

Critical Limitations of Digital Epidemiology: Why COVID-19 Apps Are Useless

Jonas Klingwort
University of Duisburg-Essen, and
Statistics Netherlands

Rainer Schnell
University of Duisburg-Essen

During the current COVID-19 pandemic, “digital epidemiology” has been proposed to complement traditional reporting and surveillance systems. Instruments such as smartphone contact apps, fitness trackers, and apps for voluntary reporting are intended to be used to monitor or limit the spread of COVID-19. The methodological drawbacks and limitations of these instruments and devices are insufficiently addressed in the public discussion. Therefore, we review these weaknesses and limitations, using the Total Survey Error framework to address sampling and non-sampling errors of these approaches. We argue that no useful results can be obtained by any of the suggested methods of digital epidemiology for COVID-19 research. Finally, we suggest feasible alternative data sources for valid and population covering COVID-19 indicator systems.

Keywords: COVID-19; Total Survey Error; Nonresponse; Coverage; Selection Bias; Non-probability Samples; Big Data; Contact Apps; Smartphones; Fitness Tracker; Health Care Insurance Data; Leopoldina Statement; Digital Public Health; Digital Tracing; Corona-App; SARS-CoV-2

1 Introduction

The German Academy of Natural Scientists (Leopoldina) suggested in their third ad-hoc-statement (13 April 2020) complementing traditional reporting and surveillance systems with methods from digital epidemiology. Digital epidemiology can be defined as the “(...) use of epidemiologic knowledge and digital technologies to enable disease surveillance and epidemiological research” (Porta, 2014). Specifically, Leopoldina (2020) suggested the use of

1. nationwide surveys via a smartphone app to provide data on the population’s current state of health,
2. apps for voluntary reporting of symptoms and information on the course of the illness, and
3. data recorded by activity trackers and other wearables on the wearer’s resting pulse and sleep rhythm to indicate signs of fever and the emergence of flu-like symptoms.

The aim given by Leopoldina for all reporting systems is to predict the probable development of the COVID-19 pandemic over the course of one to two weeks and to compare the expected efficacy of measures prior to their application.

By using this approach, Leopoldina (2020, pp. 6–7) seeks to identify recurrent regional clusters in which infection rates increase locally within a short timeframe and to establish targeted, regional measures for controlling them. Finally, the cited ad-hoc statement mentions the

4. use of voluntarily provided personal data such as movement profiles (GPS data) in combination with contact tracing.

The suggestion of smartphone contact apps received most public attention. Similar systems – without GPS-tracks – are already in use in Australia (CovidSafe) and Singapore (TraceTogether).¹ Many other countries are currently in the process of introducing such COVID-19 apps, but due to privacy concerns few of these apps rely on GPS data, but on Bluetooth signals.²

From a methodological perspective, these suggestions and the use of digital epidemiology, in general, raise several problems. We will discuss statistical problems resulting from selective sampling and doubtful measurement validity of some of the suggested methods for digital epidemiology for use as pandemic surveillance tools.

¹ See www.health.gov.au/resources/apps-and-tools/covidsafe-app and www.tracetgether.gov.sg/.

² GPS data seems to be used by COVID-19 apps only in Bahrain, Iceland, Italy and Norway. For details of the protocols, see https://en.wikipedia.org/wiki/COVID-19_apps, access date 9/5/2020.

Contact: Jonas Klingwort, Research Methodology Group, University of Duisburg-Essen (E-mail: Jonas.Klingwort@uni-due.de)

2 Digital Epidemiology

Digital epidemiology uses “data that was not generated with the primary purpose of doing epidemiology” (Salathé, 2018, p. 2). This implies that the data is not generated according to a statistical research design. Such data has been named “found data” (Harford, 2014). In the absence of a study design and without a known data generating mechanism, no valid inferences can be drawn by design-based approaches. The population coverage is not complete or unknown, so it is impossible to define the population to which inferences could be made.

Proponents of digital epidemiology rarely discuss this central problem. For example, the widely cited papers by Salathé (2018), Salathé et al. (2012) do not address sampling or non-sampling errors such as missing data, coverage of the population, or population inference at all.

Given the inherent problems of data collected without any statistical research design, from a statisticians perspective, Höhle (2017) answered the question “Does the use of emerging technologies and digital tools, especially Big Data, present an epistemic shift in epidemiology?” (Eckmanns & Hempel, 2015) with “No”. In a more recent publication, Zeeb, Pigeot, and Schüz (2020, p. 138) states the fact of a substantial lack of well planned and executed evaluation studies in the fields of “Digital Public Health”.

