Curr Envir Health Rpt (2017) 4:463–471
DOI 10.1007/s40572-017-0163-y

Towards Personal Exposures: How Technology Is Changing Air Pollution and Health Research

A. Larkin [1] · **P. Hystad** [2]

Published online: 5 October 2017

Abstract
Purpose of Review We present a review of emerging technologies and how these can transform personal air pollution exposure assessment and subsequent health research.
Recent Findings Estimating personal air pollution exposures is currently split broadly into methods for modeling exposures for large populations versus measuring exposures for small populations. Air pollution sensors, smartphones, and air pollution models capitalizing on big/new data sources offer tremendous opportunity for unifying these approaches and improving long-term personal exposure prediction at scales needed for population-based research. A multi-disciplinary approach is needed to combine these technologies to not only estimate personal exposures for epidemiological research but also determine drivers of these exposures and new prevention opportunities. While available technologies can revolutionize air pollution exposure research, ethical, privacy, logistical, and data science challenges must be met before widespread implementations occur.
Summary Available technologies and related advances in data science can improve long-term personal air pollution exposure estimates at scales needed for population-based research. This will advance our ability to evaluate the impacts of air pollution on human health and develop effective prevention strategies.

Keywords Air pollution · Sensors · Smartphones · Big data · Exposure assessment · Epidemiology

This article is part of the Topical Collection on *Air Pollution and Health*

✉ P. Hystad
Perry.Hystad@oregonstate.edu

A. Larkin
Andrew.Larkin@oregonstate.edu

[1] College of Public Health and Human Sciences, Oregon State University, Milam 20A, Corvallis, OR 97331, USA

[2] College of Public Health and Human Sciences, Oregon State University, Milam 20C, Corvallis, OR 97331, USA

Introduction

The human health burden from air pollution is extremely large. In 2015, long-term exposure to ambient fine particle matter air pollution ($PM_{2.5}$) was associated with 4.2 million deaths and 103.1 million years of healthy life lost (representing 7.6% of global mortality) [1]. While our understanding of the complex relationships between air pollution and human health has markedly improved over the last several decades, knowledge gaps and consequent uncertainties remain that limit our ability to mitigate the adverse impacts of air pollution.

Exposure assessment is one, if not the greatest, challenge to further understanding and reducing air pollution health impacts. Estimating *personal* air pollution exposures for large populations remains an elusive goal, but central to determining health impacts, evaluating exposure sources and pathways, detecting susceptible populations, and identifying intervention opportunities. Recently, the concept of the exposome "the totality of environmental exposures from conception onwards" [2] has catalyzed exposure scientists to develop new methods to assess a range of personal exposures, using both internal biomarkers and external exposure measures. While there is rapidly growing potential for internal biomarkers of environmental exposures [3], these remain limited for the air pollutants of greatest concern. Exposure scientists must therefore leverage new data sources, methods, and technologies to better assess *external* personal air pollution exposures.

The current state of air pollution exposure science can be split broadly into methods for modeling exposures for large populations versus measuring exposures for small populations. These approaches are not mutually exclusive, as individual measurements are used to build and evaluate models, but they differ in their respective study designs, applications, strengths, and limitations. Geographic information systems (GIS), deterministic models (e.g., AIRMOD, RLINE, SHEDS), and remotely sensed data have been the foundation of most air pollution modeling efforts to date [4], leveraging spatiotemporal estimates of ambient air pollution concentrations to derive exposure estimates from residential locations, typically for large populations. Personal measures of air pollution concentrations have been restricted to small sample sizes and short durations of time [5•], due primarily to sensor limitations and the logistical and cost constraints of sampling large numbers of individuals for long periods of time. Both paradigms (modeling and measuring) are rapidly evolving with new technologies and large-scale data analytics to provide new opportunities for personal exposure assessment methods.

Here we present a high-level review of available technologies and related advances in data science, and how together they can transform air pollution exposure assessment and health research. Several reviews exist focusing on air pollution exposure assessment for epidemiological studies [4, 6, 7], current methods for assessing the exposome [5•, 8•], and personal sensor technologies [5•, 9•, 10–12]. We therefore focus our review broadly on technological advances that fall under air pollution sensors (for personal measurements), smartphones (mHealth and GPS applications), and air pollution models using "big data" (i.e., large volumes of poly-dimensional data collected from traditional and novel data sources). We provide an example of an ongoing study (PURE-Air) that is attempting to combine new technologies and methods to examine air pollution impacts on cardiopulmonary disease in a global cohort. Together, these technologies will push the boundary of what is feasible in personal air pollution exposure science, epidemiology, and prevention. Ultimately, this review will provide some guidance for how the field can move forward to capitalize on these exciting opportunities.

Framework

Figure 1 illustrates the conceptual framework of this review and how available technologies and related data science can contribute to long-term, large-scale, personal exposure assessments. Technologies covered here belong to one or more of the following three domains: air pollution sensors, smartphone applications, and air pollution models. All three domains provide relevant and unique information to personal air pollution exposure, and interact to provide novel exposure information that cannot be derived from each domain alone. For example, air pollution models alone cannot capture personal exposures without detailed time-activity pattern information and, even then, models applied to smartphone-based GPS location data remain models (and not measures) of air pollution concentrations. Personal measurements are also unlikely to capture long-term (i.e., years to decades) exposures and therefore require integration with air pollution models and time-activity patterns. In addition, "contextual characteristics" (i.e., economic, social, environmental, cultural, institutional, and political attributes) influence all three domains and the relationship between air pollution exposures, health effects, and prevention opportunities. Increasing the accuracy of comprehensive long-term personal exposure estimates for large populations will therefore depend on the integration of these domains through cross-disciplinary collaboration. We provide an example from the ongoing PURE-Air study that is attempting to integrate these technologies and methods to better estimate air pollution exposures and health effects in a global cohort study. We also summarize challenges and opportunities that are presented by these technologies.

