Abbreviations

**CDR:** call detail record

**GSMA:** Groupe Spéciale Mobile Association

**LMIC:** low- to middle-income country

**MNO:** mobile network operator

*Edited by C Dias; submitted 29.01.18; peer-reviewed by N Lea, W Dixon, K Yigzaw; comments to author 20.03.18; revised version received 18.04.18; accepted 08.06.18; published 19.07.18*

*Please cite as:*

*Jones KH, Daniels H, Heys S, Ford DV*

*Challenges and Potential Opportunities of Mobile Phone Call Detail Records in Health Research: Review*

*JMIR Mhealth Uhealth 2018;6(7):e161*

*URL: <http://mhealth.jmir.org/2018/7/e161/>*

*doi: [10.2196/mhealth.9974](https://doi.org/10.2196/mhealth.9974)*

*PMID: [30026176](https://pubmed.ncbi.nlm.nih.gov/30026176/)*

©Kerina Helen Jones, Helen Daniels, Sharon Heys, David Vincent Ford. Originally published in JMIR Mhealth and Uhealth (<http://mhealth.jmir.org>), 19.07.2018. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR mhealth and uhealth, is properly cited. The complete bibliographic information, a link to the original publication on <http://mhealth.jmir.org/>, as well as this copyright and license information must be included.