The most widely known demonstration of potential uses of digital epidemiology was Google Flu Trends, where Google searches were aggregated to predict flu prevalence. However, the results are mixed and – if uncorrected – potentially misleading (Butler, 2013; Cervellin, Comelli, & Lippi, 2017; Lippi & Cervellin, 2019). Furthermore, due to Google’s data ownership, identification of regional infection clusters would be possible only for Google. Using search engines for prevalence estimation is after more than a decade of research, still an unproven and debated method. In sum, the flagship demonstration of digital epidemiology is, in no way, a reliable epidemiological tool.

This absence of positive evidence for the usefulness of digital epidemiology is also characteristic of the suggested use of fitness trackers to monitor the actual development of an illness in a population. Besides selectivity (fitness trackers usually do not aim at older and fragile subgroups), there are measurement problems due to the clear classification of activity and non-activity in the data, the non-specificity of heart rate variations, and the recognition of duplicates due to multiple users of a device. Furthermore, the lack of demographics (due to privacy protection) limits the potential analysis options (Fujibayashi et al., 2018; Hswen, Brownstein, Liu, & Hawkins, 2017; Radin, Wineinger, Topol, & Steinhubl, 2020). Finally, local weather variations have to be controlled if variations in activity are used as health indicators.

3 The Total Survey Error Framework for the Evaluation of Digital Epidemiology

The methods of digital epidemiology proposed by the Leopoldina and others require the population to have access to the electronic devices and motivation to participate. Both conditions must be met in all population subgroups, or at least the relationship between the two conditions and the variable of interest must be known. If neither is given, the suggested technique is not suited for general population studies. Neither smartphones nor fitness trackers are uniformly distributed in a population, nor is the functional relationship between their use and health status known in a way useful for correction. Therefore, population studies based on this kind of devices will have coverage and nonresponse problems (for both types of errors in general, see Biemer and Lyberg (2003)).

3.1 Bias in Non-probability Samples

The model by Bethlehem and Biffignandi (2012, pp. 309–312) allows the estimation of the expected difference between the estimated mean of a non-probability sample \bar{Y}_{ns} from the population mean \bar{Y} . In the case of digital epidemiology data, the non-probability sample is the subgroup of the population owning and using a smartphone or fitness tracker. The difference is given by

$$\bar{Y} - \bar{Y}_{ns} = \frac{R_{\rho Y} S_{\rho} S_Y}{\bar{\rho}}.$$

Since not all persons in the population owning smartphones or fitness trackers will participate in all data collection efforts, it might be assumed that every person has a response propensity ρ . The overall mean of the response propensity is $\bar{\rho}$. The standard deviation of ρ is S_{ρ} . $R_{\rho Y}$ is the correlation between Y and ρ , and S_Y is the standard deviation of Y .

Hence, the difference between the non-probability sample and the population depends on three quantities:

1. the correlation between the response propensity and the variable to be estimated,
2. the variance of the response propensity,
3. and the variance of the variable of interest.

The difference gets smaller if participation rates in the non-probability sample increases or if there is no correlation between response propensity and the variable of interest or the variance of the variable of interest is small. Larger differences are to be expected if the probability of participation depends on the variable of interest, the more the response propensity differ between persons, the larger the differences in persons concerning the variable of interest, and the smaller the overall response propensity.

4 Critical Limitations of Digital Epidemiology for COVID-19 Research

The target population must own and use a smartphone or fitness tracker to be eligible for studies in digital epidemiology. According to an estimate of German official statistics, 81.6% of the German general population uses a smartphone (Statistisches Bundesamt, 2020). Given the fact that the underlying survey is a self-recruited sample, this number is most likely an overestimate. However, the resulting coverage problem is evident when different age groups are considered. Based on a quota sample, Generali Deutschland AG (2017, pp. 116–117) reports 37% of the 65–74 old owning a smartphone and only 17% of the 75–85 years old. In general, elderly with higher socioeconomic status are more likely to own smartphones (47% in the high SES group vs. 14% in the low SES group).³ Since about 21% of the German general population is older than 65 years, disregarding differential undercoverage in high age-groups will cause biased estimates. Finally, a similar selection effect can be observed in the lower age groups: Younger children do not own smartphones. The percentage seems to increase from about 54% at the age of 6–7 to approximately 82% at the age of 11 (Berg, 2020). Given first evidence in Germany of children being as infectious as adults (Jones et al., 2020), this subgroup cannot be disregarded.