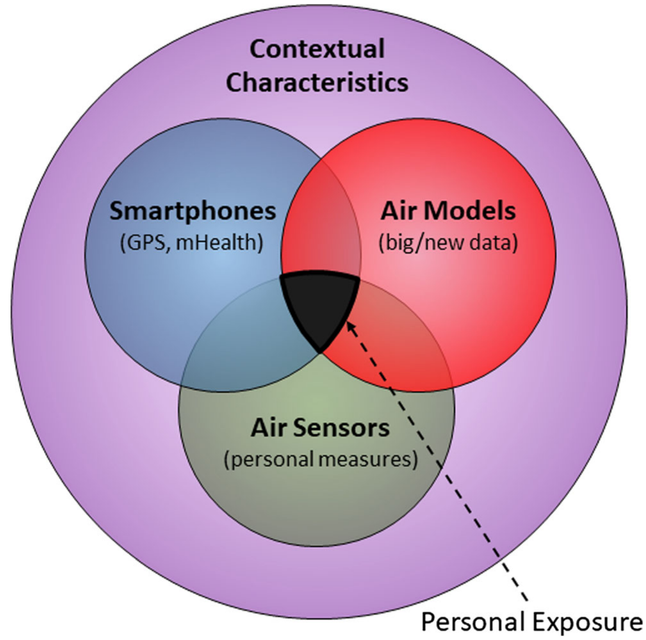

Fig. 1 Conceptual framing of technologies and related advances in data science and how together these can improve long-term personal air pollution exposure estimates at scales needed for population-based research

Air Pollution Sensors

The gold standard of air pollution exposure assessments is personal measurements (with high-quality validated instruments ideally over long periods of time). To date, collecting such measures for large populations remains a major challenge due to cost and logistical constraints. As such, most personal measurement studies have included relatively small sample sizes and/or short time periods [13–16]. For example, in a birth cohort in Sao Paulo, Brazil, NO_2 and O_3 were measured for 366 pregnant women in each trimester for 7–18 days using passive personal samplers [17], and in Barcelona, 122 adults were monitored over a 3-week period for black carbon using the micro-aethalometer AE51 [15]. These studies, while large for today's standards, are a long way from our target of personal measurements for large study populations (i.e., sufficient sample sizes needed to capture the chronic impacts of air pollution) and time periods relevant to the disease process being studied (i.e., years to decades for most chronic diseases).

While the paradigm of air pollution monitoring is changing—from large regulatory fixed-site stations to smaller mobile sensors [11]—there are still few inexpensive and accessible sensors that can measure personal air pollution concentrations accurately [10]. Wearable sensors are being developing rapidly as start-up companies attempt to produce inexpensive sensors, which cost a fraction of traditional scientific monitors. Available inexpensive sensors have been summarized elsewhere [9•], and while most are not yet suitable as wearable monitors, some are in development. Examples of personal air sensors include the TZOA (http://www.tzoa.com) particle monitor (the consumer version costing ~ US$140); AIRBEAM (http://aircasting.org) $PM_{2.5}$ monitor (~ US$250); Flow (https://plumelabs.com/en/products/flow) $PM_{2.5}$, NO_x, O_3, and VOCs monitor (price tbd); ATMOtube (https://atmotube.com/) VOCs and CO monitor (price tbd); and the CleanSpace Tag (https://store.clean.space/) CO monitor (~ US$55). These types of monitors can be easily worn and connect to a smartphone via Bluetooth to stream data online as well as provide warnings based on concentration levels.

Given the rapid pace of low-cost air pollution sensor development, there is an immediate need to ensure the accuracy of new sensors. A current search for "air pollution monitor" in crowdsourcing websites such as Indiegogo, Kickstarter, or GoFundMe reveals hundreds of new air pollution monitors under development. The issue of un-validated air pollution sensors has been highlighted in other commentaries [9•, 18, 19] and toolkits proposed for evaluating new monitors [20]. Determining the capabilities of new sensors to accurately capture pollutant concentrations is essential for ensuring individual measurements are valid and can be used for scientific research. The Air Quality Sensor Performance Evaluation Center (AQ-SPEC) (http://www.aqmd.gov/aq-spec) was created for this purpose and to inform the public about monitor performance. They have tested 19 "low"-cost particle monitors and 10 gaseous sensors against federal reference standards and report extremely variable correspondence (with R^2 values ranging from 0 to 0.99). The EPA has also developed an Air Sensor Toolbox for Citizen Scientists (https://www.epa.gov/air-sensor-toolbox) to provide information on how to select and use low-cost portable air sensors.

Once sensors are validated, and the price and ease of use are reduced, they are likely to be used widely by individuals outside of research studies, supplementing the existing measurements available in the quantified self-movement and increasing citizen science air pollution monitoring activities [9•, 21]. For example, CitiSense is a participatory air quality monitoring project that is developing a sensor-based citizens' observatory in several cities across Europe [22]. The CleanSpace Community (https://our.clean.space/cleanspace-movement/) is another initiative that leverages a smartphone app to view local air quality data, interfaced with the CleanSpace Tag air monitor, and offers "CleanMiles" for making changes to travel behaviors. iSPEX is another citizen science measurement strategy that uses a low-cost optical attachment for smartphones to measure aerosol optical thickness and contribute these measurements through an app to create fine-scale spatial and temporal maps [23]. These studies suggest how the collection of personal air pollution measures might eventually be used for large epidemiological analyses once high-quality personal air pollution sensors reach the consumer market.