The number of people in Germany using a fitness tracker or smartwatch has increased steadily since 2015. However, in 2019, only 29% used fitness trackers and 36% smartwatches (Deloitte and Bitkom, 2020, p. 40). Age-related differences in reported usage are also noticeable here. Based on a web survey, Statista (2019) reports 59% of people older than 60 years are not interested in such devices, and 6% of the respondents own but do not use such a device. Due to the lack of validation studies, empirical evidence on the use of digital health apps is sparse. In general, the usage of such apps seems to be associated with age, health, and SES (for a recent review, see Müller, Wachtler, and Lampert (2020)).

It seems to be useful to visualize the necessary steps in the selection process of digital epidemiology data (see Figure 1). Each of the steps 1–7 might introduce selection bias by excluding specific subgroups of the general population. Step 1 and step 5 exclude non-owners of specific smart devices (a coverage problem). The remaining steps might be similar in their causes and effects to nonresponse in surveys, where step 7 is identical to item-nonresponse. Please note that some disabled or vulnerable subgroups are excluded by design.

The population passing all steps in this selection process will most likely not be a random sample of the target population. Given the evidence on the differential use of health apps discussed above, the necessary selection steps will result in biased non-probability samples. People in these samples are likely to be younger, having a higher socioeconomic status, being more healthy, more physically active, and more inter-

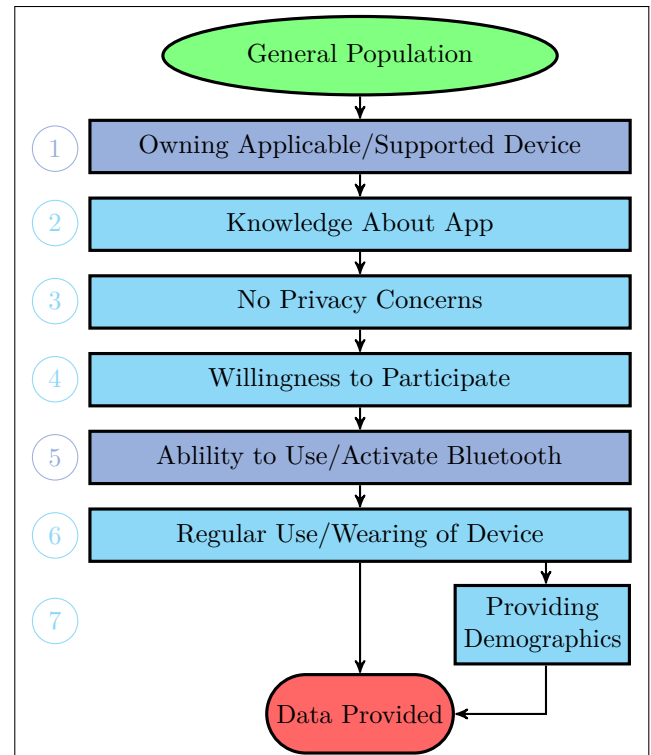


Figure 1. Selection processes in Digital Epidemiology using Smart Devices. Light blue nodes are similar to survey nonresponse, dark blue nodes are similar to undercoverage.

ested in new technologies.

4.1 The RKI Data-donation App

Despite such selection effects, the German federal government agency responsible for disease control and prevention (RKI) seems to consider the identification of regional infection clusters down to postcode levels (about 8,200 in Germany) as possible using data of such apps (Robert Koch-Institut, 2020b). The app used by the RKI (denoted as data-donation app) has been installed on about 509,000 devices (date: 5.5.2020). However, this is about 0.5% of the population. This proportion varies between 0.2%–1.2% (Robert Koch-Institut, 2020a) depending on the administrative unit (at the European NUTS-3 level; about 290 units in Germany). Even if all measurement problems (sensitivity and specificity of using pulse and sleep rhythms to detect infections) and all selection problems would have been solved, the required sample size for each postcode area exceeds 1,853 persons to achieve a power of 0.8 ($\alpha = 0.05$), resulting in an overall sample of more than 15 million persons (Schnell & Smid, 2020), if a change of a prevalence of 1% (from 1%

³Similar results are reported by Keusch, Bähr, Haas, Kreuter, and Trappmann (2020) based on a probability survey.

to 2%) has to be detected. This sample size seems to be unlikely to achieve. In sum, the data-donation app suffers from unproven sensitivity and specificity, sample selection bias, and insufficient statistical power. From a statistical point of view, it is hard to see any epidemiological use of this app.