Smartphones

Smartphones will allow for personal air pollution exposure assessments at scales needed for population-based research by facilitating personal air pollution sensors and GPS time-activity collection as well as providing a platform for new types of air pollution health studies. There are currently 3.8 billion global smartphone users, projected to nearly double to 6.8 billion users by 2022 [24]. Seventy-seven percent of adult smartphone users in the USA have downloaded an app, with 29% downloading an app that tracks or manages health [25]. The quantified-self movement [26], where individuals use sensors to measure and improve their own health and behavior, has also dramatically increased. Smartphones and mHealth (defined as all mobile health technologies that can contribute to health research, including smartphones, monitors (e.g., Fitbit), electronic health records, etc.) are becoming commonplace in all aspects of health research, offering numerous opportunities for advancing air pollution exposure assessment and epidemiology.

The most direct application of smartphones to enhancing air pollution exposure estimates (beyond facilitating personal exposure measures) is the collection of time-activity patterns using GPS. Most users (71%) continuously carry and sleep within arm's reach of their smartphones [27]. The collection and application of GPS data and time-activity patterns for air pollution exposure prediction have been documented extensively elsewhere [28]. The key distinction to stress here is that the prevalence of smartphones in the general population (and growing acceptance of health and research apps) allows for collection of time-activity patterns on potentially hundreds of thousands of individuals for long periods of time (i.e., months to years). Glasgow et al. [29] demonstrated this utility by collecting GPS locations every 5 min for 3 months for 42 participants using the smartphone application “Apolus (Air, Pollution, Exposure).” While there are challenges to cleaning and analyzing the volume of data collected by GPS [30], Gonzalez et al. [31] examined 100,000 anonymized mobile phone users tracked for 6 months and observed a high degree of temporal and spatial regularity in time-activity patterns. This suggests that continuous GPS monitoring may not be required to assess long-term activity patterns in health studies and that, for example, seasonal measurements of a week in duration may capture much of the time-activity variation important for air pollution exposures.

Smartphones can also serve as the primary platform for new air pollution health studies, including recruiting participants, obtaining electronic consent, collecting survey and biometric data, assessing outcomes, and transmitting data for linkage to other databases, such as medical health records. As an example, one of the key components of the NIH precision medicine cohort initiative [32], which aims to recruit one million participants, is a patient technology systems center, which “…taps into converging trends of increased connectivity, through social media and mobile devices, and Americans' growing desire to be active partners in medical research” ([32], p. 793). New open-source platforms for creating smartphone apps and mHealth applications are being developed to ease access to these technologies. For example, the open-source Apple ResearchKit (https://www.apple.com/researchkit/) allows individual researchers to create a research app, use code from previous apps, and leverage the awareness and reputation of the Apple ResearchKit community for recruiting participants.

In the air pollution field, the Asthma Mobile Health Study (AMHS) (http://apps.icahn.mssm.edu/asthma/) was created using this platform to examine asthma triggers (including local air pollution concentrations) and treatment. The AMHS app was downloaded nearly 50,000 times in the 6 months after its launch. The study was able to demonstrate the utility of conducting research entirely through a smartphone app, successfully linking asthma symptoms to changes in heat, pollen, and air pollution, including the 2015 wildfires in Washington State [33]. However, this study also documented several challenges that can inform future air pollution studies using smartphones for epidemiology. These included selection bias, retention, reporting bias, and privacy concerns. Of the 50,000 downloads of the AMHS app, 8524 individuals completed the consent process and only 2317 individuals were classified as robust users [33]. Not surprisingly, these individuals tended to be younger, whiter, wealthier, and more educated when compared to the CDC asthma registry [33]. Based on their experience with the AMHS app, Chan et al. [33] concluded that studies conducted entirely through smartphone applications are best suited for studies requiring rapid enrollment, pose minimal risks, examine hypothesis with short time frames, require frequent data collecting, use passive data collection (e.g., GPS), do not seek representative samples, and use an analysis plan that accounts for attrition and missing data. Several of these fit well within an air pollution context, while several others do not.

Air Pollution Models

It is unlikely that personal air measurements and individual GPS data will be collected continuously over the time periods needed to capture chronic (i.e., decade-long) air pollution exposures. As a result, environmental models of air pollution concentrations are needed to predict long-term exposures. The current modeling approaches for air pollution exposure assessment have been covered thoroughly in other reviews [4, 6, 7]. Briefly, one of the greatest strengths of the air pollution modeling domain is the ability to leverage multiple sources of data, and with the advent of “big data,” there are numerous opportunities to advance air pollution modeling.

One particular data source that is changing rapidly with new technological developments is remote-sensed air pollution data. The availability and resolution of remotely sensed data have grown exponentially in the last decade and have expanded the geographic coverage of many spatial models by providing estimates of air pollution where there have previously been no or very sparse ground-level data. Satellite-based estimates of $PM_{2.5}$ that have been calibrated to ground-based monitored data are now available for every location on earth at a ~ 1 × 1 km resolution [34]. These satellite-derived measures of air pollution can also be combined with detailed land use characteristics, such as emissions sources (e.g., roads, population density, land use), to model fine-scale spatiotemporal air pollution patterns. For example, we developed a global model of NO_2 concentrations at a 100 m × 100 m resolution (using satellite estimates and land use variables) that predicts 54% of the NO_2 variation from 5220 air monitors in 58 countries [35]. Satellite-based air pollution exposure estimates are likely to continue to improve for the foreseeable future, as new technologies are increasing the spatial and temporal resolution of new satellites, including the European Space Agency Sentinel-5 and Sentinel-5P, scheduled to launch in September 2017.