4.2 COVID-19 Apps for Automatic Contact Tracing

Many different kinds of apps for COVID-19 monitoring have been suggested. The most widely proposed kind of a COVID-19 app uses Bluetooth signals to track encounters with people who are later diagnosed as infected after the encounter.

Not considering any privacy concerns, the accuracy of such automatic contact tracing apps suffers from Bluetooth-based measurement errors. Such errors are due to devices' different signal strengths and the fact that Bluetooth modules are not transmitting the same signal strength in all directions. Furthermore, physical environment features such as windows, walls, or doors may impact the range of detectable devices.

These physical variations will cause false-positive alarms (detection of contacts when no infection risk was present) and false-negative alarms (failing to detect a contact which might have been dangerous) (Schneier, 2020). Depending on the details of the technical implementation, it might be possible that such apps may be misused by people deliberately generating false positive alarms (Soltani, Calo, & Bergstrom, 2020). It is also conceivable that false negatives are deliberately generated because infected people do not want to reveal their actual status.

However, the main methodological problem is the selection process described in section 4. The sample of the population actually using the app will not be a random sample. It seems likely that especially subpopulations with a higher prevalence of undetected infections will have lower coverage rates by the apps: Older people, children, persons without smartphones due to lower-income and members of vulnerable populations. Moreover, to be useful beyond individual cases of preventing infections, a high rate of adaption within a population is required. Privacy concerns relating to such apps and the legal impossibility to force a population to use an app may yield much lower covering rates. As currently observed in Iceland, the app is not of any epidemiological use with a covering rate of 40% in the general population (Johnson, 2020). Overall, the information gained to control the spread of the infections seems to be very limited. Due to the inbuilt privacy mechanisms, the resulting data for scientific research based on these apps are limited to counts of positive or negative encounters of selective populations, where encounter probabilities can not be computed. Therefore, at best, such apps might prevent some individual infections. They are neither a panacea nor an epidemiological research tool.

5 Feasible Alternatives for COVID-19 Population Research in Germany

Heller (2020) suggested using health care insurance data to monitor and analyze the pandemic. In the case of inpatient treatment of patients with COVID-19, health insurances receive data within a few days. Such data contains required covariates, such as age, concomitant diagnoses, procedures, ventilation hours, and survival of inpatient treatment. Statistical analysis of outcome (e.g., duration of mechanical ventilation, the survival of the inpatient stay) depending on accompanying diseases (esp. pneumonia) or previous conditions (COPD, asthma, diabetes, heart failure) conditional on demographic characteristics (e.g., age, gender or occupational codes) are possible with data already available. This approach is fast, no additional data collection is required, and can be analyzed within existing legal regulations.

Instead of mapping data of a COVID-19 app, we suggest mapping data resulting from infections reported to municipal health departments. The number of inhabitants is known for all 100*100m-areas in Germany. Local density estimates can be published using a tool freely available to administrations using the infections reported to the municipal health department or the data of the health insurances (infas 360, 2020). The areas can be aggregated automatically until the number of people in an area meet the requirements of the interpretation given by local data protection officers to the General Data Protection Regulation (GDPR). This way, each municipality can have its own desired level of aggregation. This tool would allow the identification of regional infections clusters.

Finally, we cite the recommendations given by Schnell and Smid (2020): There is no alternative to selecting a large random sample ($n > 30,000$), using local population registers as sampling frame to estimate the proportion of infected persons as well as the proportion of immune persons in the population. Second, a longitudinal sample is required to study the course of the disease and for the study of symptom-free infected persons. Third, a randomly selected post-mortem sample is needed to estimate the proportion of infected persons among the deceased and determine the cause of death. Fourth, a small sample should be randomly selected from the population to describe changes in attitudes and reported behavior due to COVID-19. This survey should not be implemented as a web survey to avoid bias due to health, age, and education. Such a bias is also to be expected in the data obtained with digital epidemiology devices. Bias caused by health issues are most likely not missing at random and therefore cannot be corrected by any weighting procedure (Schnell, Noack, & Torregroza, 2017).