With the availability of big data sources, such as satellite air pollution estimates, new data integration and modeling methods are needed. Machine learning is one such method that is being used, for example, to combine remote-sensed data, meteorology, and ground-based observations to predict daily $PM_{2.5}$ from 1997 to 2015 globally [36]. Such nonlinear and nonparametric modeling approaches present numerous advantages over traditional linear regression-based methods for resolving the spatial and temporal variability of air pollution concentrations. Deep learning approaches are also being developed for air pollution predictions [37], and applications of deep learning to high-resolution satellite imagery, combined with other ground-based images, will likely enhance our ability to predict air pollution [38] and ease the ongoing refinements of these predictions. Similar approaches have already been developed to predict poverty from satellite imagery [39]. Ultimately, new data mining techniques are providing newer, larger, more varied, and more highly resolved datasets of ambient air pollution, as well as the characteristics that predict these exposures, to inform the advanced modeling of spatiotemporal air pollution concentrations.

The amount of ground-level air monitoring data available to calibrate ambient air pollution exposure models is also expanding rapidly. For example, hourly air quality index data can now be viewed from regulatory monitoring data from 9000 stations in 800 major cities from 70 countries (http://aqicn.org/map/world/). Nontraditional measurement sources are also contributing new measurement data that can improve the accuracy of air pollution models. Citizen science initiatives like the CitiSense, CleanSpace Community, and the iSPEX monitoring initiative summarized above are examples of community-sourced air quality data. Air monitors have also been attached to multiple mobile agents in the environment, including Google Street View cars [40]. These measurements provide information on local sources that are often missed by regulatory monitoring stations, which tend to capture regional population exposures. In the future, these types of community-sourced and mobile measurements can further reduce error in spatiotemporal models by targeting measurement collection to high-priority space-time locations, increasing representation during model building and contributing to more robust model evaluations.

Contextual Characteristics

Contextual characteristics (i.e., economic, social, environmental, cultural, institutional, and political attributes of a place) are essential to consider when evaluating the relationship between air pollution and health, as well as for translating research into policy and prevention. Technological changes have transformed our ability to look upstream at the contextual conditions that influence individual behaviors, air pollution concentrations and exposures, health impacts (and health disparities), and prevention opportunities.

Smart cities/communities can be viewed as the contextual equivalent of the quantified self-concept [41•], where cities use sensors and big data to quantify community characteristics, many of which are important to air pollution [42, 43]. New data streams include those from connected infrastructure, autonomous vehicles, street view imagery, citizen science monitoring networks, and cellular data, to name a few. For example, geo-referenced Google Street View imagery analyzed with machine learning or deep learning algorithms can be used to derive a wealth of contextual characteristics important to air pollution exposures, such as vehicle congestion, vehicle fleet mix, street canyons, street vegetation buffers, pedestrian traffic, and other important modifiers of air pollution exposures. Cellular network data can also be used to quantify population time-activity patterns and population mobility to improve air pollution exposure estimates [44, 45]. In addition, the amount of data now available for cities provides opportunities to evaluate multiple environmental and social exposures together (rather than in isolation). Quantifying exposures that are spatially correlated with air pollution (e.g., noise, green space, poverty, exercise) is an important step towards understanding cumulative exposures and new prevention opportunities.

The Internet and social media are additional data streams that can further capture the complex social and political contextual characteristics of communities that may influence air pollution exposure and control. For example, from January to March 2017, we have collected more than 15 million tweets (text messages posted on the social media platform Twitter) related to air pollution in more than 30 languages. Tweets can be linked to air quality concentrations from regulatory monitors to examine personal views and sentiment about air quality, self-described change in physical and mental state, and changes in behaviors attributed to air quality conditions. More than one third of these tweets also contained images that can be evaluated for pollution-related characteristics as described for Google Street View imagery above. Although not widely utilized in air pollution research, the Internet and social media have been successfully used to capture context in other health studies [58, 59]. Such measures of societal context are rarely included in air pollution research, despite air pollution risk awareness, regulations, air pollution forecasts/notifications, and support for clean air and pollution mitigation directly [46] or indirectly [47, 48] influencing the physical and psychological impacts of air pollution on human health.

Case Example: the PURE-Air Study

An example of integrating the three domains highlighted above (smartphones, sensors, and air models) to estimate personal air pollution exposures is occurring in the ongoing Prospective Urban and Rural Epidemiology (PURE) study (http://health.oregonstate.edu/labs/spatial-health/research/pure-air). The PURE cohort includes ~ 225,000 adults aged 35–70 years at recruitment living in 850 communities in 25 countries. The PURE-Air study is funded to examine the associations between air pollution and cardiopulmonary disease.

To measure household and personal exposures, we are using a new filter-based $PM_{2.5}$ monitor, the ultrasonic personal air sampler (UPAS) [49]. Forty-eight-hour measurements are being collected for 4000 households and personal $PM_{2.5}$ for 1200 individuals living in 10 countries with over 10% of PURE households using solid fuel use for cooking. The UPAS is small, easy to use, and relatively inexpensive compared to existing monitors, allowing for large numbers of monitors to be shipped to existing field teams for data collection. A smartphone interface allows field staff to program monitors, download run-log and GPS data, and automatically send these data to our secure servers where they are checked for errors. Individuals also wear a passive silicone wristband sampler that measures exposure to 1200 organic chemicals [50]. Even this initiative, however, is designed primarily for building exposure models for the entire cohort of ~ 225,000 individuals to examine cardiovascular events and mortality, although direct analysis of measured $PM_{2.5}$ will be conducted for blood pressure and lung function.