6 Conclusion

Leopoldina (2020) and Robert Koch-Institut (2020b) proposed the use of smartphone apps to monitor the spread of

COVID-19. From a statistical and methodological perspective, the use of smart devices suffers from under-coverage and nonresponse, which are rarely addressed by proponents of digital epidemiology. Furthermore, the sensitivity and specificity of the suggested apps are unknown and undiscussed. For COVID-19 surveillance, we recommend that instead of digital epidemiology, available routine data and random samples should be used.⁴

References

- Berg, A. (2020). *Kinder und Jugendliche in der digitalen Welt*. Berlin: bitkom. Retrieved from https://www.bitkom.org/sites/default/files/2019-05/bitkom_pk-char-ts_kinder_und_jugendliche_2019.pdf
- Bethlehem, J., & Biffignandi, S. (2012). *Handbook of web surveys*. Hoboken: Wiley.
- Biemer, P. P., & Lyberg, L. E. (2003). *Introduction to survey quality*. Hoboken: Wiley.
- Butler, D. (2013). When Google got flu wrong: US outbreak foxes a leading web-based method for tracking seasonal flu. *Nature*, 494, 155–156.
- Cervellin, G., Comelli, I., & Lippi, G. (2017). Is Google trends a reliable tool for digital epidemiology? Insights from different clinical settings. *Journal of Epidemiology and Global Health*, 7(3), 185–189.
- Deloitte and Bitkom. (2020). *Zukunft der Consumer Technology – 2019: Marktentwicklung, Trends, Mediennutzung, Technologien, Geschäftsmodelle*. Berlin: Bitkom e. V.
- Eckmanns, T., & Hempel, L. (2015). DELSI—digital epidemiology and its ethical, legal and social implications. Call for Abstracts by the Robert-Koch-Institute.
- Fujibayashi, K., Takahashi, H., Tanei, M., Uehara, Y., Yokokawa, H., & Naito, T. (2018). A new influenza-tracking smartphone app (flu-report) based on a self-administered questionnaire: Cross-sectional study. *JMIR Mhealth Uhealth*, 6(6), e136.
- Generali Deutschland AG (Ed.). (2017). *Generali Altersstudie 2017: Wie ältere Menschen in Deutschland denken und leben*. Heidelberg: Springer.
- Harford, T. (2014). Big data: A big mistake? *Significance*, 11(5), 14–19.
- Heller, G. (2020). Exposé zur Surveillance von behandlungsbedürftigen COVID-19 Infektionen auf der Basis von (stationären) Abrechnungsdaten. Memo, 13.04.2020.
- Höhle, M. (2017). A statistician's perspective on digital epidemiology. *Life Sciences, Society and Policy*, 13(17), 1–4.
- Hswen, Y., Brownstein, J. S., Liu, J., & Hawkins, J. B. (2017). Use of a digital health application for influenza surveillance in china. *American Journal of Public Health*, 107(7), 1130.
- infas 360. (2020). Wo sind die Neuinfektionen in der Stadt? Retrieved from <https://blog.infas360.de/2020/04/29/georaster-hilft-bei-neuinfektionen/>
- Istituto Nazionale di Statistica. (2020). Indagine sierologica su covid-19 condotta da ministero della salute e istat. Retrieved from <https://www.istat.it/it/archivio/242676>
- Johnson, B. (2020). Nearly 40% of Icelanders are using a Covid app—and it hasn't helped much. Retrieved from <https://www.technologyreview.com/2020/05/11/1001541/iceland-rakning-c19-covid-contact-tracing/>
- Jones, T. C., Mühlemann, B., Veith, T., Zuchowski, M., Hofmann, J., Stein, A., . . . Drosten, C. (2020). An analysis of SARS-CoV-2 viral load by patient age. Manuscript of the Charité Berlin.
- Keusch, F., Bähr, S., Haas, G.-C., Kreuter, F., & Trappmann, M. (2020). Coverage error in data collection combining mobile surveys with passive measurement using apps: Data from a german national survey. *Sociological Methods & Research*, 1–38.
- Khunti, K., Singh, A. K., Pareek, M., & Hanif, W. (2020). Is ethnicity linked to incidence or outcomes of COVID-19? *BMJ*, 369. doi:10.1136/bmj.m1548
- Leopoldina. (2020). 3rd ad-hoc statement: Coronavirus pandemic. Retrieved from <https://www.leopoldina.org/en/publications/detailview/publication/coronavirus-pandemic-sustainable-ways-to-overcome-the-crisis-13-april-2020/>
- Lippi, G., & Cervellin, G. (2019). Is digital epidemiology reliable? Insight from updated cancer statistics. *Annals of Translational Medicine*, 7(1), 15.
- Müller, A. C., Wachtler, B., & Lampert, T. (2020). Digital Divide – Soziale Unterschiede in der Nutzung digitaler Gesundheitsangebote. *Bundesgesundheitsblatt*, 63, 185–191.
- Office for National Statistics. (2020). COVID-19 Infection Survey (CIS). Retrieved from <https://www.ons.gov.uk/surveys/informationforhouseholdsandindividuals/householdandindividualsurveys/covid19infectionsurvey>
- Porta, M. (Ed.). (2014). *A dictionary of epidemiology* (6th ed.). Oxford: Oxford University Press.
- Radin, J. M., Wineinger, N. E., Topol, E. J., & Steinhubl, S. R. (2020). Harnessing wearable device data to improve state-level real-time surveillance of influenza-