Figure 2 illustrates the PURE study communities and the global distribution of $PM_{2.5}$ (estimated from satellite data fused with ground-based monitoring data [34]) and NO_2 (estimated from a global LUR model [35]). The $PM_{2.5}$ and NO_2 concentrations in Beijing and 48-h GPS data (collected from the UPAS air pollution monitor) for one participant are shown. Three other individuals' time-activity spaces are also illustrated. All of these time-activity patterns are from the same PURE study community, highlighting the variability in time-activity patterns (and resulting exposure differences) for individuals living in the same community. Household and personal $PM_{2.5}$ measurements and GPS data will be integrated with models of ambient $PM_{2.5}$ and NO_2 concentrations and with baseline and follow-up questionnaire data on household characteristics and fuel cooking types to predict long-term air pollution exposures. The geographic scale, sample size, and types of measurements being collected in the PURE-Air study are only possible through the rapid technological changes occurring in our field and would not have been possible even 5 years ago. Nevertheless, this study is only a step in the right direction towards what we have proposed here—that technology can facilitate long-term personal air pollution exposure estimates at scales needed for population-based research.

Challenges and Opportunities

The greatest barrier to estimating long-term personal air pollution exposures remains data science approaches for dealing with large, dynamic, multi-level data. Technological advances are driving the big data revolution as well as the accompanying advancements in data sciences that are needed to process these data. Estimating personal air pollution exposures through the integration of smartphone and air monitoring data streams, modeling of air pollution concentrations, and characterization of the contextual characteristics relevant to air pollution and health research will require complex and powerful data processing approaches. Recent advances in general-purpose computing on graphics processing units (GPGPU) hardware and software have dramatically reduced the time required for terabyte-level data processing. In addition, online platforms are reducing the barriers to accessing and processing huge amounts of data, such as the Google Earth Engine, which provides a platform for petabyte-scale analysis of global satellite data (https://earthengine.google.com/).

The fields of health care and personalized medicine were early adopters of large-scale data analytics [51]. It is very likely that personal exposure assessment and environmental epidemiology will similarly benefit from large-scale data science developments. We can reduce barriers to new data science approaches by sharing data processing scripts as open-source code in creative common communities such as GitHub (see the authors' GitHub page at https://github.com/larkinandy/LUR-NO2-Model, for example, code from the global NO_2 model development), which will help identify, evaluate, reproduce, and validate successful methods. The rapid adoption of open-source code sharing by the data science community will increase transparency and reproducibility [52], reducing barriers to more complex (and integrated) exposure assessment approaches.

The technological advances we highlighted here also have important implications for precision medicine/health. The convergence of epidemiology and personal medicine/health is occurring rapidly [53], but how air pollution exposure assessment (and environmental exposures more generally) will fit into this equation has not been adequately explored. Most precision medicine/health initiatives do not include environmental components beyond common biomarkers (e.g., lead), but delivery of these initiatives through apps makes GPS data collection possible (to inform modeled air pollution exposure estimates), as well as the inclusion of low-cost air pollution sensors. Clearly, however, robust methods (e.g., air pollution sensors, exposure algorithms) will be needed to ensure reliable information is used to characterize and communicate the exposures to individuals.

Within both mHealth and precision medicine/health paradigms, individuals as passive participants in environment health studies (having information drawn from

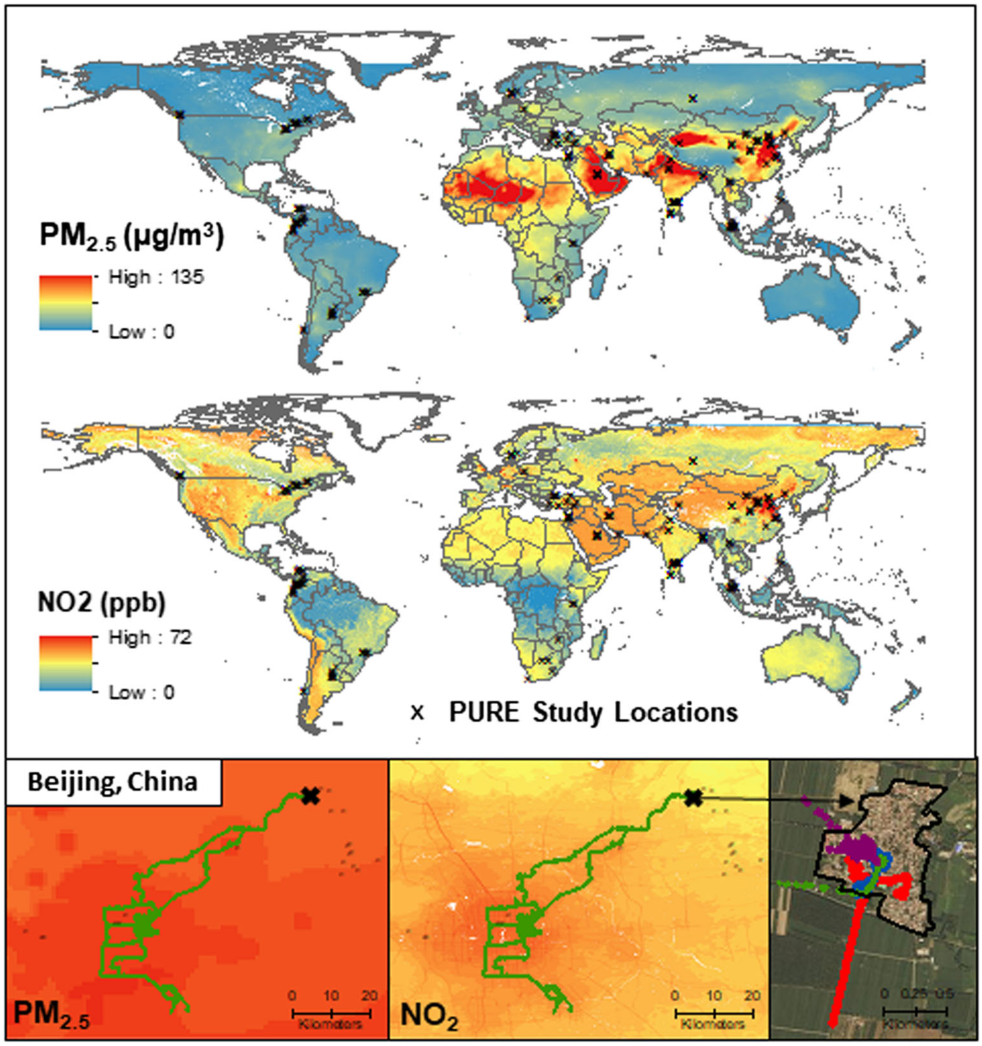

Fig. 2 Global maps of $PM_{2.5}$ (estimated from satellite data fused with ground-based monitoring data [34]) and NO_2 (estimated from a global LUR model [32]) concentrations and the location of PURE study community. Beijing is highlighted with an example of a 48-h GPS time-activity pattern from one PURE participant. Time-activity patterns for three additional PURE participants are shown (all from the same PURE community), highlighting potential differences in exposures for individuals based on mobility. Personal $PM_{2.5}$ exposures were measured with the UPAS air pollution monitor, and individuals wore a passive silicone wristband sampler to measure exposure to organic chemicals. A total of 4000 households and 1200 individuals living in 10 countries will participate in air monitoring

them) will change to individuals being active suppliers of information, choosing how and with whom to share their data. This will require providing valuable (and actionable) knowledge back to participants to ensure they stay engaged and interested in a research study. This will demand more personal environmental exposure assessments that can inform risk communication and behavioral changes to reduce exposures. Providing participants with their air pollution exposures, comparisons to their cohort or local community, and advice to reduce exposures will therefore be an important component to any mHealth or precision environmental health-based initiative. This will present methodological challenges for etiology research, given that the research study itself may promote behavioral changes and air pollution reductions. This, however, offers exciting opportunities to evaluate prevention opportunities, which to date have been limited primarily to air quality advisories.

Ethical and privacy issues remain major concerns that demand thorough examination. Several reviews and commentaries address these issues in detail [54–57]. In terms of the domains reviewed here, there are specific ethical and privacy concerns that should be highlighted. Concerns surrounding GPS data collection, storage, and analyses present obvious privacy issues that need to be addressed. Equity issues surrounding smartphone, personal sensor, and personalized health availability will also present ethical issues, considering that low socio-economic groups experience the largest burden from air pollution, but have the least resources to capitalize on these new technologies. This will also present a major issue in the generalizability of results derived from these types of technologies and under certain circumstances could even jeopardize the findings of the study due to selection bias. As air pollution exposure science moves further towards personal measures, differentiating between public versus commercial usage of the data will become imperative, as such differences have important ethical implications. Finally, there are concerns that personalized exposure science may distract from population health approaches to reducing air pollution health impacts.

Conclusions

Estimating personal air pollution exposures is currently split broadly into methods for modeling exposures for large populations versus measuring exposures for small populations. Air pollution sensors, smartphones, and modeling air pollution concentrations using big/new data offer tremendous opportunities for unifying these approaches and improving long-term personal air pollution exposure prediction at scales needed for population-based research. A multi-disciplinary approach is needed to not only estimate personal exposures for epidemiological research but also determine drivers of these exposures and new prevention opportunities. While available technologies can revolutionize air pollution research, ethical, privacy, logistical, and data science challenges need to be met before widespread applications occur.

Acknowledgments Research reported in this publication was supported by the Office of the Director, National Institutes of Health, under Award Number DP5OD019850. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health. We would also like to acknowledge the reviewers' contributions to refining the messages presented in this article.

Compliance With Ethical Standards

Conflict of Interest The authors declare that they have no conflict of interest.

Human and Animal Rights and Informed Consent All epidemiological studies cited by the authors were in accordance with the ethics standards of Oregon State University.

References

Papers of particular interest, published recently, have been highlighted as
• Of importance