⁴Examples are surveys by two European national statistical institutes. In the UK, ONS is conducting the COVID-19 Infection Survey (CIS), based on 20,000 households, to identify the percentage of the population tested positive for COVID-19 and whether they had symptoms or not (Woodhill, 2020; Office for National Statistics, 2020). In Italy, ISTAT is planning to observe 150,000 individuals in Italy using a random sample, to identify the number of people with antibodies even in the absence of symptoms (Istituto Nazionale di Statistica, 2020).

- like illness in the USA: A population-based study. *Lancet Digital Health*, 2(2), e85–e93.
- Robert Koch-Institut. (2020a). Die Spender:innen-Karte – regionale Verteilung. Retrieved from <https://corona-date.nspende.de/science/reports/users/>
- Robert Koch-Institut. (2020b). *Mit Daten von Fitnessarmbändern und Smartwatches mehr über die Verbreitung des Coronavirus erfahren*. Press statement of the Robert Koch-Institute on 7 Apr. 2020.
- Salathé, M. (2018). Digital epidemiology: What is it, and where is it going? *Life Sciences, Society and Policy*, 14(1), 1–5.
- Salathé, M., Bengtsson, L., Bodnar, T. J., Brewer, D. D., Brownstein, J. S., Buckee, C., ... Vespignani, A. (2012). Digital epidemiology. *PLoS Computational Biology*.
- Schneier, B. (2020). Me on COVID-19 contact tracing apps. Retrieved from https://www.schneier.com/blog/archives/2020/05/me_on_covid-19_.html
- Schnell, R., Noack, M., & Torregroza, S. (2017). Differences in general health of internet users and non-users and implications for the use of web surveys. *Survey Research Methods*, 11(2), 105–123.
- Schnell, R., & Smid, M. (2020). Methodological problems and solutions for sampling in epidemiological SARS-CoV-2 research. *Survey Research Methods*, 14(2), 123–129. doi:10.18148/srm/2020.v14i2.7749
- Soltani, A., Calo, R., & Bergstrom, C. (2020). Contact-tracing apps are not a solution to the COVID-19 crisis. Retrieved from <https://www.brookings.edu/techstream/inaccurate-and-insecure-why-contact-tracing-apps-could-be-a-disaster/>
- Statista. (2019). Umfrage zur Nutzung von Smartwatches und Fitness-Trackern am Handgelenk in Deutschland nach Alter und Geschlecht im Jahr 2017. Retrieved from <https://de.statista.com/statistik/daten/studie/454312/umfrage/nutzung-von-smartwatches-und-fitness-trackern-nach-alter-und-geschlecht/>
- Statistisches Bundesamt. (2020). Wirtschaftsrechnungen: Private Haushalte in der Informationsgesellschaft – Nutzung von Informations- und Kommunikationstechnologien.
- Woodhill, N. (2020). *Coronavirus (COVID-19) Infection Survey Pilot: England, 21 May 2020*. London: Office for National Statistics (ONS).
- Zeeb, H., Pigeot, I., & Schüz, B. (2020). Digital Public Health – Ein Überblick. *Bundesgesundheitsblatt*, 63, 137–144.
- not been assessed properly, and, in some cases, they even suffer from a poor statistical quality. In addition, technologies are used in various ways to address important research questions. Certainly, the urgency to deal with the pandemic requires timely statistical analysis to inform policy-makers. However, we strongly believe that statistical quality of all the estimates and analysis should not be neglected.
- We would like to thank the authors of the article titled “Critical Limitations of Digital Epidemiology: Why COVID-19 Apps Are Useless” for their interesting and well-articulated contribution on the problems of Digital Epidemiology in the COVID-19 pandemic. This article generates an important debate that must be considered by policy-makers and in particular by national health systems in the fight of COVID-19.
- The authors criticise the use of devices, such as smartphone contact apps, fitness trackers, and apps for voluntary reporting to monitor or limit the spread of COVID-19. This crucial topic is evaluated from a survey statistics as well as official statistics point of view. In particular, the article adopts the Total Survey Error (TSE) framework to investigate the problems arising from those techniques.
- The article begins with a very good introduction on the recognition of the important issues related to devices in survey statistics and methodology that are studied in the literature. Then, the use of Digital Epidemiology in the context of COVID-19 is discussed. Specifically, the authors treat the issues of missing data, coverage of the target population, selectivity, and more broadly, the possibility to carry out statistical inference. Regarding the use of fitness trackers, the authors also point out another important problem that should be investigated i.e. measurements problems arising from this data collection mechanism.
- Furthermore, the authors discuss in detail the limitations of digital epidemiology with a particular attention to COVID-19 in Germany. First, the problem of coverage is highlighted. The authors focus mainly on age groups and socioeconomic status. We remember here the problem of care homes that have been particularly vulnerable in this pandemic. Moreover, we believe that a further coverage problem might be related to ethnic groups. Information on ethnic minorities would be very helpful for policy-makers, indeed in the UK and the US ethnic minority populations seem to be disproportionately affected by COVID-19 (Khunti, Singh, Pareek, & Hanif, 2020). We want also to stress that ethical issues connected with the use of digital epidemiology might influence the coverage of these tools as well. Some people might be reluctant to share their information for privacy reasons, and some people may voluntarily hide some movements by simply turning off their Bluetooth.
- The authors also discuss some possible alternatives for COVID-19 population research in Germany. Interestingly, they stress important recommendations from survey method-

Commentary

In the past few months, there has been a huge amount of research around COVID-19 pandemic. This research is usually fast and often based on data sources whose quality has

not been assessed properly, and, in some cases, they even suffer from a poor statistical quality. In addition, technologies are used in various ways to address important research questions. Certainly, the urgency to deal with the pandemic requires timely statistical analysis to inform policy-makers. However, we strongly believe that statistical quality of all the estimates and analysis should not be neglected.

ology literature in the COVID-19 data collection problem. The authors mention excellent points regarding the necessity of randomly selected samples. Indeed, some European countries have already started to select random samples representative for the national population. Good examples of these strategies are the Italian Statistical Institute and the Office for National Statistics (UK) which launched some sample surveys based on probabilistic experiments. Moreover, Understanding Society in the UK is conducting an interesting survey on the participants' experience during COVID-19. We are not totally convinced that "a small population survey" could be adequate "to describe changes in attitudes and reported behaviour due to COVID-19", as the authors stated in Section 5. Indeed, this survey would require a complex survey design with representativeness of the sample at sub-national levels. Therefore, the sample should not be small to avoid extremely large variances in the estimates for sub-national areas. Indeed, COVID-19 has an important geographic distribution component in its aspects.

A crucial recommendation highlighted at the end of the article is the following: "This survey should not be implemented as a web survey to avoid bias due to health, age, and education". The literature has widely discussed this issue and we want to stress again that ethnic groups as well as characteristics related to economic well-being may affect web surveys.

The numerous issues arising from data collected via the technologies mentioned in this article may be approached in different ways. The paper correctly discusses Bethlehem and Biffignandi (2012) model, and thus the TSE framework. We agree with the authors that Digital Epidemiology has important limitations in the COVID-19 pandemic analysis. However, can some information collected by those be included and integrated with data coming from probabilistic experiments? For example, the issue of data integration of non-probability samples with probability samples has been studied in the literature. Also, can survey calibration using auxiliary information help in this context? In addition, how can measurement error issues be tackled?

We really hope that this article will be read by policy-makers that diffuse information on COVID-19 every day and especially by those governments that are planning to adopt apps to investigate aspect of COVID-19. To contrast this pandemic, we need organised data collection plans to provide accurate and precise estimates related to the multiple aspects of the phenomena.

Angelo Moretti
Manchester Metropolitan University, U.K.
Catarina Santi
University College Cork, Ireland