1. Cohen AJ, Brauer M, Burnett R, Anderson HR, Frostad J, Estep K, et al. Estimates and 25-year trends of the global burden of disease attributable to ambient air pollution: an analysis of data from the Global Burden of Diseases Study 2015. Lancet. 2017;389:1907–18.
2. Wild CP. Complementing the genome with an "exposome": the outstanding challenge of environmental exposure measurement in molecular epidemiology. Cancer Epidemiol Prev Biomark. 2005;14:1847–50.
3. Neveu V, Moussy A, Rouaix H, Wedekind R, Pon A, Knox C, et al. Exposome-explorer: a manually-curated database on biomarkers of exposure to dietary and environmental factors. Nucleic Acids Res. 2017;45:D979–84.
4. Özkaynak H, Baxter LK, Dionisio KL, Burke J. Air pollution exposure prediction approaches used in air pollution epidemiology studies. J Expo Sci Environ Epidemiol. 2013;23:566–72.
5.• Nieuwenhuijsen MJ, Donaire-Gonzalez D, Foraster M, Martinez D, Cisneros A. Using personal sensors to assess the exposome and acute health effects. Int J Environ Res Public Health. 2014;11:7805–19. **This study provides a thorough review of how personal sensors can measure multiple exposures and acute health effects**.
6. Jerrett M, Arain A, Kanaroglou P, Beckerman B, Potoglou D, Sahsuvaroglu T, et al. A review and evaluation of intraurban air pollution exposure models. J Expo Anal Environ Epidemiol. 2005;15:185–204.
7. Zou B, Gaines Wilson J, Benjamin Zhan F, Zeng Y. Air pollution exposure assessment methods utilized in epidemiological studies. J Environ Monit. 2009;11:475–90.
8.• Turner MC, Nieuwenhuijsen M, Anderson K, Balshaw D, Cui Y, Dunton G, et al. Assessing the exposome with external measures: commentary on the state of the science and research recommendations. Annu Rev Public Health. 2017;38:215–39. **This commentary highlights how the external exposome can be quantified using modeling and measurement methods**.
9.• Thompson JE. Crowd-sourced air quality studies: a review of the literature & portable sensors. Trends Environ Anal Chem. 2016;11:23–34. **This study provides a review of the relatively low-cost air pollution sensors that are currently available for air pollution research as well as citizen science initiatives**.
10. Wang A, Brauer M. Review of next generation air monitors for air pollution. 2014 [cited 2017 Apr 30]; Available from: https://open.library.ubc.ca/cIRcle/collections/facultyresearchandpublications/52383/items/1.0132725
11. Snyder EG, Watkins TH, Solomon PA, Thoma ED, Williams RW, Hagler GSW, et al. The changing paradigm of air pollution monitoring. Environ Sci Technol. 2013;47:11369–77.
12. Loh M, Sarigiannis D, Gotti A, Karakitsios S, Pronk A, Kuijpers E, et al. How sensors might help define the external exposome. Int J Environ Res Public Health. 2017;14:434.
13. Harrison R, Thornton C, Lawrence R, Mark D, Kinnersley R, Ayres J. Personal exposure monitoring of particulate matter, nitrogen dioxide, and carbon monoxide, including susceptible groups. Occup Environ Med. 2002;59:671–9.
14. Steinle S, Reis S, Sabel CE, Semple S, Twigg MM, Braban CF, et al. Personal exposure monitoring of PM2.5 in indoor and outdoor microenvironments. Sci Total Environ. 2015;508:383–94.
15. Nieuwenhuijsen MJ, Donaire-Gonzalez D, Rivas I, de Castro M, Cirach M, Hoek G, et al. Variability in and agreement between modeled and personal continuously measured black carbon levels using novel smartphone and sensor technologies. Environ Sci Technol. 2015;49:2977–82.
16. Nethery E, Leckie SE, Teschke K, Brauer M. From measures to models: an evaluation of air pollution exposure assessment for epidemiological studies of pregnant women. Occup Environ Med. 2008;65:579–86.
17. Carvalho MA, Bernardes LS, Hettfleisch K, Pastro LDM, Vieira SE, Saldiva SRDM, et al. Associations of maternal personal exposure to air pollution on fetal weight and fetoplacental Doppler: a prospective cohort study. Reprod Toxicol. 2016;62:9–17.
18. Lewis A, Edwards P. Validate personal air-pollution sensors. Nat News. 2016;535:29.
19. Castell N, Dauge FR, Schneider P, Vogt M, Lerner U, Fishbain B, et al. Can commercial low-cost sensor platforms contribute to air quality monitoring and exposure estimates? Environ Int. 2017;99:293–302.
20. Fishbain B, Lerner U, Castell N, Cole-Hunter T, Popoola O, Broday DM, et al. An evaluation tool kit of air quality micro-sensing units. Sci Total Environ. 2017;575:639–48.
21. Reis S, Cowie H, Riddell K, Semple S, Steinle S, Apsley A, et al. Urban air quality citizen science. Phase 1: review of methods and projects. 2013 [cited 2017 Apr 30]. Available from: http://www.environment.scotland.gov.uk/media/68215/Urban-air-quality-citizen-science-Phase-1.pdf

22. Nikzad N, Verma N, Ziftci C, Bales E, Quick N, Zappi P, et al. CitiSense: improving geospatial environmental assessment of air quality using a wireless personal exposure monitoring system. Proc Conf Wirel Health New York, NY, USA. 2012 [cited 2017 Apr 30]. p. 11:1–11:8. Available from: https://doi.org/10.1145/2448096.2448107
23. Snik F, Rietjens JHH, Apituley A, Volten H, Mijling B, Di Noia A, et al. Mapping atmospheric aerosols with a citizen science network of smartphone spectropolarimeters. Geophys Res Lett. 2014;41: 2014GL061462.
24. Ericsson. 5G subscriptions to reach half a billion in 2022: Ericsson Mobility Report [Internet]. Ericsson.com. 2016 [cited 2017 Apr 28]. Available from: https://www.ericsson.com/news/2056743
25. Purcell K. Half of adult cell phone owners have apps on their phones [Internet]. Pew Res Cent Internet Sci Tech. 2011 [cited 2017 Apr 28]. Available from: http://www.pewinternet.org/2011/11/02/half-of-adult-cell-phone-owners-have-apps-on-their-phones/
26. Swan M. Health 2050: the realization of personalized medicine through crowdsourcing, the quantified self, and the participatory biocitizen. J Pers Med. 2012;2:93–118.
27. Bank of America. Trends in consumer mobility report [Internet]. 2015. Available from: http://newsroom.bankofamerica.com/files/doc_library/additional/2015_BAC_Trends_in_Consumer_Mobility_Report.pdf
28. Steinle S, Reis S, Sabel CE. Quantifying human exposure to air pollution—moving from static monitoring to spatio-temporally resolved personal exposure assessment. Sci Total Environ. 2013;443: 184–93.
29. Glasgow ML, Rudra CB, Yoo E-H, Demirbas M, Merriman J, Nayak P, et al. Using smartphones to collect time–activity data for long-term personal-level air pollution exposure assessment. J Expo Sci Environ Epidemiol. 2016;26:356–64.
30. Hu M, Li W, Li L, Houston D, Wu J. Refining time-activity classification of human subjects using the global positioning system. PLoS One. 2016;11:e0148875.
31. González MC, Hidalgo CA, Barabási A-L. Understanding individual human mobility patterns. Nature. 2008;453:779–82.
32. Collins FS, Varmus H. A new initiative on precision medicine. N Engl J Med. 2015;372:793–5.
33. Chan Y-FY, Wang P, Rogers L, Tignor N, Zweig M, Hershman SG, et al. The Asthma Mobile Health Study, a large-scale clinical observational study using ResearchKit. Nat Biotechnol. 2017;35:354–62.
34. van Donkelaar A, Martin RV, Brauer M, Hsu NC, Kahn RA, Levy RC, et al. Global estimates of fine particulate matter using a combined geophysical-statistical method with information from satellites, models, and monitors. Environ Sci Technol. 2016;50:3762–72.
35. Larkin A, Geddes JA, Martin RV, Xiao Q, Liu Y, Marshall JD, et al. Global land use regression model for nitrogen dioxide air pollution. Environ Sci Technol. 2017;51:6957–64.
36. Lary DJ, Lary T, Sattler B. Using machine learning to estimate global PM2.5 for environmental health studies. Environ Health Insights. 2015;9:41–52.
37. Li X, Peng L, Hu Y, Shao J, Chi T. Deep learning architecture for air quality predictions. Environ Sci Pollut Res Int. 2016;23:22408–17.
38. Zhang C, Yan J, Li C, Rui X, Liu L, Bie R. On estimating air pollution from photos using convolutional neural network. Proc. 2016 ACM Multimed. Conf. [Internet]. New York, NY, USA: ACM; 2016 [cited 2017 Sep 8]. p. 297–301. Available from: https://doi.org/10.1145/2964284.2967230
39. Jean N, Burke M, Xie M, Davis WM, Lobell DB, Ermon S. Combining satellite imagery and machine learning to predict poverty. Science. 2016;353:790–4.
40. Apte JS, Messier KP, Gani S, Brauer M, Kirchstetter TW, Lunden MM, et al. High-resolution air pollution mapping with Google Street View cars: exploiting big data. Environ Sci Technol. 2017;51:6999–7008.
41.• Barrett MA, Humblet O, Hiatt RA, Adler NE. Big data and disease prevention: from quantified self to quantified communities. Big Data. 2013;1:168–75. **This commentary highlights the role of big data for quantifying communities and how this approach can facilitate public health**.
42. Albino V, Berardi U, Dangelico RM. Smart cities: definitions, dimensions, performance, and initiatives. J Urban Technol. 2015;22: 3–21.
43. Solanas A, Patsakis C, Conti M, Vlachos IS, Ramos V, Falcone F, et al. Smart health: a context-aware health paradigm within smart cities. IEEE Commun Mag. 2014;52:74–81.
44. 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. Atmos Environ. 2016;131:289–300.
45. Nyhan M, Grauwin S, Britter R, Misstear B, McNabola A, Laden F, et al. "Exposure track"—the impact of mobile-device-based mobility patterns on quantifying population exposure to air pollution. Environ Sci Technol. 2016;50:9671–81.
46. Clougherty JE, Levy JI, Kubzansky LD, Ryan PB, Suglia SF, Canner MJ, et al. Synergistic effects of traffic-related air pollution and exposure to violence on urban asthma etiology. Environ Health Perspect. 2007;115:1140–6.
47. Clougherty JE, Kubzansky LD. A framework for examining social stress and susceptibility to air pollution in respiratory health. Environ Health Perspect. 2009;117:1351.
48. Downey L, Van Willigen M. Environmental stressors: the mental health impacts of living near industrial activity. J Health Soc Behav. 2005;46:289–305.
49. Volckens J, Quinn C, Leith D, Mehaffy J, Henry CS, Miller-Lionberg D. Development and evaluation of an ultrasonic personal aerosol sampler. Indoor Air. 2017;27:409–16.
50. O'Connell SG, Kincl LD, Anderson KA. Silicone wristbands as personal passive samplers. Environ Sci Technol. 2014;48:3327–35.
51. Raghupathi W, Raghupathi V. Big data analytics in healthcare: promise and potential. Health Inf Sci Syst [Internet]. 2014 [cited 2017 Aug 4];2. Available from: http://www.ncbi.nlm.nih.gov/pmc/articles/PMC4341817/
52. Sandve GK, Nekrutenko A, Taylor J, Hovig E. Ten simple rules for reproducible computational research. PLoS Comput Biol. 2013;9: e1003285.
53. Arnett DK, Claas SA. Precision medicine, genomics, and public health. Diabetes Care. 2016;39:1870–3.
54. Petersen C, DeMuro P. Legal and regulatory considerations associated with use of patient-generated health data from social media and mobile health (mHealth) devices. Appl Clin Inform. 2015;6:16–26.
55. Vayena E, Salathé M, Madoff LC, Brownstein JS. Ethical challenges of big data in public health. PLoS Comput Biol. 2015;11: e1003904.
56. Clarke A, Steele R. Smartphone-based public health information systems: anonymity, privacy and intervention. J Assoc Inf Sci Technol. 2015;66:2596–608.
57. Cortez NG, Cohen IG, Kesselheim AS. FDA regulation of mobile health technologies. N Engl J Med. 2014;371:372–9.
58. Oscar N, Fox PA, Croucher R, Wernick R, Keune J, Hooker K. Machine learning, sentiment analysis, and tweets: an examination of Alzheimer's disease stigma on Twitter. J Gerontol B Psychol Sci Soc Sci [Internet]. 2017 [cited 2017 Apr 30]; Available from: https://www.ncbi.nlm.nih.gov/pubmed/28329835.
59. Cavazos-Rehg PA, Krauss MJ, Sowles SJ, Bierut LJ. "Hey everyone, I'm drunk." An evaluation of drinking-related Twitter chatter. J Stud Alcohol Drugs. 2015;76:635–43.